

České vysoké učení technické v Praze
Fakulta elektrotechnická
Otvorená informatika



Diplomová práce

Veľkoobjemové úložisko emailov

Bc. Patrik Lenárt

Vedúci práce: Ing. Ján Šedivý, CSc.

Študijný program: Otvorená informatika, Magisterský

Obor: Softwarové inžinierstvo

2. marca 2011

Pod'akovanie

Prehlásenie

Prehlasujem, že som svoju diplomovú prácu vypracoval samostatne a použil som iba podklady uvedené v priloženom zozname.

Nemám závažný dôvod proti užitiu tohto školského diela v zmysle §60 Zákona č. 121/2000 Sb., o autorskom práve, o právach súvisiacich s právom autorským a o zmene niektorých zákonov (autorský zákon).

V Prahe dňa 1.3.2011

.....

Abstract

Abstrakt

Obsah

1	Úvod	1
1.1	Osnova	1
2	Databázové systémy	3
2.1	História	3
2.2	ACID	4
2.3	Škálovanie databázového systému	4
2.3.1	Replikácia	5
2.3.2	Rozsekávanie dát (sharding)	6
2.4	BASE	7
2.5	CAP	8
2.5.1	Konzistencia verzus dostupnosť	9
2.6	Eventuálna konzistencia	10
2.6.1	Konzistencia z pohľadu klienta	10
2.6.2	Konzistencia na strane servera	11
2.7	MapReduce	12
2.7.1	Príklad	13
2.7.2	Architektúra	13
2.7.3	Použitie	13
2.8	Zhrnutie	13
3	NoSQL	15
4	Definícia problému	17
4.1	Archivácia elektronickej pošty	17
4.2	Požiadavky na systém	18
4.2.1	Obecne vlastnosti	18
5	test	21
5.1	Komponenty IEEE 802.15.4	21
5.2	Sieťová topológia	22
5.3	Architektúra	22
5.3.0.1	Typy rámcov	22
5.4	Architektúra	23
5.5	Sieťové komponenty	23
5.6	Sieťová topológia	24

5.7	XBee Series 1	24
5.7.1	Technické parametre	25
6	Teória antén	27
6.0.2	Definícia pojmov	27
6.0.3	Propagácia rádiového signálu	28
6.0.3.1	Pojmy	28
6.0.4	Typy antén	29
6.0.5	Stratovosť voľného priestoru (Free-space path loss)	29
6.0.6	Antény a simulácia	30
7	Analýza a návrh riešenia	33
7.0.6.1	Výkonová závislosť	33
7.0.7	Štruktúra MF	35
7.0.7.1	Štruktúra modulu NIC	36
8	Realizácia	41
8.0.8	Model kolízie	42
8.0.8.1	Popis kolízie	43
9	Testovanie	47
9.0.9	Testy zamerané na pohyb XBee zariadení	47
9.0.10	Testy s kolíziou	48
10	Záver	53
	Literatúra	55
A	Zoznam použitých skratiek	57
B	Inštalčná a užívateľská príručka	59
B.0.11	Inštalácia simulátoru OMNeT++ pre platformu Linux	59
B.0.12	Inštalácia mnou modifikovaného Mobility Frameworku	59
B.0.13	Práca s modelom IEEE 802.15.4	60
C	Obsah priloženého CD	61

Zoznam obrázkov

5.1	Topológia štandardu 802.15.4	22
5.2	Štruktúra štandardu 802.15.4	23
5.3	Štruktúra štandardu ZigBee	24
6.1	Model rádiového systému	28
6.2	Charakteristika zisku antény Whip	30
6.3	Model zoslabenia signálu	30
7.1	Popis výkonov a zisku u antén pre vysieláč a prijímač	35
7.2	Štruktúra stanice v OMNeT++	36
7.3	Štruktúra NIC v OMNeT++	38
7.4	Prechody medzi stavmi v module snrEval	39
8.1	Model kolízie	44
8.2	Kolízia na MAC vrstve	45
9.1	Pohyb na vzdialenosť 5m, vysielací výkon 1mW	49
9.2	Pohyb na vzdialenosť 5m, vysielací výkon 10mW	50
9.3	Pohyb na vzdialenosť 250m, vysielací výkon 10mW	50
9.4	Pohyb na vzdialenosť 250m, vysielací výkon 1mW, rotácia okolo vlastnej osi	51
C.1	Výpis priloženého CD	61

Zoznam tabuliek

5.1	Špecifikácia výkonu a rýchlosti	25
5.2	Obecné parametre	25
5.3	Sieťové parametre	25
9.1	Stratovosť rámcov	48
9.2	Prípad č. 1	48
9.3	Prípad č. 2	49

Kapitola 1

Úvod

S neustálym rozvojom informačných technológií súčasne narastá objem informácií, ktoré je potrebné spracúvať. Tento fakt podnietil vznik databázových systémov, ktoré slúžia na organizáciu, uchovávanie a prácu s veľkým objemom dát. V dnešnej dobe existuje množstvo databázových systémov, ktoré sa navzájom líšia svojou architektúrou, dátovým modelom, výrobcom atď.

Od začiatku sedemdesiatych rokov 20. storočia sú v tejto oblasti dominantou relačné databázové systémy (Relational Database Management Systems). Vďaka neustálemu prudkému rozvoju internetových technológií a rapídneho rastu dát v digitálnom univerze [7] začínajú byť tieto systémy nepostačujúce. Medzi hlavné faktory pre výber relačného databázového systému doposiaľ patrili výrobca, cena a pod. V dnešnej dobe so vznikom moderných aplikácií (napríklad sociálne siete, dátové sklady, analytické aplikácie a iné), požadujeme od týchto systémov vlastnosti ako vysoká dostupnosť, horizontálna rozšíriteľnosť a schopnosť pracovať s obrovským objemom dát (petabyte). Novo vznikajúce databázové systémy, spĺňajúce tieto požiadavky sa spoločne označujú pod názvom NoSQL (Not Only SQL). Pri ich výbere je v tomto prípade dôležité porozumenie architektúry, dátového modelu a dát, s ktorými budú tieto systémy pracovať.

Táto práca si kladie za cieľ viacero úloh, ktorými sú pochopenie a popis základných konceptov, ktoré tieto systémy využívajú, určenie kritérií vďaka ktorým môžeme tieto systémy navzájom porovnávať. Ďalej je úlohou analyzovať a popísať požiadavky pre systém veľkoobjemového úložiska elektronickej pošty, ktorý bude schopný spracovávať milióny emailov. Poslednou úlohou je na základe našich požiadaviek vybrať, čo najlepšie odpovedajúci NoSQL systém a s jeho použitím implementovať prototyp aplikácie.

1.1 Osnova

...

Kapitola 2

Databázové systémy

V tejto časti stručne popíšeme históriu vzniku databázových systémov, základné problémy pri tvorbe distribuovaných relačných databázových systémov a uvedieme možné spôsoby ich riešenia. Ďalej popíšeme základné koncepty, ktoré sa využívajú pri tvorbe distribuovaných databázových systémov a techniku MapReduce, ktorá slúži na prácu s veľkým objemom dát uloženým v systémoch NoSQL.

2.1 História

V polovici šesťdesiatych rokov 20. storočia bol spoločnosťou IBM vytvorený informačný systém IMS (Information Management System), využívajúci hierarchický databázový model. IMS je po rokoch vývoja využívaný dodnes. Po krátkej dobe, v roku 1970, publikoval zamestnanec IBM, Dr. Edger F. Codd článok pod názvom „A Relational Model of Data for Large Shared Data Banks“ [5], ktorým uviedol relačný databázový model. Prvým databázovým systémom, ktorý tento model implementoval bol System R od IBM. Tento systém používal jazyk pod názvom SEQUEL, ktorý je predchodca dnešného SQL (Structured Query Language) slúžiacého na manipuláciu a definíciu dát v relačných databázových systémoch. Tento koncept sa stal základom pre relačné databázové systémy, ktoré vďaka širokej škále vlastností ako napríklad podpora transakcií, dotazovací jazyk SQL, patria v dnešnej dobe medzi najpoužívanejšie riešenia na trhu.

V minulosti boli objem dát, s ktorým tieto systémy pracovali a výkon hardvéru mnohonásobne nižšie. Dnes napriek tomu, že výkon procesorov a veľkosť pamäťových zariadení rapídne stúpa, je najväčšou slabinou počítačových systémov rýchlosť prenosu dát medzi pevným diskom a hlavnou pamäťou. Ako príklad si vezmeme bežnú konfiguráciu počítačového systému, ktorá obsahuje pevný disk o veľkosti 2TB a operačnú pamäť veľkosti 64Gb. Napriek týmto vysokým kapacitám tento systém bohužiaľ dokáže v daný moment spracúvať maximálne 64Gb dát, čo je zlomok veľkosti v porovnaní s kapacitou pevného disku. Vznik nových webových aplikácií napr. sociálne siete, zavádzanie cloud computingu vyžadujú od systémov podporu škálovania, ktorá zabezpečuje vysokú dostupnosť, spoľahlivosť a ich nároky na spracovávaný objem dát sa neustále zvyšujú. Tieto nové požiadavky efektívne riešia distribuované systémy pod spoločným názvom NoSQL, ktoré popisuje nasledujúca kapitola.

2.2 ACID

Relačné databázové systémy poskytujú veľkú množinu operácií, ktoré sa vykonávajú nad ich dátami. Tranzakcie [7][8] sú zodpovedné za korektné vykonanie operácií v prípade, že spĺňajú množinu vlastností ACID. Význam jednotlivých vlastností akronýmu ACID je nasledovný:

- Atomicita (Atomicity) - zaisťuje, že sa daná tranzakcia vykoná celá, čo spôsobí korektný prechod systému do nového stavu. V prípade zlyhania tranzakcie nemá daná operácia žiaden vplyv na výsledný stav systému a prechod do nového stavu sa nevykoná.
- Konzistencia (Consistency) - každá tranzakcia po svojom úspešnom ukončení garantuje korektnosť svojho výsledku a zabezpečí, že systém prejde z jedného konzistentného stavu do druhého. Pojem konzistentný stav zaručuje, že dáta v systéme odpovedajú požadovanej hodnote. Systém sa musí nachádzať v konzistentnom stave aj v prípade zlyhania tranzakcie.
- Izolácia (Isolation) - operácie, ktoré prebiehajú počas vykonávania jednej tranzakcie nie sú viditeľné pre ostatné. Každá tranzakcia musí mať konzistentný prístup k dátam a to aj v prípade, že u inej tranzakcii dôjde k jej zlyhaniu.
- Trvácnosť (Durability) - v prípade, že bola tranzakcia úspešne ukončená, systém musí garantovať trvácnosť jej výsledku aj v prípade jeho zlyhania.

Implementácia vlastností ACID, ktoré zaručujú konzistenciu, zvyčajne využíva u relačných databázových systémoch metódu zamykania. Tranzakcia uzamkne dáta pred ich spracovaním a spôsobí ich nedostupnosť až do jej úspešného ukončenia, poprípade zlyhania. Pre databázový systém, od ktorého požadujeme vysokú dostupnosť alebo prácu pod zvýšenou záťažou, tento model nie je vyhovujúci. Zámky spôsobujú stavy, kedy ostatné operácie musia čakať na ich uvoľnenie. Jeho náhradou je Multiversion concurrency control, ktorý využívajú aj niektoré NoSQL databázové systémy.

Tranzakcie splňujúce vlastnosti ACID využívajú v distribuovaných databázových systémoch ¹ dvojfázový potvrdzovací protokol (two-phase commit protocol). Distribuovaný databázový systém využívajúci tento protokol, ktorého tranzakcie splňujú vlastnosť ACID zaručuje konzistentnosť a je schopný odolávať čiastočným poruchám na sieti alebo v prípade čiastočnej poruchy systému. Vlastnosti ACID nekladú žiadnu záruku na dostupnosť systému. Takéto systémy sú vhodné pre Internetové tranzakcie, aplikácie využívajúce platby apod. Existuje množstvo aplikácií, u ktorých má dostupnosť prednosť pred konzistenciou. Pri tvorbe distribuovaných databázových systémov je preto potrebné upustiť z niektorých ACID vlastností, čo spôsobilo vznik nového modelu pod názvom BASE.

2.3 Škálovanie databázového systému

Obecná definícia pojmu škálovateľnosť [2] je náročná bez vymedzenia kontextu, ku ktorému sa vzťahuje. V tejto práci budeme škálovateľnosť chápať v kontexte webových aplikácií,

¹Distribuované databázové systémy sú tvorené pomocou viacerých samostatne operujúcich databázových systémov, ktoré nazývame uzly a môžu komunikovať pomocou siete a užívateľovi alebo aplikácii sa javia ako jeden celok [ref].

ktorých dynamický vývoj kladie na databázové systémy viacero požiadavkov. Medzi hlavné patrí neustála potreba zvyšovania diskového priestoru a teda zvyšovanie veľkosti databázy alebo schopnosť obslúžiť čoraz vyšší počet užívateľov aplikácie (zvýšenie počtu operácií pre čítanie a zápis do databázového systému). V tomto prípade pod pojmom škálovateľnosť databázového systému rozumieme vlastnosť, vďaka ktorej je systém schopný spracúvať narastajúce požiadavky webovej aplikácie v definovanom čase intervalu. Typicky pridaním nových systémov, čo spôsobuje vznik distribuovaného databázového systému.

Škálovateľnosť delíme na vertikálnu a horizontálnu. Táto metóda dodáva systému nasledujúce vlastnosti [5]:

- umožňuje zväčšiť veľkosť celkovej kapacity databázy a táto zmena by mala byť transparentná z pohľadu aplikácie na dáta.
- zvyšuje celkové množstvo operácií, pre čítanie a zápis dát, ktoré je systém schopný vykonať v danú časovú jednotku.
- v niektorých prípadoch môže zaručiť, že systém neobsahuje jednotku, ktorá by v prípade zlyhania spôsobila nedostupnosť celého systému (single point of failure).

Vertikálna škálovateľnosť je metóda, ktorá sa aplikuje pomocou zvyšovania výkonnosti hardvéru, tj. do systému sa pridáva operačná pamäť, rýchlejšie viacjadrové procesory, zvyšuje sa kapacita diskov. Jednou z nevýhod tohoto riešenia je jeho vysoká cena a možná chvíľková nedostupnosť systému. Proces vertikálneho škálovania nad relačnou databázou obsahuje nasledujúce kroky:

- zámena hardvéru za výkonnejší
- úprava súborového systému (napr. zrušenie žurnálu)
- optimalizácia databázových dotazov, indexovanie
- pridanie vrstvy pre kešovanie (memcached, EHCACHE, atď.)
- denormalizácia dát v databáze, porušenie normalizácie

V tomto prípade je možné naraziť na hranice Moorovho zákona [6] a na rad nastupuje horizontálna škálovateľnosť, ktorá je omnoho komplexnejšia. Horizontálnu škálovateľnosť je možné realizovať pomocou replikácie alebo metódou rozsekávania dát (sharding).

2.3.1 Replikácia

V distribuovaných systémoch sa pod pojmom replikácia rozumie vlastnosť, ktorá má za následok že sa daná informácia nachádza v konzistentnom stave na viacerých uzloch² tohoto systému. Táto vlastnosť zvyšuje dostupnosť, spoľahlivosť a odolnosť systému voči chybám.

V prípade distribuovaného databázového systému sa časť informácií uložených v databáze nachádza na viacerých uzloch. Táto vlastnosť môže napríklad zvýšiť výkonnosť operácií, ktoré

²Pod pojmom uzol v tomto prípade myslíme samostatný počítačový systém, ktorý je súčasťou distribuovaného systému

pristupujú k dátam a to tak, že dochádza k čítaniu dát z databázy paralelne z viacerých uzlov. V systéme obsahujúcom repliku dát nedochádza k strate informácií v prípade poruchy uzlu. Replikácia a propagácia zmien v systéme sú z pohľadu aplikácie transparentné. Metóda replikácie nezvyšuje pridávaním nových uzlov celkovú kapacitu databázy. Problémom tejto techniky je zápis dát, pri ktorom sa zmena musí prejaviť vo všetkých replikách. Existuje viacero metód pomocou, ktorých je možné zabezpečiť túto funkcionálnosť:

- Read one - write all, u tejto metódy sa čítanie dát prevedie z ľubovoľného uzlu obsahujúceho repliku. Zápis dát sa vykoná na všetky uzly obsahujúce repliku a v prípade, že každý z nich potvrdí úspech tejto operácie, zmena sa považuje za úspešnú. Táto metóda nie je schopná pracovať v prípade, že dôjde k prerušeniu sieťového toku medzi uzlami (network partitioning) alebo v prípade poruchy uzlu.
- Quorum consensus - zápis na jeden uzol a následná asynchrónna propagácia repliky na ostatné uzly. Táto metóda je schopná zvládnuť stav pri ktorom dôjde k prerušeniu sieťového toku alebo poruche uzlu. Implementácie využívajú algoritmy pod názvom kôrum konsenzus (quorum consensus). ???

Výber metódy replikácie čiastočne popisuje dvojicu vlastností distribuovaného databázového systému a to dostupnosť a konzistenciu. Podľa teórie s názvom CAP (viď. nasledujúca kapitola) nie je možné aby systém disponoval súčasne vlastnosťami ako dostupnosť, konzistencia dát a schopnosť odolávať poruchám v prípade chyby v sieťovej komunikácii.

V relačných databázových systémoch sa replikácia rieši pomocou techniky Master-Slave. Uzol pod názvom master slúži ako jediný databázový stroj, na ktorom sa vykonáva zápis dát a replika týchto dát je následne distribuovaná na zvyšné uzly pod názvom slave. Touto metódou sme schopný mnohonásobne zvýšiť počet operácií, ktoré slúžia pre čítanie dát z databázového systému a v prípade zlyhania niektorého zo systémov máme neustále k dispozícii kópiu dát. Slabinou tejto techniky je uzol v roli master, ktorý znižuje výkonnosť v prípade operácií vykonávajúcich zápis a zároveň môže jeho porucha spôsobiť celkovú nedostupnosť systému.

Druhým riešením je technika Multi-master, kde každý uzol obsahujúci repliku je schopný zápisu dát a následne tieto preposiela zmeny ostatným. Tento mechanizmus predpokladá distribuovanú správu zamykania a vyžaduje algoritmy pre riešenie konfliktov spôsobujúcich nekonzistenciu dát.

2.3.2 Rozsekávanie dát (sharding)

Rozsekávanie dát je metóda založená na princípe, kde dáta obsiahnuté v databáze rozdelíme podľa stanovených pravidiel do menších celkov. Tieto celky môžeme následne umiestniť na navzájom rôzne uzly distribuovaného databázového systému. Táto metóda umožňuje zvýšiť výkonnosť operácií pre zápis a čítanie dát a zároveň pridávaním nových uzlov do systému sme schopný zvyšovať celkovú kapacitu databázy. V prípade, že architektúra distribuovaného databázového systému využíva túto techniku, zvýšenie výkonu jeho operácií a objemu uložených dát sa realizuje automaticky bez nutnosti zásahu do aplikácie.

Techniku rozsekávania môžeme považovať za architektúru známu pod názvom zdieľanie ničoho [8] (shared nothing). Táto architektúra sa používa pre návrh systémov využívajúcich multiprocesory. V takomto prípade sa medzi procesormi nezdieľa operačná ani disková

pamäť. Táto architektúra zabezpečuje takmer neobmedzenú škálovateľnosť systému a využíva ju mnoho NoSQL systémov ako napríklad Google Bigtable, Amazon Dynamo alebo technológia MapReduce.

Pri návrhu distribuovaných databázových systémov, s využitím tejto techniky, patrí medzi kľúčový problém implementácia funkcie spojenia (join) nad dátami, ktorá sa radšej neimplementuje. V prípade, že sa dáta nad ktorými by sme chceli túto operáciu vykonať, nachádzajú na dvoch rozdielnych uzloch prepojených sieťou, takéto spojenie by značne znížilo celkovú výkonnosť systému a viedlo by k zvýšeniu sieťového toku a záťaži systémových zdrojov.

Keďže sa dáta nachádzajú na viacerých uzloch systému, hrozí zvýšená pravdepodobnosť hardverového zlyhania, poprípade prerušenie sieťového spojenia a preto sa táto technika často kombinuje s pomocou využitia replikácie.

V prípade použitia tejto techniky v relačných databázach, je nutný zásah do logiky aplikácie. Dáta uložené v tabuľkách relačnej databáze zachytávajú vzájomné relácie. Týmto spôsobom dochádza k celkovému narušeniu tohoto konceptu. Príkladom môže byť tabuľka obsahujúca zoznam zamestnancov, ktorú rozdelíme na samostatné celky. Každá tabuľka bude reprezentovať mená zamestnancov, ktorých priezvisko začína rovnakým písmenom abecedy a zároveň sa bude nachádzať na samostatnom databázovom systéme. Táto technika so sebou prináša problém, v ktorom je potrebné nájsť vhodný kľúč podľa, ktorého budeme dáta rozsekať a zabezpečíme tak rovnomerné zaťaženie uzlov daného systému. Existuje viacero metód [6] a to:

- segmentácia dát podľa funkcionality - dáta, ktoré sme schopný popísať spoločnou vlastnosťou ukladáme do samostatných databáz a tieto umiestňujeme na rozdielne uzly systému. Príkladom môže byť samostatný uzol spravujúci databázu pre užívateľov a iný uzol s databázou pre produkty. Túto metódu spracoval Randy Shoup³ [10], architekt spoločnosti eBay.
- rozsekanie podľa kľúča - v dátach hľadáme kľúč, podľa ktorého sme schopný ich rovnomernej distribúcie. Následne na tento kľúč aplikujeme hašovaciu funkciu a na základe je výsledku tieto dáta umiestňujeme na jednotlivé uzly.
- vyhľadávacia tabuľka - jeden uzol v systéme slúži ako katalóg, ktorý určuje na ktorom uzle sa nachádzajú dané dáta. Tento uzol zároveň spôsobuje zníženie výkonu a v prípade jeho havárie spôsobuje nedostupnosť celého systému.

Replikácia a rozsekanie dát patria medzi kľúčové vlastnosti využívané v NoSQL systémoch.

2.4 BASE

Akroným BASE [1] bol prvýkrát použitý v roku 1997 na sympóziu SOSP (ACM Symposium on Operating Systems Principles). BASE tvoria nasledujúce slovné spojenia:

³“If you can't split, you can't scale it.” – Randy Shoup, Distinguished architect Ebay

- bežne dostupný (Basically Available) - systém je schopný zvládať jeho čiastočné zlyhanie za cenu nižšej complexity.
- zmiernený stav (Soft State) - systém nezaručuje trvácnosť dát za cieľom zvýšenia výkonu.
- čiastočná konzistencia (Eventually Consistent) - je možné na určitú dobu tolerovať nekorektnosť dát, ktoré musia byť po určitom časovom intervale konzistentné.

Tento model poľavil na požiadavku zodpovednom za konzistenciu dát, s tým že dosiahol vyššiu dostupnosť v distribuovanom databázovom systéme aj v prípade čiastočného zlyhania. Prakticky môžeme každý systém klasifikovať ako systém spĺňajúci vlastnosti ACID alebo BASE.

BASE umožňuje horizontálne škálovanie relačných databázových systémov bez nutnosti použitia distribuovaných transakcií. Použitie tejto metódy je možné vďaka rozsekávaniu dát s využitím metódy segmentácie dát podľa funkcionality, ďalej sa využívajú perzistentné fronty a princípy udalosťami riadenej architektúry (event driven architecture) [9]. Poľavením na požiadavku konzistencie dát sa v tomto prípade myslí stav, že dáta budú konzistentné po uplynutí určitého časového intervalu.

Systém obsahujúci čiastočnú konzistenciu dát je bankomatový systém. Po vybratí určitej čiastky z účtu, sa korektná informácia o aktuálnom zostatku účtu môže zobraziť až za niekoľko dní, kdežto transakcia ktorá túto zmenu vykonala musí spĺňať vlastnosti ACID. Medzi podobné webové aplikácie, u ktorých sa nepožadujú všetky vlastnosti ACID patria nákupný košík spoločnosti Amazon, zobrazovanie časovej osi aplikácií Twitter poprípade indexy Googlu. Ich nedostupnosť by znamenala obrovské finančné straty, napríklad v prípade, zlyhanie vyhľadávania pomocou Googlu by znamenalo zobrazenie nižšieho počtu reklám, nedostupnosť nákupného košíka v Amazone by spôsobila pokles predaja atď.

Aplikácia vyššie popísaných techník na relačné databázové systémy môže byť netrivialnou úlohou. Relačný model, je spôsob reprezentácie dát, ktorý umožňuje efektívne riešiť určité typy problémov, preto snaha prispôbiť tento model každému problému je nezmyselná. V tomto prípade, musíme uvažovať alternatívne riešenia, medzi ktoré patria systémy NoSQL.

2.5 CAP

Moderné webové aplikácie kladú na systémy požiadavky, medzi ktoré patrí vysoká dostupnosť, konzistencia dát a schopnosť odolávať chybám ⁴. Dr. Brewer v roku 2000 nastolil myšlienku dnes známu pod názvom teória CAP [3], ktorá tvrdí že je možné súčasne dosiahnuť len dvojicu z týchto vlastností. V roku 2002 platnosť tejto teórie pre asynchrónnu sieť matematicky dokázali Lynch a Gilbert [?]. Modelu asynchrónnej siete odpovedá svojimi vlastnosťami sieť Internet. Akroným CAP tvoria nasledujúce vlastnosti:

- Konzistencia (Consistency) - distribuovaný systém je v konzistentnom stave ak sa zmena aplikuje na všetky relevantné uzly systému v rovnakom čase.

⁴Chybou v tomto kontexte myslíme, prerušenie sieťovej komunikácie medzi uzlami daného systému, poprípade hardverová porucha uzlu???

- Dostupnosť (Availability) - distribuovaný systém je dostupný ak každý jeho uzol, ktorý pracuje korektne je schopný pri prijatí požiadavku zaslať odpoveď. V spojení s toleranciou chýb, tato vlastnosť hovorí, že prípade ak nastane sieťový problém ⁵ každý požiadavok bude vykonaný.
- Tolerancia chýb (Partition Tolerance) - uzly distribuovaného systému navzájom komunikujú pomocou siete, v ktorej hrozí stráta správ. V prípade vzniku sieťovej poruchy dané uzly medzi sebou navzájom nedokážu komunikovať. Táto vlastnosť podľa definície vid'. Gilbert a Lynch tvrdí, že v prípade vzniku zlyhania sieťovej komunikácie medzi niektorými uzlami musí byť systém schopný naďalej pracovať korektne. Neexistuje reálny distribuovaný systém, ktorého uzly na vzájomnú komunikáciu využívajú sieť a nedochádza pri tom k strate správ, teda k poruchám sieťovej komunikácie.

Pravdepodobnosť, že dôjde k zlyhaniu ľubovoľného uzlu v distribuovanom systéme exponenciálne narastá s počtom pribúdajúcich uzlov.

$$P(\text{ubovonhozlyhania}) = 1 - P(\text{individuálny uzol nezlyh})^{počet uzlov}$$

2.5.1 Konzistencia verus dostupnosť

V distribuovanom systéme nie je možné zaručiť súčasne vlasnosť konzistencie a dostupnosti. Ako príklad si predstavme distribuovaný systém, ktorý zaručuje obe vlasnosti aj v prípade sieťového prerušenia. Tento systém obsahuje tri uzly A,B,C, na ktorých sa nachádzajú identické (replikované) dáta. Ďalej uvažujme, že došlo k sieťovému prerušeniu, ktoré rozdelilo uzly na dva samostatné celky A,B a C. V prípade, že uzol C obdrží požiadavok pre zmenu dát má na výber dve možnosti:

1. vykonať zmenu dát, v tomto prípade sa uzly A a B o tejto zmene dozvedia až v prípade, že bude sieťové prerušenie odstránené
2. zamietnuť požiadavok na zmenu dát, z dôvodu že uzly A a B sa o tejto zmene nedozvedia až do jej odstránenia

V prípade výberu možnosti číslo 1 zabezpečíme dostupnosť systému naopak v prípade možnosti číslo 2 jeho konzistenciu, avšak nie je možný súčasný výber oboch riešení.

Ak od daného systému tolerujúceho sieťové prerušenia požadujeme konzistenciu na úkor dostupnosti jedná sa o alternatívu CP. Takýto systém zabezpečí atomickosť operácií ako zápis a čítanie a zároveň sa môže stať, že na určité požiadavky nebude schopný odpovedať. Medzi takéto systémy môžeme zaradiť distribuovaný databázový systém využívajúci dvojfázový potvrdzovací protokol (2PC).

V prípade, že poľavíme na požiadavku konzistencie tak takýto systém bude vždy dostupný aj napriek sieťovým prerušeniam. V tomto prípade sa jedná o model AP. Je možné, že v takomto systéme bude dochádzať ku konfliktným zápisom alebo operácie čítania budú po určitú dobu vracať nekorektné výsledky. Tieto problémy s konzistenciou sa v distribuovaných databázových systémoch riešia napríklad pomocou metódy vektorový časovač (Vector

⁵týmto sa nemyslí porucha uzlu

clock) alebo na aplikačnej úrovni. Príkladom systému patriaceho do tejto kategórie je služba DNS.

V prípade, že systém nebude schopný zvládať sieťové prerušenia tak takýto systém bude spĺňať požiadavok konzistencie a dostupnosti, varianta CA. Jedná sa o nedistribuované systémy pracujúce na jednom fyzickom hardvéri využívajúce tranzakcie.

Vyššie popísané vlastnosti nám umožnia vhodný výber distribuovaného databázového systému podľa požiadavkov našej aplikácie.

2.6 Eventuálna konzistencia

V ideálnom svete je predstava konzistencie v distribuovaných systémoch nasledovná: v prípade, že sa v systéme vykoná zmena (zápis dát), na všetkých uzloch sa táto zmena prejaví súčasne s rovnakým výsledkom. Konzistencia v distribuovanom databázovom systéme je úzko spojená s replikáciou. Keďže podľa CAP teórie distribuovaný systém nemôže súčasne spĺňať požiadavok dostupnosti, konzistencie v prostredí s možným sieťovým prerušením, je na našom zväžení ktorú z týchto vlastností uprednostníme pri návrhu a tvorbe aplikácií. Väčšina NoSQL systémov poskytuje eventuálnu konzistenciu. V nasledujúcej časti preto popíšeme rôzne typy konzistencie.

V predchádzajúcom texte sme už spomínali, že v dnešnej dobe existuje mnoho aplikácií, u ktorých je možné poľaviť na požiadavku konzistencie a funkčnosť systému nebude v tomto prípade ohrozená, ak sa určitá zmena prejaví s miernym oneskorením. Takáto konzistencia je odlišná od definície vlastností ACID, kde ukončenie tranzakcie spôsobí, že systém sa nachádza v konzistentom stave. Na konzistenciu sa môžeme pozeráť z dvoch pohľadov. Prvým je klientský pohľad na strane zadávateľa problému resp. programátora, ktorý rozhodne aká je závažnosť zmien, ktoré sa budú vykonávať v systéme. Druhý pohľad je serverový, ktorý zabezpečuje technické riešenie a implementáciu techník zodpovedných za konzistenciu v distribuovaných databázových systémoch.

2.6.1 Konzistencia z pohľadu klienta

Pre potrebu nasledujúcich definícií uvažujme distribuovaný databázový systém, ktorý tvorí úložisko dát a tri nezávislé procesy A,B,C, ktoré môžu v danom systéme zmeniť hodnotu dátovej jednotky, tj. vykonať zápis. Tieto procesy môžu zároveň zo systému hodnotu dátovej jednotky prečítať. Na základe toho ako dané procesy pozorujú nezávisle zmeny systému delíme konzistenciu [?] na:

Silná konzistencia (Strong consistency) - proces A vykoná zápis. Po jeho ukončení je nová hodnota dátovej jednotky dostupná všetkým procesom A, B, C, ktoré k nej následne pristúpia - vykonajú operáciu čítania. Túto konzistenciu zabezpečujú tranzakcie s vlastnosťami ACID.

Slabá konzistencia (Weak consistency) - proces A vykoná zápis novej hodnoty do dátovej jednotky. V takomto prípade systém negarantuje, že následne pristupujúce procesy A, B, C k tejto jednotke vrátia jej novú hodnotu. Definujeme pojem „nekonzistentné okno“ zabezpečujúci, že po uplynutí určitej doby sa táto nová hodnota dátovej jednotky prejaví vo všetkých procesoch, ktoré k nej pristúpia.

Eventuálna konzistencia (Eventual consistency) - je to špecifická forma slabej konzistencie. V tomto prípade systém garantuje, že ak sa nevykoná žiadna nová zmena hodnoty dátovej jednotky, po určitom čase budu všetky procesy prístupujúce k tejto jednotke schopné vrátiť jej korektnú hodnotu. Tento model má viacero variácií, niektoré z nich popíšeme v nasledujúcej časti textu.

Read-your-write consistency - v prípade, že proces A zapíše novú hodnotu do dátovej jednotky, žiadny z jeho následujúcich prístupov k tejto jednotke nevráti staršiu hodnotu ako jeho posledný zápis.

Session consistency - v tomto prípade prístupuje proces k systému v kontexte relácií. Po dobu trvania relácie platí predchádzajúci typ konzistencie. V prípade zlyhania relácie sa vytvorí nová, v ktorej môže systém vrátiť hodnotu dátovej jednotky, zapísanú pred vznikom predchádzajúcej relácie.

Monotonic read consistency - v prípade, že proces vrátil určitú hodnotu dátovej jednotky, tak v každom ďalšom prístupe, nemôže nastať situácia, kde by vrátil predchádzajúcu hodnotu dátovej jednotky.

Tieto typy konzistencie je možné navzájom kombinovať a ich hlavným cieľom je zvýšiť dostupnosť distribuovaného systému na úkor toho, že poľavíme na požiadavkách konzistencie. Príkladom môže byť asynchrónna replikácia v modernom relačnom databázovom systéme, ktorá spôsobí, že systém bude eventuálne konzistentný.

2.6.2 Konzistencia na strane servera

Kôrum je minimálny počet hlasov, ktorý musí obdržať distribuovaná tranzakcia aby mohla následne vykonať operáciu v distribuovanom systéme. Technika založená na protokoloch kvôra (quorum-based protocols) je používaná na vykonávanie konzistentných operácií v distribuovaných databázových systémoch.

Definujme nasledujúcu terminológiu:

- N - počet uzlov, ktoré obsahujú repliku dát
- W - počet uzlov obsahujúcich repliku, na ktorých sa musí vykonať zápis, aby bola zmena úspešne potvrdená
- R - počet uzlov s replikou, ktoré musia vrátiť hodnotu dátového objektu v prípade operácie čítanie

Rôzna konfigurácia týchto parametrov zabezpečí rozdielnu výkonnosť a dostupnosť distribuovaného systému. Uvažujme nasledujúce príklady, kde $N = 3$.

1. $R = 1$ a $W = N$, v takomto prípade zabezpečíme že systém bude optimalizovaný pre operácie čítania dát. Operácie budú konzistentné, pretože uzol z ktorého dáta čítame sa prekrýva s uzlami na ktorých vykonávame zápis. Nevýhodou tohoto modelu je, že v prípade nedostupnosti všetkých replík nebude možné do systému zapisovať. V prípade systémov, kde vyžadujeme rýchle čítanie a na systém je obrovský počet požiadavok čítania sa môže hodnota N pohybovať v stovkách až tisícoch, závisí to od počtu uzlov v systéme.

2. $W = 1$ a $R = N$, tento prípad je vhodný pre systémy u ktorých požadujeme rýchly zápis. Tento model môže spôsobiť stratu dát v prípade, že systém s replikou na ktorú sa vykoná zápis zlyhá.
3. $W + R \leq N$, tento model spôsobí, že uzly, na ktoré sa vykonáva zápis a čítanie sa neprekrývajú, z čoho vyplýva u distribuovaného databázového systému vlastnosť eventuálnej konzistencie.

Nekonzistentnosť dát, môže byť tolerovaná v distribuovaných systémoch, ktoré sú vysoko škálovateľné za cieľom dosiahnutia lepšieho výkonu operácií, ktoré slúžia pre zápis a čítanie dát, celkovej výkonnosti a dostupnosti systému. Hranica do akej miery je možné dovoliť nekonzistenciu je určená požiadavkom klientskej aplikácie a vyššie spomínané modely sa ju snažia riešiť. Väčšinu z týchto modelov implementujú NoSQL systémy.

2.7 MapReduce

Nárast diskových kapacít a množstva dát, ktoré na nich ukladáme spôsobuje jeden z ďalších problémov, ktorým je analýza a spracovanie dát. Kapacita pevných diskov sa za posledné roky mnohonásobne zvýšila v porovnaní s dobou prístupu a prenosových rýchlostí pre čítanie a zápis dát na tieto zariadenia.

Pre jednoduchosť uvažujme nasledujúci príklad, v ktorom chceme spracovať pomocou jedného počítačového systému 1TB dát uložených na lokálnom súborovom systéme, pri priemernej prenosovej rýchlosti diskových zariadení 100Mb/s. Za ideálnych podmienok by čas na prečítanie týchto dát presahoval dve a pol hodiny. Tento čas je z praktických dôvodov neprípustný. V prípade, že by sme tento 1TB dát vhodne rozdelili na sto počítačov a na každom z nich tento úsek spracovali, celková doba spracovania by sa znížila za ideálnych podmienok na necelé tri minúty. Spoločnosť Google v roku 2004 zverejnila programovací model pod názvom MapReduce [?], ktorý rieši tento problém pomocou paralelizácie výpočtu.

MapReduce je programovací model, ktorý slúži na paralelne spracovanie dát (PB). Model využíva vlastnosti paralelných a distribuovaných systémov, je optimalizovaný pre beh na klastrí, tvorenom vysokým počtom (tisíce) spotrebných počítačov. Jeho cieľom je pre programátora zastrieť všetky problémy, ktoré spôsobuje paralelizácia, poruchovosť systémov, distribúcia dát vzhľadom na ich lokalitu a rovnomerne rozvňovanie záťaže medzi systémami. Poskytuje rozhranie pre automatickú paralelizáciu a rozsiahly distribuovaný výpočet.

Pre použitie tohoto nástroja musí programátor zadať dve funkcie pod názvom map a reduce. Funkcia map na jednotlivých uzloch systému, transformuje vstupné data na základe zadefinovaného kľúča (k_1) na medzivýsledok, ktorý obsahuje nové kľúče (g_1, \dots, g_n) a k nim odpovedajúce hodnoty. Tieto hodnoty sa zoradia podľa ich príslušnosti ku kľúčom (g_1, \dots, g_n), následne sa odošlú na uzly s funkciou reduce. Užívateľom definovaná funkcia reduce nad hodnotami priradenými pre kľúč (g_1, \dots, g_n) prevedie operáciu, ktorej typickým výsledkom je jedna výsledná hodnota.

```
map(kľúč k1, hodnota) -> list(kľúč(z g1,...,gn), hodnota2)
reduce(kľúč(z g1,...,gn), list(hodnota2)) -> list(hodnota3)
```

2.7.1 Príklad

? histogram slov

2.7.2 Architektúra

Obrázok

2.7.3 Použitie

MapReduce nie je vhodný na spracovanie dát v reálnom čase, online spracovanie dát, ktoré sú citlivé na latenciu a to z dôvodu jeho optimalizácie pre dávkový beh.

Výhodou tohoto modelu je, že dokáže pracovať s neštruktúrovanými dátami. Jeho implementácia spoločnosťou Google, ktorá zároveň využíva distribuovaný súborový systém GFS nie je verejná. V rámci hnutia NoSQL vzniklo open-source riešenie pod názvom Hadoop, ktoré implementuje tento model na vlastnom distribuovanom súborovom systéme HDFS. Zároveň vznikli frameworky ako HIVE alebo PIG, ktoré sú nadstavbou modelu MapReduce, majú syntax podobnú jazyku SQL a využívajú ich NoSql databázové systémy pre spracovanie dát.

2.8 Zhrnutie

Cieľom tejto časti bolo pochopiť základné princípy, ktoré platia v distribuovaných databázových systémoch a sú súčasťou systémov NoSQL. Taktiež sme identifikovali problémy, ktoré môžu nastať v prípade tvorby distribuovaných relačných databázových systémov a stručne popísali možnosti ich riešení.

Kapitola 3

NoSQL

Názov NoSQL bol prvýkrát použitý v roku 1998 ako názov relačnej databázy, ktorej implementácia bola prevažne v interpretovaných programovacích jazykoch a neobsahovala jazyk SQL. V druhej polovici roku 2009 [?] sa názov NoSQL začal používať v spojení s databázovými systémami, ktoré nepoužívajú SQL dotazovací jazyk a tradičný relačný model, sú schopné horizontálneho škálovania pracujú na bežných spotrebných počítačoch, vyznačujú sa vysokou dostupnosťou a používajú jednoduchú schému.

Novo vznikajúce webové aplikácie ako napríklad sociálne siete spracúvajú čoraz väčší objem dát, musia byť schopné v daný moment obslúžiť čoraz väčší počet užívateľov a neustále dostupné. Pôvodným cieľom hnutia NoSql, bolo vytvoriť koncept, pre tvorbu moderných databáz, ktoré by boli schopné riešiť tieto nové požiadavky. Idea týchto systémov je založená na filozofii, ktorá tvrdí, že nemá zmysel sa za každú cenu snažiť prispôbiť dáta modelu relačnej databáze. Cieľom je vybrať systém, ktorý bude čo najvhodnejšie opovedať požiadavkom na uloženie a spracovanie našich dát. NoSQL obecné nepopisuje, žiaden konkrétny databázový systém, namiesto toho je to obecný názov pre nerelačné (non-relational) databázové systémy, ktoré majú odlišné vlastnosti a umožňujú prácu s rôznymi dátovými modelmi. Medzi ich ďalšie znaky patrí slabá konzistencia, možnosť spracúvať obrovské objemy dát (PB), jednoduché API a možnosť asynchrónneho zápisu dát. Tieto systémy nepodporujú operáciu databázového spojenia z dôvodu, že znižuje výkonnosť a zvyšuje zaťaženie siete (v prípade, že by sa táto operácia mala vykonať nad dátami, ktoré sa nachádzajú na rôznych uzloch). Pre tieto databázové systémy ďalej platí, že sú distribuované, podporujú automatickú replikáciu a rozsekávanie dát. Keďže sa jedná o relatívne mladé systémy, jedným z ich nedostatkov je malá podpora frameworkov, neustále sa meniace API a taktiež u mnohých chýbajúce rôzne grafické utility pre ich správu a monitoring. Medzi dátami, ktoré do nich ukladáme, je možné vytvárať vzájomné závislosti až na aplikačnej vrstve. Cieľom tohoto konceptu je riešiť spomínané novo vznikajúce problémy a zároveň koexistovať s relačnými databázovými systémami.

Kapitola 4

Definícia problému

Množstvo digitálnych informácií, každým rokom prudko narastá. Podľa štatistík spoločnosti IDC [7] v roku 2006 dosahovala kapacita digitálneho univerza veľkosť 161 exabytov. Podiel elektronickej pošty bez spamu, tvoril 3% z tohoto objemu. Podľa odhadov na rok 2010 mala kapacita digitálneho univerza dosiahnuť veľkosť 988 EB (exabyte), čo je takmer šesťnásobok nárastu pôvodnej kapacity v období štyroch rokov. V rozmedzí rokov 1998 až 2006 sa mal počet schránok elektronickej pošty zvýšiť z 253 miliónov na 1.6 miliardy. Predpoveď IDC ďalej uvádzala, že po ukončení roku 2010 tento počet presiahne hodnotu dvoch miliard. Počas obdobia medzi rokmi 1998 až 2006 celkový počet odoslaných správ elektronickej pošty rástol trikrát rýchlejšie ako počet jej užívateľov, dôvodom tohoto prudkého nárastu bola nevyžiadaná elektronická pošta, nazývaná spam. Predpokladá sa, že až 85% dát z celkového odhadovaného objemu 988 EB budú spracovávať, prenášať alebo zabezpečovať organizácie. Napriek tejto explózii digitálnych informácií je potrebné správne porozumieť hodnote týchto dát, nájsť vhodné metódy pre ich ukladanie do pamäti počítačových systémov, ich archiváciu a to tak aby sme ich mohli ďalej spracúvať a efektívne využiť. Táto časť práce si preto kladie za cieľ pochopiť štruktúru dát, ktoré reprezentujú elektronickú poštu tj. emaily a následne pomocou využitia vhodných databázových technológií popísať a navrhnúť model systému pre ich archiváciu.

4.1 Archivácia elektronickej pošty

S neustálym nárastom informácií obsiahnutých v digitálnom univerze, sa zároveň zvyšuje objem dát, ktorý reprezentuje elektronickú poštu. Je preto potrebné porozumieť štruktúre emailových správ a následne ich vhodne spracovať. Tieto dáta je potrebné uložiť tak aby sme dosiahli úsporu diskového priestoru, boli sme nad nimi schopný vykonávať operácie ako fulltextové vyhľadávanie, zber údajov pre tvorbu štatistík alebo ich opätovné sprístupnenie. Emaily obsahujú čoraz viac podnikových obchodných infomácií a iný dôležitý obsah, z tohto dôvodu musia byť organizácie všetkých rozmerov schopné uchovávať tento obsah pomocou vhodných archivačných nástrojov. S problémom archivácie zároveň úzko súvisí problém bezpečnosti. Pod pojmom bezpečnosti v tejto oblasti máme na mysli hlavne ochranu proti nevyžiadanej pošte tj. spamu, spyware, malware a phishingu. Na boj proti týmto hrozbám využívajú organizácie anti-spamové a anti-vírusové systémy. Možné dôvody prečo archivovať elektronickú poštu sú následovne: ??? strucne popisat jednotlivé pojmy

- súkromie dát
- regulatory compliance
- e-discovery
- knowledge management
- self-service for end users
- disaster recovery

4.2 Požiadavky na systém

V nasledujúcej časti popíšeme vlastnosti systému, ktorý bude slúžiť na archiváciu veľkého objemu emailových správ.

4.2.1 Obecné vlastnosti

Primárnym požiadavkom na systém je jeho rozširiteľnosť, dostupnosť a nízkonakladová administrácia. Predpokladané množstvo uložených dát v tomto systéme bude dosahovať desiatky až stovky TB (terabyte). Takúto kapacitu dát nie je možné uchovať na bežne prístupnom spotrebnom hardvéri (commodity hardware). Systém musí byť neustále dostupný a dáta zalohované v prípade vzniku havárie niektorej z jeho častí. Nad uloženými dátami, je potrebné vykonávať výpočtovo náročné operácie napr. generovanie štatistík. Tieto požiadavky prirodzene implikujú využitie distribuovaného databázového systému. Medzi hlavných kandidátov, vďaka ktorým sme tieto požiadavky schopný vyriešiť patria NoSQL databázové systémy, ktorých zrovnanie popíšeme v nasledujúcej kapitole.

4.2.1.1 Rozširiteľnosť

Predpokladáme použitie bežne dostupného spotrebného hardvéru, namiesto superpočítačov. Systém musí byť schopný horizontálneho škálovania, ktoré využijeme pre zvýšenie celkovej kapacity dátového úložiska systému a na prípadné efektívnejšie spracovanie dát pomocou techniky MapReduce. U databázového systému je nutná podpora replikácie, ktorá zvýši výkonnosť operácií pre čítanie a zápis do systému a vďaka nej nebude nutné riešiť zálohovanie systému pomocou externého riešenia.

4.2.1.2 Dostupnosť

Systém musí byť neustále dostupný a schopný odolávať poruchám v sieťovej komunikácii (ang. network partitions), preto je vhodné aby sa dáta replikovali vo vnútri datacentra a následne aspoň medzi dvoma datacentrami, ktoré budú umiestnené na geograficky odlišných miestach.

4.2.1.3 Nízkonákladová administrácia

Bolo by vhodné ak by systém neobsahoval bod, ktorého porucha by spôsobila celkovú nedostupnosť systému (ang. single point of failure). Tento problém rieši vlasnosť decentrelizácie, ktorá zabezpečuje, že každá jednotka systému vykonáva rovnakú funkciu.

4.2.1.4 Funkčné požiadavky

Uložené dáta je potrebné vhodne štruktúrovať aby bolo s nimi možné efektívne pracovať. Od systému požadujeme nasledujúce funkcie:

- vyhľadávanie emailov podľa údajov tvoriacich obálku, podľa hlavičiek (predmet, odosielateľ, príjemca), podľa textu obsiahnutom v tele správy a podľa názvu prílohy
- prístup koncových užívateľov k archívu, jeho prehľadávanie, export vybraných emailov
- tvorba štatistík (ich konkrétny popis)
- expirácia emailov uložených v databáze po predom určenej dobe
- unikátne ukladanie príloh emailov z dôvodu úspory diskového priestoru

Podľa popisu databázových systémov v prechádzajúcej kapitole vidíme, že použitie relačnej databázy nie je vhodné pre riešenie nášho problému. Medzi základné problémy patrí náročné horizontálne škálovanie a problémy s dostupnosťou. Emailová správa je štruktúrovaný objekt, časť z našich funkcií systému bude s touto štruktúrou a jej časťami pracovať, rozhodli sme sa, že náš výber noSQL systému zúžime na tie systémy, ktoré využívajú stĺpcový dátový model.

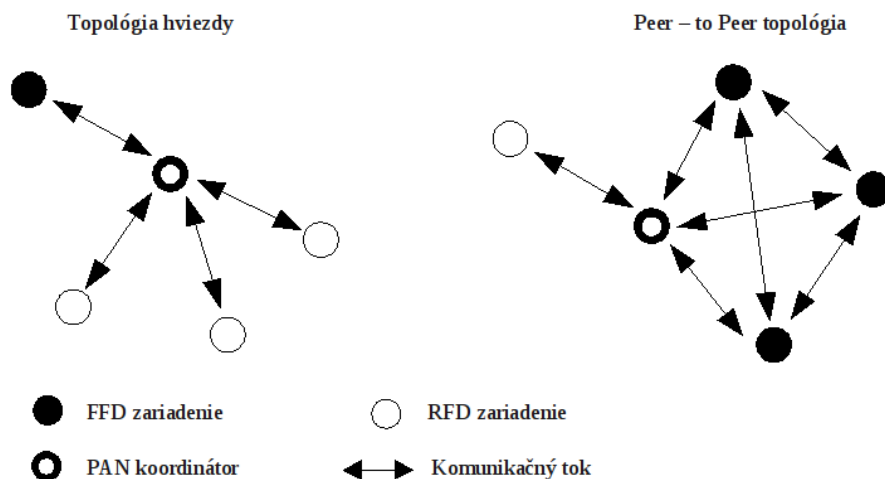
Kapitola 5

test

5.1 Komponenty IEEE 802.15.4

Zariadenia delíme na dva druhy a to zariadenie poskytujúce úplnú funkčnosť FFD (Full-function device) a redukované zariadenia RFD (Reduced-function device). Zariadenie FFD môže operovať v troch módoch, ktorými sú PAN (Personal area network) koordinátor, koordinátor, alebo koncové zariadenie. Zariadenie FFD ďalej dokáže komunikovať so zariadeniami typu RFD alebo FFD a zariadenie RFD je schopné komunikácie len so zariadením FFD. Výhodou RFD zariadení, je že neposielaajú veľké objemy dát a teda na ich realizáciu je potrebná minimálna pamäťová kapacita. Samotná WPAN je tvorená dvoma alebo viacerými zariadeniami, ktoré komunikujú na tom istom fyzickom kanály.

5.2 Sieťová topológia



Obr. 5.1: Topológie štandardu 802.15.4

5.3 Architektúra

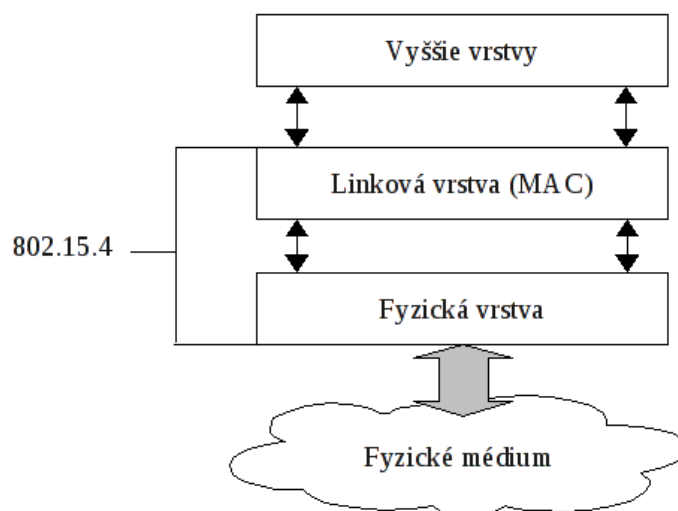
Architektúra IEEE 802.15.4 vid'. obrázok 5.2 je tvorená pomocou vrstiev. Každá vrstva je zodpovedná za časť štandardu a poskytuje služby vyšším vrstvám. Vrstvy sú zavedené z dôvodu aby bol štandard ľahko popísateľný modelom ISO-OSI.

LR-WPAN zariadenie je tvorené fyzickou vrstvou, ktorá zahŕňa rádiový vysielateľ, prijímač a mechanizmus potrebný na jeho obsluhu. Ďalšiu vrstvu tvorí linková vrstva, ktorá je zodpovedná za prístup všetkých prenosov ku komunikačnému kanálu. Jednotlivé vrstvy medzi sebou komunikujú pomocou prístupových bodov SAP (Service access point). Medzi vyššie vrstvy patrí sieťová vrstva a aplikačná vrstva, ktoré poskytujú zamýšľanú funkčnosť zariadenia a patria do štandardu ZigBee. Architektúra LR-WPAN môže byť implementovaná u embeded zariadení a taktiež aj u zariadení vyžadujúcich pre svoj chod PC.

5.3.0.5 Typy rámcov

Štruktúry rámcov boli navrhnuté aby zostala komplexnosť na minime, pri súčasnom zachovaní robustnosti pri posielaní dát na kanáli obsahujúcom šum. Štandard definuje štyri druhy rámcov a to:

- beacon rámeč
- dátový rámeč, ktorý sa používa pre prenos všetkých dát
- potvrdzovací rámeč, použitý k potvrdeniu úspešne prijatého rámcu
- MAC rámeč, ktorý sa používa k spracovaniu všetkých servisných MAC prenosov



Obr. 5.2: Štruktúra štandardu 802.15.4

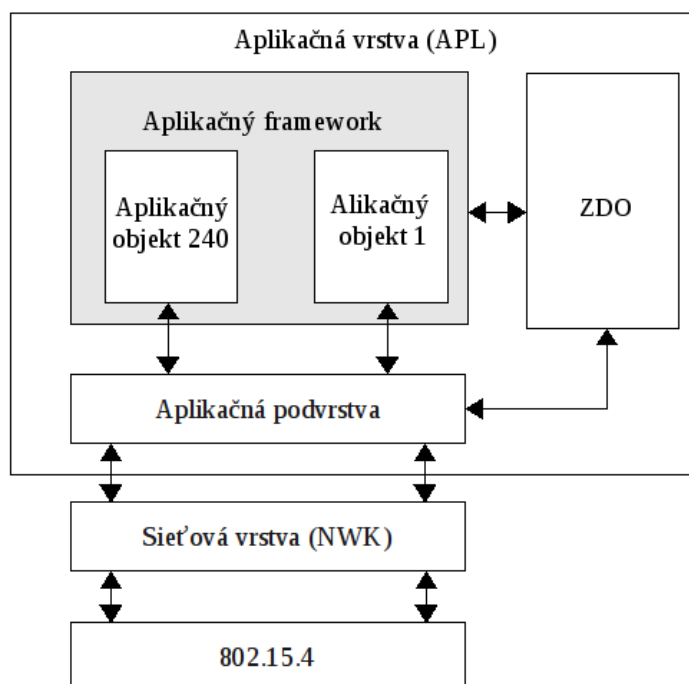
5.4 Architektúra

Architektúra štandardu ZigBee je tvorená viacerými blokmi, ktoré nazývame vrstvy. Každá z vrstiev vykonáva špecifickú množinu operácií, ktoré následne využívajú nadradené vrstvy. Ďalej sú použité dátové jednotky, ktoré sú zodpovedné za prenos dát a riadiace jednotky vykonávajúce všetky ostatné služby. Každá jednotka poskytuje svoje funkcie vyšším vrstvám pomocou rozhrania, ktoré sa nazýva SAP. Každé SAP rozhranie poskytuje množstvo servisných primitív, pomocou ktorých sa dosahuje požadovaná funkcionálna. ZigBee aliancia stavia na základoch štandardu IEEE 802.15.4-2003, ktorý definuje dve vrstvy a to fyzickú vrstvu (PHY) a spojovú vrstvu (MAC), viď. predošlá kapitola. Nadstavbu týchto vrstiev tvorí sieťová vrstva (NWK) a framework aplikačnej vrstvy. Tento framework je ďalej tvorený vrstvou APS (Application support sub-layer) a vrstvou nazývanou ZDO (ZigBee device objects). Špecifické aplikácie daných výrobcov následne využívajú daný framework, vrstvy APS a objekty ZDO. Samotná architektúra je zobrazená na obrázku 5.3

5.5 Sieťové komponenty

U siete typu ZigBee rozoznávame nasledujúce tri druhy zariadení:

- ZigBee koordinátor, odpovedá PAN koordinátoru štandardu IEEE 802.15.4-2003
- koncové zariadenie ZigBee, FFD alebo RFD zariadenie štandardu IEEE 802.15.4-2003, ktoré v ZigBee sieti nevystupuje ako ZigBee koordinátor alebo ZigBee smerovač
- ZigBee smerovač, FFD zariadenie štandardu IEEE 802.15.4-2003, ktoré nemôže vystupovať v sieti ako ZigBee koordinátor, ale vystupuje ako koordinátor štandardu IEEE 802.15.4-2003, ktorý smeruje správy medzi zariadeniami a umožňuje asociáciu



Obr. 5.3: Štruktúra štandardu ZigBee

5.6 Sieťová topológia

Sieťová vrstva štandardu ZigBee podporuje topológiu hviezda, ďalej stromovú a mesh (je to topológia používaná hlavne u bezdrôtových sietí, pri ktorej je každé sieťové zariadenie prepojené so všetkými ostatnými zariadeniami v danej sieti) topológiu.

5.7 XBee Series 1

Konkrétnu implementáciu predchádzajúcich štandardov tvoria moduly XBee/XBee PRO OEM od spoločnosti Digi, ktorá patrí do zoznamu certifikovaných výrobcov. V tejto práci som použil modul XBee Series 1, ktorý implementuje len štandard 802.15.4. Tento modul bol použitý pri meraniach v anténnej komore a jeho popisujúce parametre v samotnej simulácii. Linková vrstva modelu, ktorý som použil, pracovala s parametrami modulu TI CC1100, keď som chcel tieto parametre nahradiť, kontaktoval som spoločnosť Digi, ktorá mi však tieto parametre nebola schopná poskytnúť z dôvodu, že časť modulu, ktorá je zodpovedná za hodnoty týchto parametrov tvorí Ember 3.1 Stack, ktorý dodáva spoločnosť Ember. Spoločnosť Ember, však na moju žiadosť o sprístupnenie týchto parametrov (čas medzi prechodom z režimu spánku do režimu vysielania/prijímania, čas pri prechode zo stavu vysielania do stavu prijímania a opačne) vôbec nezareagovala.

5.7.1 Technické parametre

Modul popisujú nasledujúce parametre[4].

Dosah vnútri	30m
Dosah vonku	90m
Sila výstupného signálu	1mW (0 dBm)
Rýchlosť prenosu dát	250 000 bps
Rýchlosť sériového rozhrania	1200 bps - 250 kbps
Citlivosť pri prijíme	-92 dBm

Tabuľka 5.1: Špecifikácia výkonu a rýchlosti

Frekvenčné pásmo	ISM 2.4 GHz
Rozmery	2.438cm x 2.761cm
Operačná teplota	-40 až 85C

Tabuľka 5.2: Obecné parametre

Podporované sieťové technológie	Point-to-point, Point-to-multipoint, Peer-to-peer
Počet kanálov	16
Anténa	Whip
Adresácia	PAN ID

Tabuľka 5.3: Sieťové parametre

Kapitola 6

Teória antén

Anténa je zariadenie, ktoré slúži na vysielanie a príjem rádiových signálov. Toto zariadenie konvertuje elektromagnetické vlny na elektrickú energiu a opačne. Podľa toho ako sú signály vysielané ich delíme na všesmerové (vysielanie vo všetkých smeroch) a smerové (vysielaajú len v danom smere). Medzi chovaním sa vysielacej a prijímacej antény nepozorujeme žiadne rozdiely.

Následujúce parametre slúžia na popis základných vlastností antén:

- smerovosť, určuje v akom smere sú elektromagnetické vlny vysielané. Je posudzovaná na základe vyžarovaných charakteristík, ktoré delíme na vertikálne a horizontálne. Meria sa pomocou parametrov zisk antény a vyžarovací uhol.
- vyžarovací uhol
- impedancia antény
- zisk antény
- frekvenčná šírka prenášaného pásma
- polarizácia
- účinnosť

H rovina, je rovina v ktorej sa šíri vektor magnetického poľa a sleduje sa v nej ako sa mení intenzita elektrického poľa. Naopak v E rovine sa šíri vektor elektrického poľa a sleduje sa zmena intenzity magnetického poľa.

6.0.2 Definícia pojmov

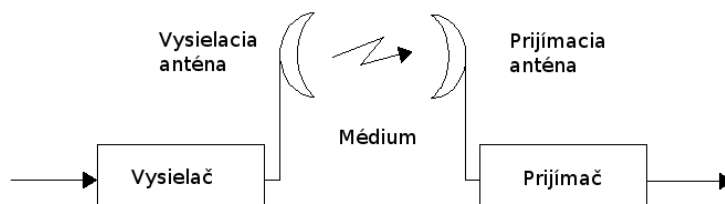
Účinnosť je pomer vyžarovaného výkonu k výkonu, ktorý privádzame na vstup antény.

Zisk určuje mieru smerovosti antény. Definujeme ho ako pomer vyžarovanej intenzity antény v danom smere k intenzite, ktorá je vyprodukovaná ideálnou anténou vyžarujúcou do všetkých smerov rovnomerne, bez strát a obe antény majú na vstupe rovnaký výkon. Zisk berie v úvahu

okrem smerovosti aj účinnosť antény. Pre zisk ďalej platí, ak má anténa pre daný smer väčší zisk ako je celkový zisk antény, tak v nejakom inom smere musí byť zisk menší aby bola zachovaná celková energia. Tohoto faktu si je možno všimnúť u grafov popisujúcich zisk antén, ktoré boli vytvorené meraním v anténnej komore viď obrázok 6.2.

6.0.3 Propagácia rádiového signálu

Následujúci obrázok 6.1 zachytáva typický rádiový systém. Informácie vstupujú do vysielача. Informácie sú následne vysielané anténou, ktorá konvertuje rádiový signál na elektromagnetické vlny. Médium na prenos elektromagnetických vln je voľný priestor. Následne sú elektromagnetické vlny odchytené pomocou prijímacej antény, ktorá ich konvertuje späť na rádiový signál. Ideálny stav je ak tento signál odpovedá signálu generovanému vysielачom. Originálna informácia, ktorá vstupovala do vysielача je následne demodulovaná na svoju pôvodnú formu.



Obr. 6.1: Model rádiového systému

6.0.3.1 Pojmy

- dB - je skratka pre decibel. Je to matematické vyjadrenie používané na zobrazenie závislosti medzi dvoma hodnotami.
- Rádiový výkon - je buď výkon vysielача alebo prijímača vyjadrený vo Wattoch. Taktiež môže byť vyjadrený v dBm. Vzťah medzi dBm a Wattmi je vyjadrený nasledovne $P_{dBm} = 10 * \log P_{mW}$
- Zoslabenie signálu modeluje nasledujúci obrázok 6.3. Kde P_{in} je vstupný výkon a P_{out} je hodnota výstupného výkonu.
Zoslabenie je vyjadrené v dB podľa nasledujúceho vzťahu: $P_{dB} = 10 * \log(P_{out}/P_{in})$. Napríklad predpokladajme, že dôjde k strate 1/3 vysielaného signálu ($P_{out}/P_{in} = 2/3$), potom hodnota zoslabenia v dB je $10 * \log(2/3) = -4.05$ dB
- Citlivosť prijímača je minimálna hodnota výkonu radio frekvenčného signálu potrebná na vstupe prijímača, aby bol signál ďalej spracovaný.
- Stratovosť je oslabenie výkonu rádiového signálu, ktorý je šírený v priestore. Je vyjadrená v dB a ďalej závisí na vzdialenosti medzi vysielacou a prijímacou anténou, na viditeľnosti medzi vysielacou a prijímacou anténou a na veľkosti antén.

6.0.4 Typy antén

Izotropická anténa sa používa pre teoretické účely, vysielané vlny majú rovnaké parametre, ktoré popisujú anténu vo všetkých smeroch. Používa sa hlavne pri popisovaní a porovnávaní vlastností reálnych antén.

Horn anténa sa používa v situácia, kde je potrebné dosiahnuť vysokého zisku, vlnová dĺžka je krátka, môže byť širokopásmová alebo úzko-pásmová, čo závisí na jej tvare. Taktiež dokáže pracovať s akoukoľvek frekvenciou. Keďže charakteristiky tejto antény sú známe a dobre matematicky popísané používa sa táto anténa ku kalibrácii iných systémov, tento fakt bol využitý aj počas meraní v anténnej komore. Hlavným parametrom, ktorý bol podstatný u meraní s touto anténou je jej zisk.

Whip anténa, je model antény, ktorý používajú ZigBee zariadenia, ktoré som simuloval. Anténa je poväčšine vertikálna a u ZigBee zariadení upevnená na dosičke plošného spoja. Je to anténa, ktorá vysiela horizontálne do všetkých smerov a hluché zóny sú vertikálne v bode upevnenia a ukončenia.

6.0.5 Stratovosť voľného priestoru (Free-space path loss)

Stratovosť signálu vo voľnom priestore sa používa k predikcii sily rádiového signálu. Napriek tomu, že nemodeluje dôveryhodne realitu obsahujúcu prekážky, odrazy atď, má veľký význam pre základné pochopenie šírenia sa signálu v reálnych podmienkach. Využíva sa taktiež pri tvorbe simulačných modelov a pri vývoji v oblasti bezdrôtových systémov.

Definujeme ju ako stratu sily signálu elektromagnetickej vlny, ktorá vzniká medzi dvoma priamo viditeľnými bodmi vo voľnom priestore, kde nie sú žiadne prekážky, odrazy a nedochádza k ohybu vln.

Formula pre výpočet stratovosti je nasledovná:

$$FSPL = \left(\frac{4\pi d}{\lambda} \right)^2 = \left(\frac{4\pi df}{c} \right)^2,$$

kde:

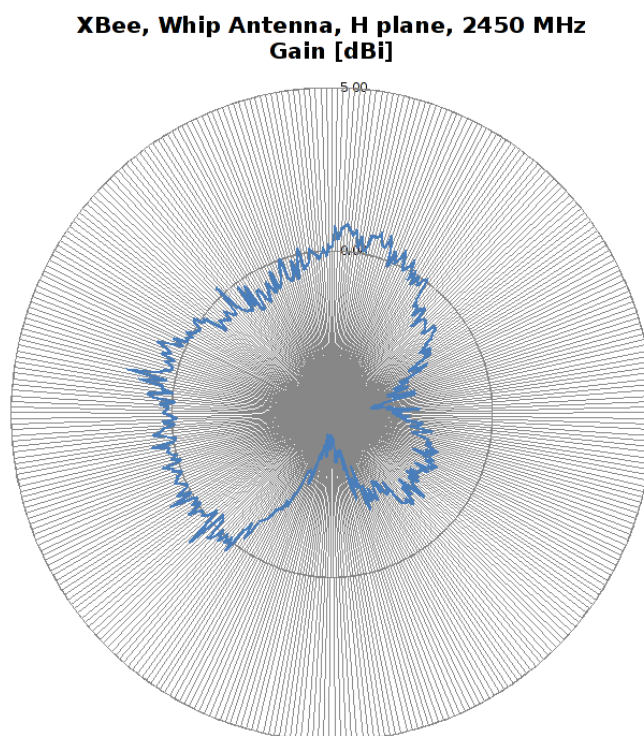
- λ je vlnová dĺžka (m)
- f je frekvencia signálu (Hz)
- d je vzdialenosť od vysielača (m)
- c je rýchlosť svetla vo vákuu $2.99792458 * 10^8$ m/s

Daná formula vyjadrená v dB:

$$\begin{aligned} FSPL(dB) &= 10 \log_{10} \left(\frac{4\pi df}{c} \right)^2 \\ &= 20 \log_{10} \left(\frac{4\pi df}{c} \right) \\ &= 20 \log_{10}(d) + 20 \log_{10}(f) + 20 \log_{10} \left(\frac{4\pi}{c} \right) \end{aligned}$$

6.0.6 Antény a simulácia

Jednotlivé antény, na ktorých bolo vykonané meranie v anténnej komore, boli popísané hodnotami výkonu (v dBm a W) v rozmedzí 0-360 stupňov pre každý stupeň. Z týchto hodnôt som pre každý stupeň spočítal hodnotu zisku. Samotný XML súbor, ktorý popisuje anténu a je použitý k simulácii potom obsahuje hodnoty ziskov po 10 stupňov, ktoré som spočítal ako priemer z hodnôt z meraní, čo samozrejme spôsobí ďalšiu odchýlku od reality.



Obr. 6.2: Charakteristika zisku antény Whip



Obr. 6.3: Model zoslabenia signálu

inicializácia - vybudovanie modelu a vloženie prvých udalostí do FES

```
while(FES nie je prázdna a simulácia neskončila)
{
vyber prvú udalosť z FES
t:= sprav časovú značku tejto udalosti
spracuj udalosť
(spracovanie udalosti môže pridať novú udalosť alebo zmazať existujúcu z FES)
}

ukončenie simulácie (zápis výsledkov, atď.)
```


Kapitola 7

Analýza a návrh riešenia

7.0.6.1 Výkonová závislosť

Následujúci obrázok 7.1 zobrazuje vysielateľ (T), u ktorého vystupuje dvojica výkonov P_1 , P_2 a hodnota zisku pre daný smer $G_T(\alpha_T)$, ďalej u prijímača (R) vystupujú výkony P_3 , P_4 a zisk v danom smere $G_R(\alpha_R)$. V simulácii bude zohrávať hlavnú úlohu hodnota výkonu P_4 , hodnota výkonu P_1 je pred vyslaním rámcu známa, je to hodnota špecifikovaná výrobcom zariadenia. V nasledujúcej časti odvodím vzťah pomocou, ktorého určím hodnotu výkonu P_4 , na základe ktorej sa prijímač rozhodne či sa naozaj jedná o prijímané dáta alebo sa vysielaný signál zoslabil na takú úroveň, kedy bude považovaný za šum na kanály.

Nasleduje formula pre výpočet hodnoty výkonu P_4 , pomocou FSPL:

$$\left[\frac{P_3}{P_2} \right]_W = \left(\frac{\lambda}{4\pi} \right)^2 \frac{1}{d^\alpha}$$

$$\begin{aligned} \left[\frac{P_3}{P_2} \right]_{dB} &= 10 \log_{10} \left(\frac{\lambda}{4\pi} \right)^2 \frac{1}{d^\alpha} \\ &= 10 \log_{10} \left(\frac{c}{4\pi f} \right)^2 \frac{1}{d^\alpha} \\ &= 20 \log_{10} \left(\frac{c}{4\pi f} \right) - 10\alpha \log_{10} d \\ &= -40.2251 - 10\alpha \log_{10} d, \end{aligned}$$

kde:

- λ je vlnová dĺžka (m), $\lambda = \frac{c}{f}$
- f je frekvencia signálu (Hz), pre použité zariadenia XBee $f = 2450$ Mhz
- d je vzdialenosť od vysielateľa (m)
- c je rýchlosť svetla vo vákuu $2.99792458 * 10^8$ m/s

- α je koeficient útlmu prostredia (pre vzduch $\alpha = 2$)

Zavediem nasledujúce označenie:

$$L = -40.2251 - 10\alpha \log_{10} d$$

Ďalej platí:

$$P_2 = P_1 + G_T(\alpha_T)$$

$$P_3 = P_2 + L$$

$$P_4 = P_3 + G_R(\alpha_R),$$

z čoho následne vyplýva nasledujúca rovnosť:

$$P_4 = P_1 + G_T(\alpha_T) + G_R(\alpha_R) + L + K,$$

kde:

- P_1 - výkon privedený do antény vysielacza (dBm)
- P_2 - výkon vysielaný vysielacou anténou (dBm)
- P_3 - hodnota výkonu prijatého prijímacou anténou (dBm)
- P_4 - výkon vystupujúci z káblu prijímacej antény a vstupujúci do prijímača (dBm)
- $G_T(\alpha_T)$ - zisk vysielacej antény v danom smere (dBi)
- $G_R(\alpha_R)$ - zisk prijímacej antény v danom smere (dBi)
- K - konštanta, ktorá spôsobuje ďalšie straty existujúce v reálnom prostredí (odrazy vo vodičoch, konektoroch, atď.)

Taktiež zároveň platí:

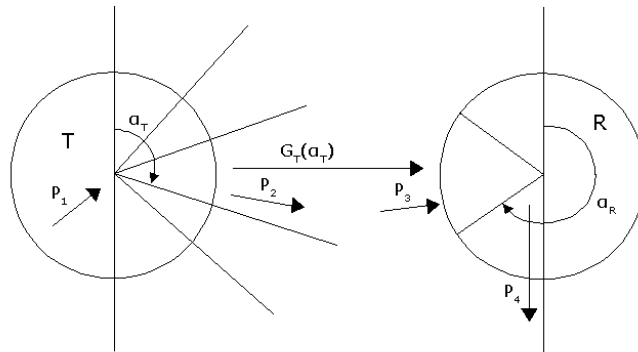
$$P'_4 < P''_4,$$

kde:

- P'_4 - hodnota výkonu v reálnom prostredí (W)
- P''_4 - spočítaná hodnota výkonu P_4 (W)

$$[P_3]_W < [P_2]_W \Rightarrow \left[\frac{P_3}{P_2} \right]_{dB} < 0.$$

Daný vzťah uvažuje straty, ktoré vznikajú na konektoroch antén, prepojovacím káblom medzi anténou a zariadením, vlastným odporom vodiča antény a iné. v podobe konštanty K .



Obr. 7.1: Popis výkonov a zisku u antén pre vysielateľ a prijímač

7.0.7 Štruktúra MF

V tejto časti sa zameriam na detailný popis funkcie modulov a ich následnú implementáciu. Obrázok 7.2 zachytáva štruktúru stanice, ktorú popisujeme pomocou jazyka NED. Na základe konvencie musí názov súboru vo svojom názve obsahovať reťazec „host“. Ako môžeme vidieť daná stanica sa skladá z jednotlivých podmodulov a to aplikačná vrstva, sieťová vrstva a vrstva NIC, ktoré navzájom komunikujú. Stanica taktiež obsahuje modul *Mobility*, ktorý zabezpečuje pohyb a geografickú pozíciu uzlu.

Podmoduly modulu stanica majú spoločného predka, ktorým je trieda *BasicModule*. Táto trieda je ďalej potomkom triedy *cSimpleModule*, ktorá je súčasťou Omnetu. Trieda *BasicModule*, obsahuje dvojfázovú inicializáciu, ktorá je vhodná aj v prípade použitia modulu *Blackbox*¹, ktorý v tomto texte nebudem popisovať. Medzi ďalšie vlastnosti patrí napríklad metóda *logName*, ktorá vracia názov NED modulu, čo je možné využiť pri ladení za pomoci makra *EV*, ktoré slúži pre výpis ladiacich správ.

Samotná komunikácia medzi jednotlivými podmodulmi modulu stanice je zabezpečená pomocou metód:

- *handleUpperMsg* - táto metóda je volaná zakaždým, keď modul obdrží správu z vyššej vrstvy, následne ju spracuje a predá nižšej vrstve
- *handleLowerMsg* - metóda je volaná po príchode správy z nižšej vrstvy, správu ďalej spracuje a predá nadradenej vrstve

Ďalej sú k dispozícii pomocné metódy slúžiace k enkapsulácii a dekapulácii správy predávanej medzi vrstvami (*encapsMsg*, *decapsMsg*) a metódy na opozdenie správy pri jej vysielaní na kanál.

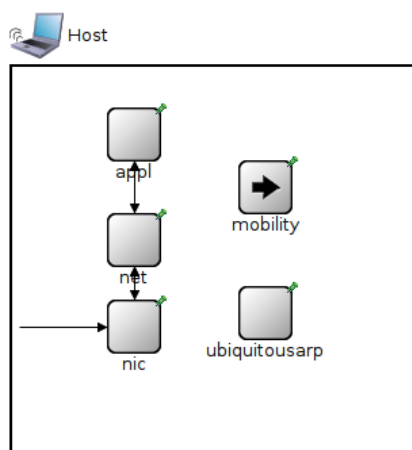
Keďže model poskytuje enkapsuláciu a dekapuláciu správ, z toho vyplýva, že pre každú vrstvu definujeme osobitný typ správy, ktorej štruktúra je popísaná pomocou súboru s príponou *.msg*. K dispozícii sú napríklad správy typu *AirFrame* (správa pre komunikáciu

¹Je to modul umožňujúci medzivrstvovú komunikáciu, bez nutnosti zasielania správ

na fyzickej vrstve), *MacPkt* (správa pre komunikáciu na linkovej vrstve), *NtwPkt* (správa pre sieťovú vrstvu) a *ApplPkt* (správa pre vrstvu aplikačnú). Všetky typy správ sú rozšíriteľné, resp. si pomocou nich môžeme odvodiť vlastný typ správy. Pre potreby mojej simulácie bola najdôležitejšia správa typu *AirFrameRadioAccNoise3*, zdedená z triedy *AirFrame*, ktorej štruktúra obsahuje nasledujúce položky

- pSend - výkon, ktorým bol rámec odoslaný
- channelId - kanál, na ktorý bola správa zaslaná
- duration - čas, ktorý bol spotrebovaný na zaslanie rámcu (v sekundách)
- hostMode - štruktúra popisujúca pozíciu stanice, jej rýchlosť, smer pohybu, čas kedy sa pohyb začal
- SnrControlInfo - trieda, ktorá obsahuje pomocné informácie SNR (Signal to noise ratio), ktoré sa predávajú modulu *Decider*

Danú správu som ďalej rozšíril o parameter *moduleId*, ktorý obsahuje globálny identifikátor modulu, z ktorého bola správa odoslaná.



Obr. 7.2: Štruktúra stanice v OMNeT++

7.0.7.1 Štruktúra modulu NIC

NIC (Network card interface) je časť sieťového adaptéru zodpovedná za fyzický prístup k médiu a adresácii na úrovni MAC vrstvy, popisuje ju fyzická a linková vrstva ISO-OSI modelu. Štruktúra tohoto modulu je znázornená na obrázku 7.3. V mojom prípade je teda fyzická vrstva tvorená modulom *snrEval*, *decider* a vrstvu MAC tvorí modul *mac*. Vzájomná úzka kooperácia medzi týmito vrstvami je dôvodom prečo sú zapuzdrené v jednom module. Všetky submoduly modulu NIC, sú zdedené z triedy *ChannelAcces*, ktorá je ďalej odvozená z triedy *BasicModule*. Táto trieda poskytuje funkcionality umožňujúcu komunikáciu

jednotlivých staníc, ktoré sú vo vzájomnom dosahu. Na úrovni fyzickej vrstvy ma hlavne zaujíma modelovanie oslabenia signálu a výpočet chybovosti na kanále. Tým, že je fyzická vrstva tvorená osobitne modulom *snrEval*, som schopný modelovať výpočet stratovosti na kanále pomocou rôznych metód a na základe výsledku sa v module *decider* rozhodnem pomocou akého kritéria tieto dáta vyhodnotím. Napríklad modul *decider* bude rozhodovať o tom či dané dáta prijme na základe porovnania hodnoty SNR, spočítanej z modulu *snrEval*, s definovanou hraničnou hodnotou, alebo sa môže rozhodovať na základe počítania pomocou formúl pre výpočet chybovosti na kanále (napr. BER). Vďaka tomu môžem tieto moduly navzájom rôzne kombinovať. V nasledujúcej časti detailnejšie priblížim štruktúru modulov *snrEval* a *decider*.

Modul *snrEval*

Tento modul zabezpečuje príjem a vysielanie dát na kanál. Ďalej vytvára správy typu *AirFrame* z *MacFramu* a opačne, počas toho ako odpočúva kanál zároveň mení stavy rádia, ktoré sú reprezentované stavovým automatom a taktiež zabezpečuje simuláciu oneskorenia vo vysielaní alebo prijíme pomocou pomocných funkcií (*bufferMsg*, *unbufferMsg*). V mojom modeli ma zaujímala jedna z jeho ďalších vlastností a to ukladanie a spracúvanie SNR hodnôt pri prijímaní rámcov. Rámec fyzickej vrstvy (*AirFrame*) obsahuje pomocnú štruktúru *SnrList*, ktorú reprezentuje štruktúra *List* programovacieho jazyka C++ a záznam tejto štruktúry obsahuje dve položky a to časovú značku prijatia rámcu a k nej odpovedajúcu hodnotu SNR. V mojom prípade je hodnota SNR spočítaná na základe vzťahu [hodnota výkonu vstupujúca do prijímača (P_4) / hodnota šumu na kanále], kde hodnotu P_4 počítam za využitia modifikovanej formule FSPL. Túto hodnotu následne predávam do modulu *Decider*, kde je ďalej spracovaná.

Práca modulu *snrEval* je znázornená vývojovým diagramom na obrázku 7.4. Keď sa nachádza modul v stave SYNC prijíma správu. Pred jej prijatím sa najskôr vykoná kontrola, či nedošlo k poškodeniu SFD (Start frame delimiter), následne sa spracuje zvyšok správy a spočíta sa hodnota SNR. V prípade, že sa modul nenachádza v stave SYNC a obdrží ďalšiu správu je tato správa považovaná za šum, hodnota šumu sa zvýši o hodnotu výkonu, ktorým bola táto správa prijatá a spočíta sa nová hodnota SNR, ku ktorej je pripojená časová značka. Detailnejšie sa budem zaoberať modelom kolízie v nasledujúcej kapitole.

V tomto module je metóda *handleLowerMsg* rozdelená na dve časti a to:

- *handleLowerMsgStart* - volá sa hneď po prijatí správy, volá metódy na výpočet hodnoty prijatého výkonu (P_4) a následne predáva spracovanie ďalším metódam na základe stavu rádia
- *handleLowerMsgEnd* - slúži na samotné odoslanie správy vyššej vrstve a zároveň pripája *SnrList* ako parameter.

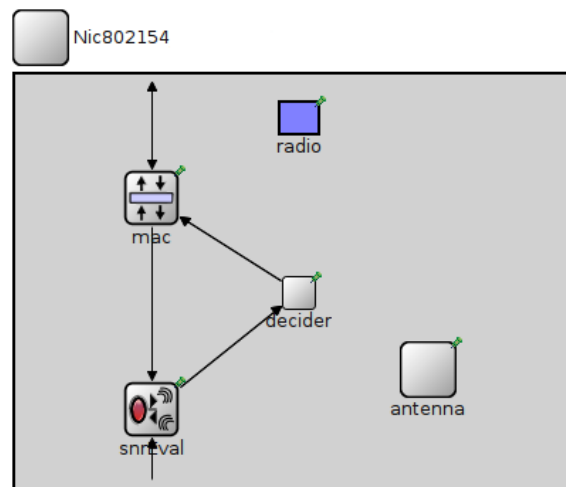
Modul *decider*

Modul spracúva len správy, ktoré prichádzajú z kanálu cez modul *SnrEval*. Správy z vyšších vrstiev, ktoré sa posielajú na kanál neprechádzajú týmto modulom a to z dôvodu, že tento modul len rozhoduje, či sa daná správa zahodí alebo prepošle vyššej vrstve. Rozhodovanie je založené na základe výpočtov ako napríklad chybovosť bitov alebo sa rozhoduje či sa má

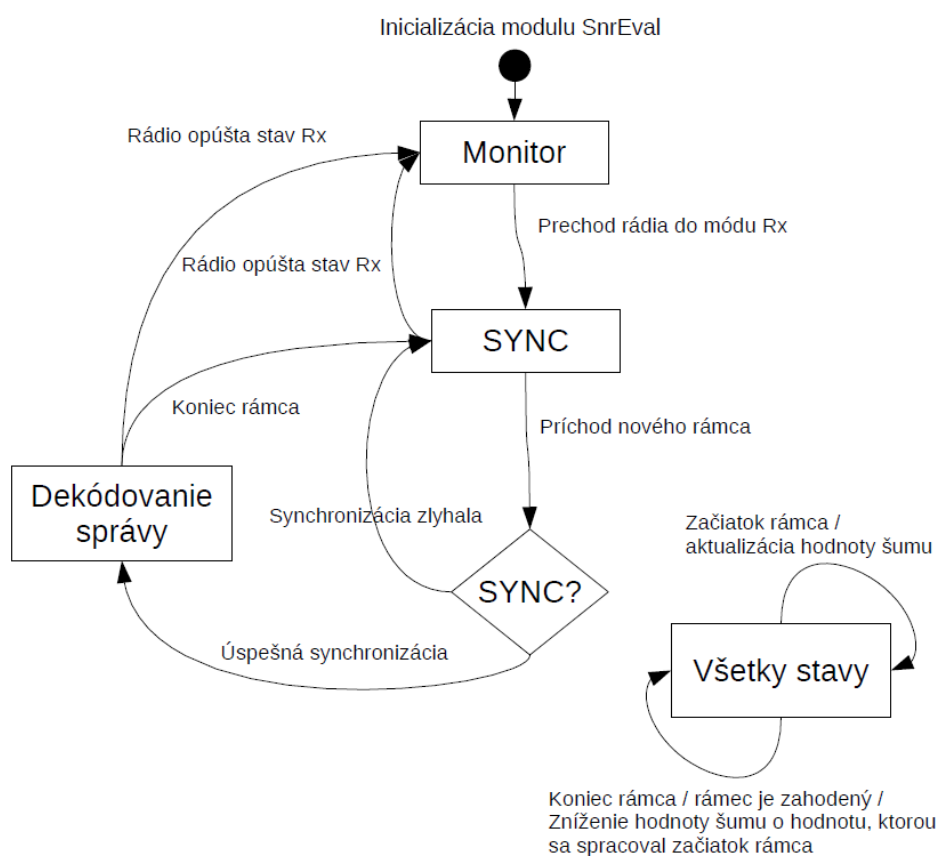
správa zahodiť na základe vysokej hodnoty šumu na kanále porovnaním s hraničnou hodnotou. Tieto vlastnosti však sledujem len u správ prichádzajúcich z kanálu. *Decider* teda slúži na výpočet chybných bitov (BER) v správe, čo počíta pomocou hodnôt uložených v štruktúre *SnrList*. Ďalej je v ňom možné implementovať rôzne opravné kódy. V simulácii je použitý vzorec na výpočet chybných bitov (BER) pre moduláciu MSK.

$$BER = 0.5 * \exp(-0.5 * SNR)$$

Modul *mac* ďalej poskytuje funkcionality metódy na riadenie prístupu k médiu CSMA. Modul *radio* je centrálne zdieľaný a moduly ako *snrEval* alebo *mac* prepínajú jeho stavy (napr. RX, TX) podľa potreby. Posledným modulom, je modul *antenna* ktorý popisuje parametre antény a je využívaný modulom *snrEval*.



Obr. 7.3: Štruktúra NIC v OMNeT++

Obr. 7.4: Prechody medzi stavmi v module `snrEval`

Kapitola 8

Realizácia

Pre potrebu simulácie som teda ako prvý vytvoril modul *Antenna*, ktorý popisuje anténu ZigBee zariadení nasledujúcimi parametrami:

- zisk - ideálny zisk antény, ktorý nájdeme v popise antény (dBi)
- impedancia (Ω)
- config - odkazuje na xml súbor popisujúci zisk antény (*.xml)

Modul sa snaží byť čo najobecnejší pre prípadne využitie aj v iných modeloch a pridal som ho medzi ostatné moduly MF. Z predošlých parametrov je najdôležitejší parameter *config*. Pomocou tohoto parametru sa odkazujem na súbor, ktorým popisujem zisk antény pre daný uhol, v ktorom anténa prijíma alebo vysiela. Hodnoty tohoto súboru sú z meraní, ktoré boli vykonané v anténnej komore. Nasledujúce riadky zachytávajú príklad popisu konkrétnej antény.

```
<?xml version="1.0" encoding="UTF-8"?>
<!DOCTYPE antenna SYSTEM "Antenna.dtd">
<antenna>
  <angle min="0" max="0" gain="0.3" />
  <angle min="0" max="30" gain="0.4" />
  <angle min="30" max="360" gain="0.9" />
</antenna>
```

Jednotlivé záznamy v elementoch `<angle ...>`, reprezentujú hodnotu zisku pre uhol z intervalu (min, max). Jednotlivé intervaly musia byť zoradené vzostupne, bez prekryvania sa. V prípade, že je rámec odoslaný pod uhlom ktorý, sa v popisnom xml súbore nenachádza, je použitá hodnota ideálneho zisku, ktorá je zapísaná v .ned súbore popisujúcom štruktúru modulu antény. Táto hodnota môže byť prepísaná v konfiguračnom súbore danej simulácie omnetpp.ini pomocou nasledujúceho zápisu: `sim.host[*].nic.antenna.gain = 2.1dBi`

Parametru *config*, ktorý odkazuje na xml súbor je potrebné priradiť xml súbor aj v prípade ak chcem v simulácii použiť len hodnotu teoretického zisku antény. V aktuálnej verzii Omnetu nie je možné priradiť tomuto parametru napríklad hodnotu „false“ a zabezpečiť

týmto neodkazovanie sa na xml súbor. Túto funkcionálnosť som ďalej do Omnetu neimplementoval, pretože po konzultácii s autorom Omnetu, som sa dozvedel, že táto možnosť pribudne v jeho novej verzii. Tento prípad riešim tak, že parametru config priradím xml súbor obsahujúci samotný koreňový element `<root />`, ktorý vyhodnotím pri samotnom spracúvaní xml súboru.

Samotný xml súbor popisujúci konkrétny model antény priradím danej stanici následovne: `sim.host[*].nic.antenna.config = xmldoc(„antenna1.xml“)`

Po spustení simulácie sa v inicializačnej časti modulu validuje daný xml súbor pomocou súboru `Antenna.dtd`, ďalej sa načíta do pamäti, z dôvodu, že simulátor poskytuje len DOM parsér a sprístupním si odkaz na jeho prvý element `<angle ...>`. Modul ďalej obsahuje metódu `findGainValue`, ktorá v danom xml súbore vyhľadá hodnotu zisku pre daný uhol. Z dôvodu optimalizácie som pre vyhľadávanie v štruktúre xml súboru použil binárne vyhľadávanie.

Ďalej som do modulu `SnrEvalRadioAccNoise3` implementoval výpočet modifikovanej formule FSPL. Samotný priestor, takzvaný playground, v ktorom sa odohráva simulácia som rozdelil na štyri kvadranty, vďaka čomu dokážem veľmi efektívne počítať uhol, pod ktorým bol rámec vyslaný z vysielačnej stanice a uhol, pod ktorým bol rámec prijatý na prijímacej stanici. Tieto uhly sú prepočítané na strane príjemcu, z ich hodnôt zistím pomocou modulu antény, konkrétne hodnoty daných ziskov $G_T(\alpha_T)$ a $G_R(\alpha_R)$. Hodnotu zisku $G_R(\alpha_R)$ určím priamo pomocou metódy `findGainValue` modulu `antenna`, ktorý obsahuje prijímač. Prijatý rámec obsahuje hodnotu jedinečného identifikátoru modulu (`moduleId`), z ktorého bol vyslaný. Pomocou neho sprístupním odkaz na modul vysielača a taktiež zavolám jeho metódu `findGainValue`, ktorá mi vráti hodnotu zisku $G_T(\alpha_T)$. Následne môžem spočítať hodnotu výkonu P_4 , z tejto hodnoty sa ďalej spočíta hodnota SNR pomocou vzťahu $SNR = P_4 / [\text{hodnota šumu prostredia}]$, kde hodnota šumu prostredia je rovná -100dBm. SNR sa ďalej pripojí ako kontrolná informácia k rámcu a odošle o úroveň vyššie vrstve `decider`. Táto vrstva následne spočíta hodnotu BER pre daný rámec a v prípade, že nedošlo k poškodeniu rámca je tento rámec predaný opäť vyššej vrstve a to vrstve MAC.

Pre potreby vyhodnocovania modelov, som ďalej upravil modul `ChannelControl`, kde bol pridaný parameter `ratio`, pomocou, ktorého je prepočítavaná vzdialenosť v modeli na reálnu vzdialenosť.

Počas implementácie a ladenia modelu som objavil v produkčnom kóde MF, dve chyby, konkrétne pri výpočte hodnoty BER v module `decider`, druhá chyba bola v module `snrEval`. Po upozornení autora boli obe chyby opravené.

8.0.8 Model kolízie

Pri komunikácii ZigBee zariadení v reálnom prostredí môže dochádzať k ich vzájomnému rušeniu. Takáto situácia môže nastať napríklad v prípade, že nastane kolízia v mechanizme, ktorý riadi prístup k médiu (CSMA-CA) alebo ak máme dve siete, kde prijímač z prvej siete práve súčasne v jednom okamihu počas svojho stavu Rx, dva rámce. Súčasné prijatie dvoch rámcov na anténe zanesie do komunikácie šum (čo je vlastne prídavný signál), ktorý môže ďalej spôsobiť chybovosť (BER). Keďže, som chcel modelovať aj situácie, u ktorých by dochádzalo k rušeniu, musel som pre tieto potreby model čiastočne modifikovať. Daný model neposkytuje vrstvy štandardu ZigBee preto nie je možné modelovať dve nezávislé siete. Danú kolíziu som preto vytvoril pomocou modifikácie MAC vrstvy, čím som dosiahol,

že dané zariadenie sa chovalo ako generátor šumu, tj. periodicky vysiela rámce, s tým, že dochádza ku kolízii a prijímač prijme viacero rámcov súčasne.

V simulácii je používaný diskretný simulátor, počas simulačného času prebiehajú udalosti. Model súčasného prijatia viacerých rámcov vyzerá tak, že v čase keď prijímač prijme rámce simulačný čas sa zastaví a samotné prijatie rámcov považujeme za udalosti v rovnakom simulačnom čase. Procesorový čas však neustále beží a teda program obsluhujúci súčasne prijatie viacerých rámcov prijme tieto rámce v skutočnosti s určitým časovým odstupom, čo je v simulácii reprezentované pomocou udalosti. Najprv je prijatý prvý rámec, je spracovaná jeho obsluha, následne sa spracuje druhý rámec atď. Tomuto popisu odpovedá nasledujúci obrázok 8.2, ktorý zachytáva situáciu pri ktorej došlo ku kolízii v modifikovanej vrstve MAC (modul *mac*).

8.0.8.1 Popis kolízie

Obrázok 8.2 je výstup z nástroja Sequence chart a detailný popis jednotlivých udalosti je možné analyzovať pomocou nástroja Event log. Oba tieto nástroje sú vhodné pre analýzu simulácie poprípadе jej ladenie a pribudli vo verzii OMNeT++ 4.0. V mojom modeli 8.1 som simuloval vznik kolízie na module *host[0]*, ktorý periodicky prijímal rámce z modulu *host[1]*. Modul *host[2]* som použil ako generátor rámcov, ktoré budú spôsobovať kolízie. Na obrázku reprezentuje sivý úsek zastavenie simulačného času. Modul *host[0]* prijal v rovnaký čas dva rámce, prvý od modulu *host[1]*, čomu odpovedá číslo udalosti 41, druhý od modulu *host[2]* s číslom udalosti 43. Oba tieto rámce boli prijaté modulom *snrEval*. Na základe predchádzajúceho popisu modulu *snrEval*, prebehne nasledujúca obsluha:

1. rámec z udalosti 41, vstupuje do modulu *snrEval*, ktorý sa sa prepne do stavu SYNC, keďže sa jedná o nový rámec, je spočítaná hodnota SNR1, nasleduje prepnutie do stavu Dekódovanie správy

$$\text{SNR1} = (\text{výkon, ktorým bol rámec prijatý} / \text{šum})$$

2. následne je prijatý rámec z udalosti 43, tento rámec bude spracúvaný ako šum, spočíta sa nová hodnota SNR2, rámec sa následne zahodí

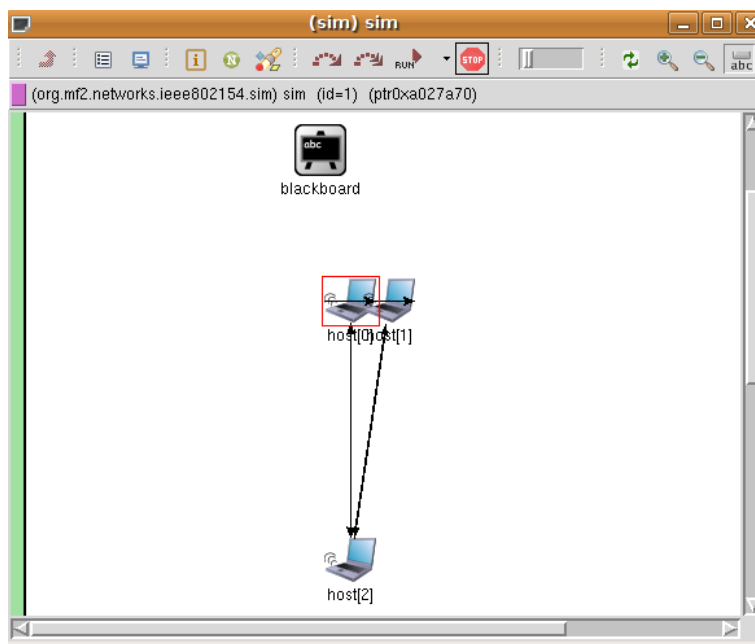
$$\text{SNR2} = (\text{hodnota výkonu, ktorým bol prijatý rámec z udalosti 41}) / (\text{šum} + \text{prijatý výkon rámcu z udalosti 43})$$

3. ukončí sa spracovanie rámcu z udalosti 41, hodnoty SNR1 a SNR2 sa pripoja ako kontrolné informácie k správe, ktorá sa následne prepošle modulu *decider*
4. *decider* na základe hodnôt SNR1, SNR2 spočíta BER a podľa jeho hodnoty sa rozhodne či sa rámec zahodí (tj. rámec obsahuje chybné bity) alebo pošle modulu *mac*

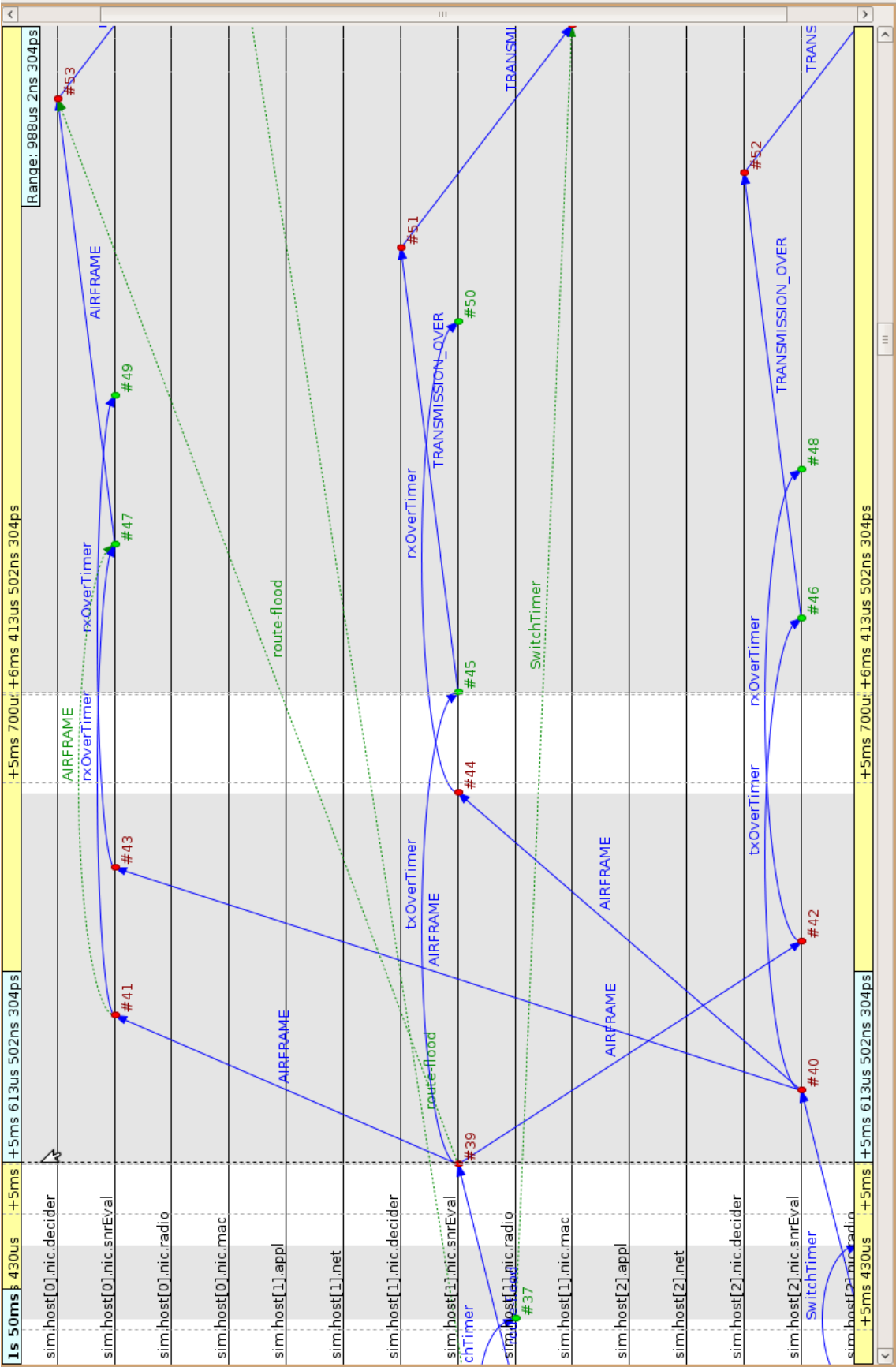
Čím je nižšia hodnota SNR, tým je vyššia pravdepodobnosť, že prijatý rámec bude obsahovať chyby. Tento fakt vychádza zo Shannonovej vety, danej nasledujúcou formulou, ktorá udáva max. teoretický limit prenosovej rýchlosti C kanálu s pásmom o šírke W a odstupom signálu od šumu (SNR).

$$C = W * \log_2(1 + SNR)[b/s, Hz]$$

V prípade, že sa hodnota SNR blíži k nule, prenosová rýchlosť kanálu sa taktiež blíži k nule, z čoho vyplýva, že dochádza k veľkej strate prenášaných dát, čo zapríčini veľkú chybovosť na prenášaných dátach.



Obr. 8.1: Model kolízie



Obr. 8.2: Kolízia na MAC vrstve

Kapitola 9

Testovanie

V tejto časti popíšem testy, ktoré som uskutočnil pomocou daného modelu a porovnáam výsledky týchto testov s reálnym meraním.

9.0.9 Testy zamerané na pohyb XBee zariadení

Pri vykonávaní týchto testov, som modeloval komunikáciu dvoch XBee zariadení, z ktorých jedno bolo v pozícii príjemcu a vysielateľ sa pohyboval. Hlavný faktor, na ktorý som kládol dôraz bolo pozorovanie ako sa mení hodnota výkonu na strane príjemcu s narastajúcou vzdialenosťou. Taktiež som sledoval počet zahodených rámcov, takzvanú stratovosť rámcov (na základe výpočtu chybovosti BER na prijímači) s narastajúcou vzdialenosťou a oslabovaním sa signálu, viď tabuľka [9.1](#)

Vykonal som nasledujúce testy:

1. vzdiaľovanie sa vysielateľa (s výkonom vysielateľa 1mW a 10mW) po kroku 0.6cm (po každom kroku bol vyslaný rámec) do vzdialenosti 5m, viď. grafy [9.1](#) a [9.2](#)
2. vzdiaľovanie sa vysielateľa (s výkonom vysielania 1mW a 10mW) po kroku 0.6cm (po každom kroku bol vyslaný rámec) do vzdialenosti 5m a súčasná náhodná rotácia oboch zariadení okolo vlastnej osi
3. vzdiaľovanie sa vysielateľa (s výkonom vysielania 1mW a 10mW) po kroku 0.6m (po každom kroku bol vyslaný rámec) do vzdialenosti 250m, viď. graf [9.3](#)
4. vzdiaľovanie sa vysielateľa (s výkonom vysielania 1mW a 10mW) po kroku 0.6m (po každom kroku bol vyslaný rámec) do vzdialenosti 250m a súčasná náhodná rotácia oboch zariadení okolo vlastnej osi, viď. graf [9.4](#)
5. náhodná rotácia vysielateľa okolo prijímateľa vo fixnej vzdialenosti 10m

Detailné výstupy z týchto testov sa nachádzajú na priloženom CD vo forme spracovaných grafov a taktiež vo forme súboru s príponou .sca, čo je jeden z výstupných formátov simulátoru Omnet, ktorý sa dá ďalej vhodne spracúvať pomocou nástroja Scave.

Analýzou týchto meraní je vidno, že krivky grafov sa takmer zhodujú s reálnymi meraniami a to aj napriek tomu, že reálne prostredie obsahuje množstvo faktorov spôsobujúcich odrazy atď. Daným modelom som schopný modelovať reálne podmienky s vysokou presnosťou aj napriek tomu, že zanedbám straty, ktoré v nich vznikajú.

Vzdialenosť [m]	Stratovosť [%]	Výkon vysielača [mW]
100	0	1
150	0.12	1
200	11.87	1
220	30.42	1
250	76.38	1
300	49.65	10

Tabuľka 9.1: Stratovosť rámcov

9.0.10 Testy s kolíziou

V týchto testoch bola poloha zariadení XBee stacionárna. Model uvažoval dve zariadenia prijímač a vysielač, kde vysielač vysiela rámce. Ďalej som do simulácie zapojil generátor šumu (zariadenie periodicky generujúce rámce, ktoré spôsobovali kolíziu). V tejto simulácii som sledoval koľko rámcov bolo zahodených (stratovosť) z dôvodu šumu spôsobeného generátormi na prijímači.

Vykonal som nasledujúce testy:

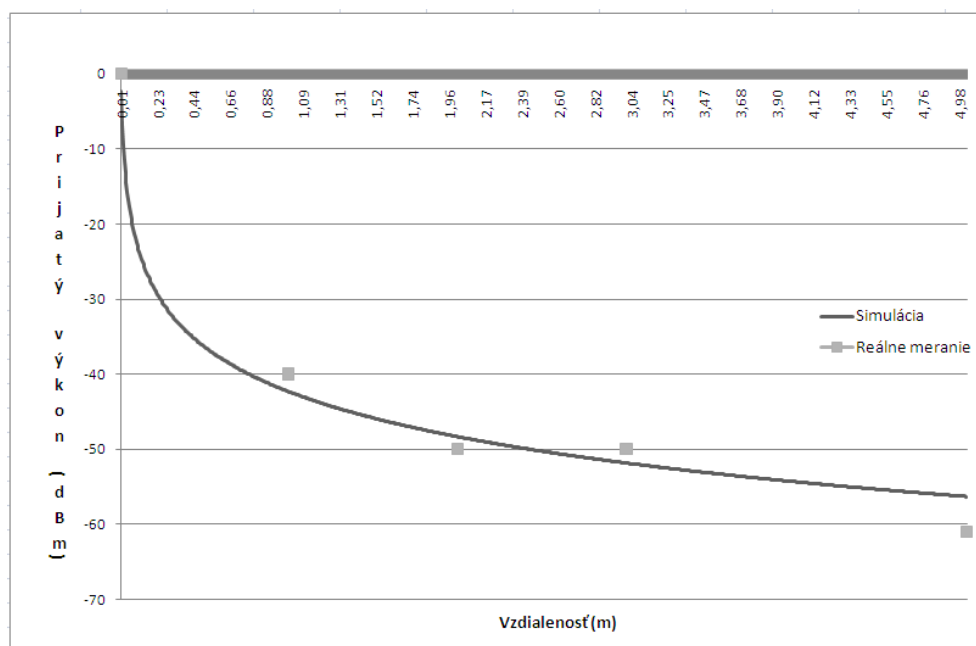
1. model vid'. obrázok 8.1 prijímač - host[0], vysielač - host[1] boli umiestnené vo vzdialenosti 30cm, generátor šumu - host[2] (vysielací výkon 10mW) vo vzdialenosti 3m a 4m
2. model totožný s predchádzajúcim no bol pridaný druhý generátor šumu, jeho umiestnenie bolo $x = \text{host}[2].x - 0.7\text{m}$, $y = \text{host}[2].y$. Vysielací výkon oboch generátorov šumu bol 10mW.

Vzdialenosť generátora šumu [m]	Stratovosť [%]	Stratovosť v reálnych podmienkach [%]
3	9.8	15
4	0	nebolo merané

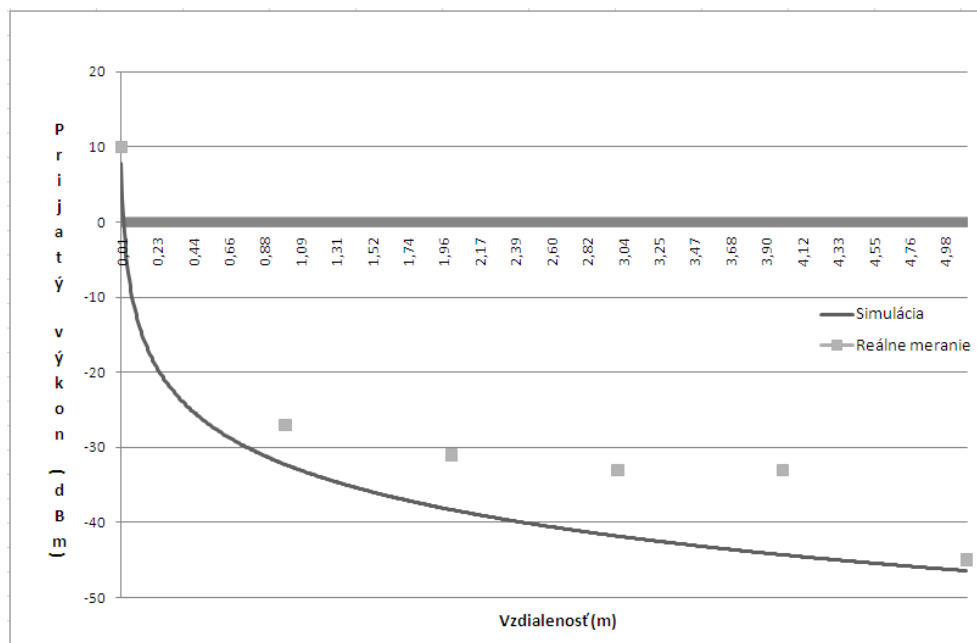
Tabuľka 9.2: Prípad č. 1

Vzdialenosť generátorov šumu [m]	Stratovosť [%]
3	98
4	80

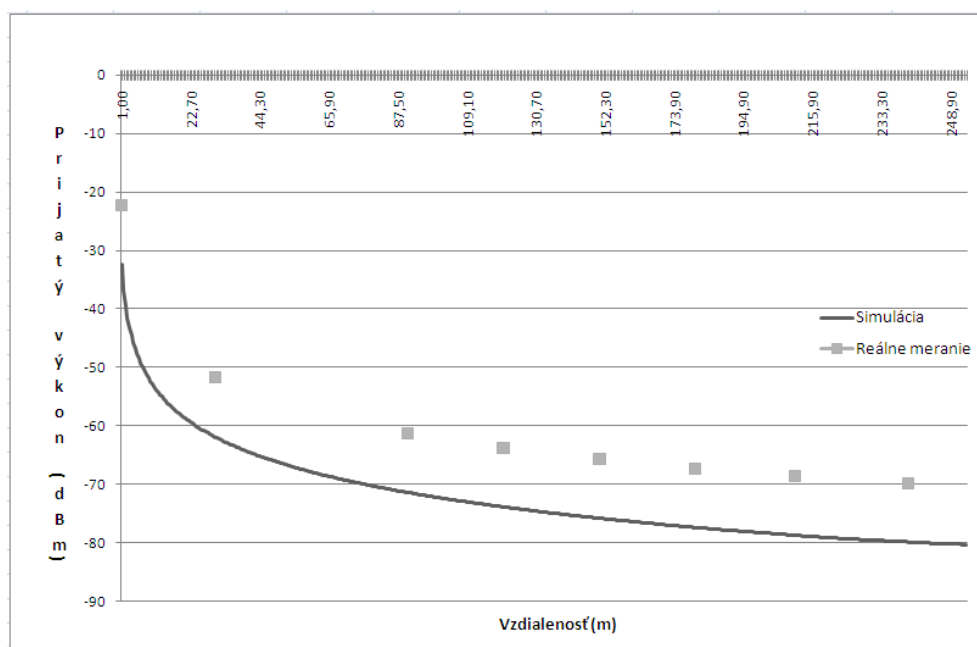
Tabuľka 9.3: Prípad č. 2



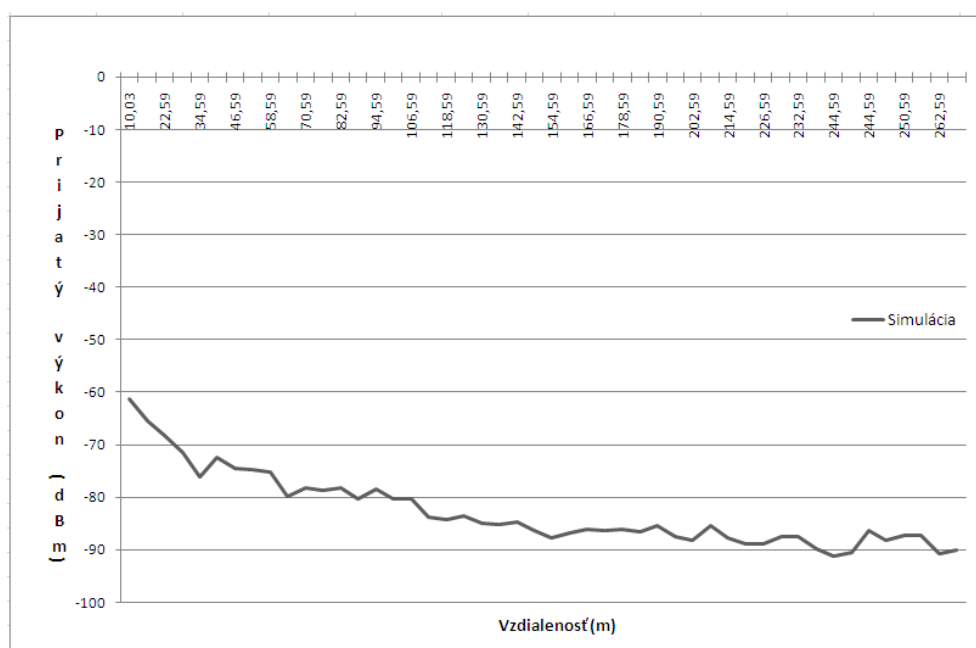
Obr. 9.1: Pohyb na vzdialenosť 5m, vysielací výkon 1mW



Obr. 9.2: Pohyb na vzdialenosť 5m, vysielací výkon 10mW



Obr. 9.3: Pohyb na vzdialenosť 250m, vysielací výkon 10mW



Obr. 9.4: Pohyb na vzdialenosť 250m, vysielač výkon 1mW, rotácia okolo vlastnej osi

Kapitola 10

Záver

Literatúra

- [1] Armando Fox, Steven D. Gribble, Yatin Chawathe, Eric A. Brewer, Paul Gauthier. Cluster-base scalable network services, 1997.
- [2] A. B. Bondi. Characteristics of Scalability and Their Impact on Performance, 2000.
- [3] Brewer. Towards robust distributed systems., 2000.
- [4] Digi International Inc. XBee/XBee-Pro ZB OEM RF Modules product manual, 2008.
- [5] E. F. Codd. A Relational Model of Data for Large Shared Data Banks, Jún, 1970.
www.seas.upenn.edu/~zives/03f/cis550/codd.pdf.
- [6] Eben Hewitt. Cassandra: The Definitive Guide, November 2010.
- [7] John F. Gantz, David Reinsel, Christopher Chute, Wolfgang Schlichting, John McArthur, Stephen Minton, Irida Xheneti, Anna Toncheva, Alex Manfrediz. The Expanding Digital Universe, A Forecast of Worldwide Information Growth Throught 2010, Marec 2007.
<http://www.emc.com/>, stav z 28.2.2011.
- [8] Michael Stonebraker. The Case for Shared Nothing Architecture, 1986.
<http://db.cs.berkeley.edu/papers/hpts85-nothing.pdf>, stav z 28.2.2011.
- [9] D. Pritchett. BASE: An Acid Alternative.
- [10] R. Shoup. The eBay Architecture, Striking a balance between site stability, feature velocity, performance, and cost, November 2006.
www.addsimplicity.com/downloads/eBaySDForum2006-11-29.pdf, stav z 28.2.2011.

Dodatok A

Zoznam použitých skratiek

APL Application Layer

APS Application Support Sub-layer

CSMA-CA Carrier Sense Multiple Access - Collision Avoidance

FFD Full Functionality DEvice

FEL Future Event List

FES Future Event Set

FSPL Free-space path loss

GTS Guaranteed Time Slot

GUI Graphical User Interface

IDE Integrated Development Environment

LQI Link Quality Indicator

LR-WPAN Low-Rate Wireless Personal Area Network

MAC Medium Access Control

MF Mobility Framework

NED Network Description

NIC Network Card Interface

NWK Network

RFD Reduced Functionality Device

PAN Person Area Network

PHY Physical

SAP Service Access Point

SFD Start Frame Delimiter

SNR Signal to noise ratio

ZDO ZigBee Device Object

WLAN Wireless Local Area Network

WPAN Wireless Personal Area Network

Dodatok B

Inštalčná a užívateľská príručka

B.0.11 Inštalácia simulátoru OMNeT++ pre platformu Linux

1. Stiahnutie archívu obsahujúceho zdrojový kód zo stránok
<http://www.omnetpp.org/omnetpp>
2. Prekopírovanie archívu do adresára /usr/local/
3. Rozbalenie archívu pomocou príkazu `tar zxvf omnetpp-4.0.src.tgz`
4. Do užívateľského profilu `.bash_profile` alebo `.profile` pridáme riadok
`export PATH=$PATH:/usr/local/omnetpp-4.0/bin`
5. Je potreba zabezpečiť prítomnosť nasledujúcich balíkov v systéme

```
sudo apt-get install build-essential gcc g++ bison flex perl tcl8.4 tcl8.4-dev  
tk8.4 tk8.4-dev blt blt-dev libxml2 libxml2-dev  
zlib1g zlib1g-dev libx11-dev
```

6. Prevedieme nasledujúce príkazy:
`cd /usr/local/omnetpp-4.0`
`./configure`
`./make`
7. Spustenie OMNeT++ s IDE pomocou príkazu `omnetpp`

B.0.12 Inštalácia mnou modifikovaného Mobility Frameworku

1. Stiahnutie súborov Mobility frameworku z svn `http://my-svn.assembla.com/svn/mframework/`,
poprípade prekopírovanie adresára `mf2o4` z priloženého CD do adresára `/usr/local/`
2. Import MF do aplikácie OMNeT++
 - (a) Po spustení aplikácie Omnet, klikneme na oblasť „Project explorer“, pravým tlačítkom
a zvolíme položku „Import...“
 - (b) Zvolíme „General->Existing project into Workspace“

- (c) V položke „Select root directory“, zvolíme cestu k adresáru mf2o4, tj. /usr/local/mf2o4
- (d) Pomocou CTRL+B, preložíme zdrojové súbory

B.0.13 Práca s modelom IEEE 802.15.4

Vo vývojom prostredí Omnetu si otvoríme v oblasti „Project explorer“ adresárovú štruktúru mf2o4, kde si následne otvoríme adresár networks a v ňom adresár ieee802.15.4. V tomto adresári sa nachádzajú aj xml súbory popisujúce antény. Otvoríme si súbor omnetpp.ini, tento súbor je hlavným konfiguračným súborom modelu simulácie. Zahŕnul som do neho ukážkové nastavenia viacerých modelov, ktoré som simuloval. Samotná simulácia sa potom spustí otvorením súboru omnetpp.ini a následným kliknutím na tlačítko „Run“ z menu aplikácie.

Dodatok C

Obsah priloženého CD

Následující obrázok [C.1](#) zobrazuje štruktúru priloženého CD.

```
.
|-- data
|   |-- XBee.xlsx           - popis charakteristik antén z meraní
|   |-- Graphs.xls         - grafy z realnych meraní
|   |-- Graphs2.xls        - grafy zo simulácie
|-- mf2o4                   - modifikovaný Mobility framework
|   |-- networks
|       |-- ieee802154      - model IEEE 802.15.4
|       |-- results        - výsledky zo všetkých uskutočnených simulácií
|-- readme.txt              - popis adresárovej štruktúry
|-- text
|   |-- Lenart-thesis-2009.pdf - text bakalárskej práce vo formáte PDF
|-- zoznamCD
```

Obr. C.1: Výpis priloženého CD