Software Defined Monitoring: Nový prístup k monitorovaniu vysokorýchlostných počítačových sietí

Lukáš Kekely
Výpočetní technika a informatika, 1. ročník, prezenčná forma
Školiteľ: Jan Kořenek

Fakulta informačních technologií, Vysoké učení technické v Brně
Božetěchova 2, 612 66 Brno

ikekely@fit.vutbr.cz

Abstrakt. Neustále sa zvyšujúce rýchlosti liniek spolu s rastúcou významnosťou dát aplikačných protokolov pre monitorovanie vedú na nutnosť vytvoriť nový princíp hardvérovej akcelerácie spracovávania sieťových dát. V rámci dizertačnej práce *Softwarově řízené monitorování síťového provozu* je preto predstavený a skúmaný úplne nový koncept hardvérovej akcelerácie monitorovania sietí nazvaný *Software Defined Monitoring* (SDM). Základná myšlienka SDM je založená na úzkom prepojení softvérových monitorovacích aplikácií s výkonným hardvérovým akcelerátorom, ktorý predspracúva sieťové dáta. Softvérové aplikácie pritom môžu jednoducho ovládať stupeň detailov zachovávaných predspracovaním pre jednotlivé sieťové toky. Vďaka tomu je možné spracovanie menej zaujímavých dát prenechať akcelerátoru a v softvéri sa zamerať už len na podrobné spracovanie naozaj zaujímavých dát. Tým SDM umožňuje prakticky realizovať flexibilné monitorovanie s podporou podrobnej analýzy paketov aj na veľmi vysokých rýchlostiach – až 100 Gb/s.

Klúčové slová. FPGA, akcelerácia, monitorovanie, bezpečnosť, vysokorýchlostné siete

1 Úvod

Monitorovanie sieťových dát hrá jednu z kľúčových úloh pre oblasti správy a bezpečnosti moderných počítačových sietí. Dnes zaužívaným štandardom pre monitorovanie sietí je meranie na bázy sieťových tokov. Monitorovacie zariadenia zbierajú základné štatistiky o všetkých paketoch a agregujú ich do záznamov o tokoch. Tie zasielajú na centrálne úložisko (kolektor) pomocou protokolu NetFlow [1] alebo IPFIX [2]. V procese zberu a agregovania dát tak dochádza k istej strate informácií a kolektor (kde sa dáta ďalej analyzujú) má preto obmedzený pohľad na sieť. Z uvedeného dôvodu je aktuálnym trendom snaha rozširovať záznamy o tokoch pridaním nejakej informáciu navyše k základným veľkostným a časovým štatistikám. Pridaná informácia pritom často býva založená na dátach z aplikačných protokolov.

Implementáciu monitorovania obohateného o analýzu aplikačných protokolov je možné celú vytvoriť v softvéri. Priepustnosť takejto realizácie je však silne obmedzená výkonnosťou súčasných procesorov. Na druhej strane, čisto hardvérové riešenia majú slabú flexibilitu, z dôvodu náročnej hardvérovej realizácie komplexných analyzátorov aplikačných protokolov. Navyše nové bezpečnostné hrozby nestále vznikajú a je potrebné na ne dostatočne rýchlo reagovať, čo je pre hardvérové riešenia problémové. Uvedené zhodnotenie dvoch základných prístupov vedie na ideu vytvoriť niečo medzi, teda výkonný hardvérový akcelerátor spracovania dát plne kontrolovaný flexibilnými softvérovými aplikáciami. Práve softvérovému riadeniu vďačí navrhnutý koncept za označenie *Software Defined Monitoring* (SDM).

Úloha hardvérového akcelerátoru v SDM spočíva v redukcii objemu dát tečúcich k softvérovým aplikáciám tým, že nad zvolenými časťami dát realizuje analýzu hlavičiek paketov a prípadne aj ich agregovanie do tokov. Akcelerátor tak posiela zaujímavú časť paketov nedotknutých do softvéru na precíznejšiu analýzu, zatiaľ čo sám realizuje základné meranie na bázy tokov nad zvyškom dát. Navyše je podporované aj filtrovanie pre prípad, že aplikácie nepotrebujú agregované informácie o všetkých paketoch.

Výber spôsobu spracovania jednotlivých paketov v akcelerátore SDM je plne kontrolovaný monitorovacím softvérom a môže byť za behu prispôsobovaný aktuálnym potrebám konkrétnej aplikácie. Realizovaný je pomocou dynamicky sa meniacej množiny pravidiel nad sieťovými tokmi vytváranej aplikáciou na základe pozorovaných paketov. Uvedené pravidlá sú do akcelerátoru nahrávané jednotným rozhraním a každé určuje ako predspracovať ďalšie príchodzie pakety jedného konkrétneho toku. Vďaka jednotnosti ovládacieho rozhrania akcelerácie je systém flexibilný a je možné ho použiť na zvýšenie výkonnosti širokého spektra rôznych monitorovacích a bezpečnostných aplikácií.

Prínos práce prezentovanej v tomto príspevku je v troch oblastiach: (1) analýza dát z reálnej vysokorýchlostnej siete s ohľadom na rozhodnutie o vhodnosti akcelerácie založenej na popísanom koncepte SDM (sekcia 2); (2) rozpracovanie návrhu konceptu SDM pre vysokorýchlostné siete, čo zahŕňa návrh hardvéru (aplikačne špecifický procesor) aj jeho riadiaceho softvéru (sekcia 3); (3) implementácia a vyhodnotenie vlastností systému v niekoľkých prípadoch použitia (sekcia 4).

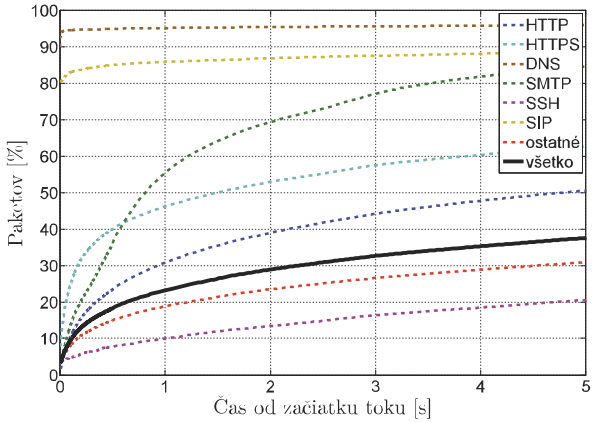
2 Analýza

Začiatok príspevku sa venuje analýze vlastností sieťových dát na reálnej vysokorýchlostnej sieti. Na základe zmeraných charakteristík je následne vytvorený podrobný návrh SDM systému tak, aby dosahoval optimálnu výkonnosť v reálnom nasadení. Všetky merania uvedené v celom príspevku sú realizované vo vysokorýchlostnej sieti CESNET2, ktorá má optické linky pracujúce na rýchlostiach do 100 Gb/s.

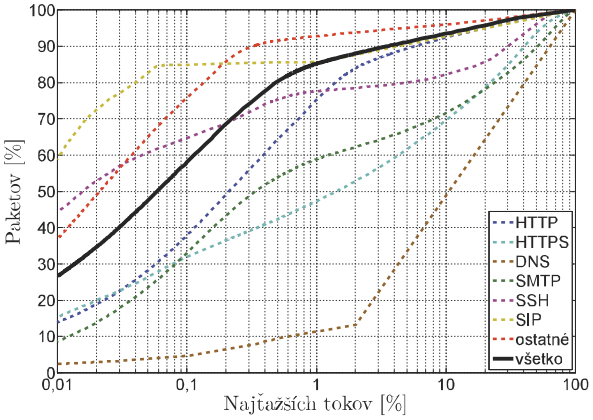
Pretože softvérové aplikácie rozhodujú o predspracovaní dát je časovanie príchodu paketov veľmi dôležité z pohľadu dosiahnuteľnej výkonnosti. Najlepší pohľad na časovanie paketov v tokoch je možné získať meraním relatívneho času príchodu paketov od začiatku toku. Čiže, prvý paket každého toku má nulový relatívny čas príchodu a jeho absolútny čas označuje moment začiatku toku. Potom relatívny čas príchodu každý nasledujúceho paketu je rozdiel absolútneho času jeho príchodu a poznačeného momentu začiatku toku. Výsledky popísaného merania sú zanesené v grafe na Obr. 1 , ktorý zobrazuje distribučné funkcie práve relatívnych časov príchodu paketov pre rôzne skupiny dát. Graf ukazuje, že všeobecne (čierna plná čiara) len malá časť paketov príde hneď po začatí toku (napr. len asi pätina paketov príde počas prvej sekundy tokov). To znamená, že aj ak bude oneskorenie softvérového riadenia pri zavádzaní pravidiel o tokoch relatívne vysoké, stále umožní pravidlami ovplyvniť spracovanie väčšiny paketov.

Ďalšou dôležitou vlastnosťou sieťových dát je charakter rozdelenia veľkostí tokov. Z grafu na Obr. 2 vidno, že podľa merania má distribúcia veľkostí tokov na reálnej sieti heavy-tailed charakter. Uvedený graf ukazuje podiel paketov prenesených istým percentom najťažších tokov. Je teda všeobecne (čierna plná čiara) vidno, že aj veľmi malé percento najťažších tokov prenáša významnú časť celkového počtu paketov (napr. 1 % tokov nesie až 85 % paketov). Z pohľadu navrhnutého SDM je tak možné aj zavedením len malého počtu pravidiel o tokoch zaistiť akceleráciu predspracovania väčšiny paketov.

Pre praktické využitie heavy-tailed charakteru v prospech výkonnosti SDM je ešte potrebné vyriešiť problém vhodného rozpoznania najťažších tokov. Presnejšie je problém definovaný ako schopnosť predpovedať, ktoré toky sú z najťažších len na základe pozorovania istých vlastností ich prvých paketov. Na riešenie uvedeného problému je možné použiť veľmi priamočiaru metódu: pre zvolenú hodnotu parametra označ za ťažký tok každý taký, o ktorom je už známe, že má aspoň paketov. Výhodou tejto jednoduchej metódy je nenáročnosť jej implementácie, pretože jedinou sledovanou vlastnosťou paketov je ich samotná existencia (netreba ich dodatočne analyzovať). Pritom aj takto jednoduchá metóda vedie na veľmi dobré výsledky rozpoznania ťažkých tokov z pohľadu konceptu SDM, ako je ukázané na grafoch 5 a 6 v sekcii s rozborom dosahovanej výkonnosti.



Obr. 1: Časovanie príchodu paketov v tokoch



Obr. 2: Heavy-tailed charakter dát

3 Architektúra

Ako už je spomenuté v úvode, základná myšlienka akcelerácie v SDM systéme spočíva v jemne kontrolovanej redukcii objemu dát dosiahnutej akcelerovaným predspracovaním paketov zo siete. Predspracovanie samotné je realizované v hardvéri, ale jeho použitie je plne kontrolované softvérovými aplikáciami. Práve preto, je niekoľko počiatočných paketov každého toku poslaných do softvéru, ktorý až na ich základe vyberie spôsob hardvérového predspracovania nasledujúcich paketov daného toku. Vhodné typy podporovaného predspracovania paketov pre oblasť monitorovania je možné rozdeliť do troch skupín:

Extrakcia zaujímavých informácií z paketov a posielanie len týchto informácií do softvéru v jednotnom formáte (unifikovaná hlavička - UH). Tým sa zníži jednak objem dát poslaných do softvéru, ale aj vyťaženie procesoru, pretože analýzu paketov realizoval už hardvér.

Agregovanie dát z paketov do záznamov o tokoch priamo v hardvéri vedúce na ešte vyššiu úsporu výkonnosti softvéru. Môžu pritom existovať rôzne spôsoby agregovania pre rôzne aplikácie.

Filtrovanie úplne nepotrebných paketov, čo môže napomôcť rôznym aplikáciám zameraným na pokročilú analýzu špecifickej podskupiny sieťových dát (napr. analyzátor HTTP).

Základnú konceptuálnu schému navrhnutého systému SDM je možné vidieť na Obr. 3. Dáta nesúce cesty sú značené plnými šípkami a kontrolné spätné väzby prerušovanými. Systém je zložený z dvoch častí (firmvér FPGA a softvér) prepojených dátovou zbernicou (napr. PCI Express). Dáta z firmvéru do softvéru prichádzajú po viacerých nezávislých kanáloch a to vo forme celých paketov, UH alebo záznamov o tokoch. Tieto dáta sú potom spracúvané užívateľom definovanou množinou monitorovacích a bezpečnostných aplikácií (napr. exportér tokov). Aplikácie, vo forme SDM zásuvných modulov, okrem čítania dát z vybraných kanálov môžu špecifikovať, ktoré toky sú pre ne nezaujímavé a môžu sa tak spracúvať vo firmvéri. Definície nezaujímavých tokov od všetkých aplikácií sú agregované v SDM radiči, ktorý na základe nich priamo konfiguruje predspracovanie vo firmvéri so snahou dosiahnuť maximálnu redukciu dát pri zachovaní dostatočnej úrovne detailov. SDM radič je tak jediným kontrolným prvkom celého systému, ktorý priamo riadi správanie sa firmvéru.

SDM firmvér začne spracovanie každého paketu jeho analýzou a extrakciou zaujímavých dát. Na základe extrahovaných dát a množiny pravidiel nakonfigurovaných od SDM radiča potom rozhodne o konkrétnom spôsobe predspracovania tohto paketu aj o smerovaní dát pre softvér do správneho kanálu. Podrobnejšie je možné spôsob realizácie akceleračné firmvéru SDM vidieť na Obr. 4. Popísané spracovanie paketov realizuje procesná zreťazená linka štyroch jednotiek. Dáta paketov pritom netečú priamo touto linkou, ale sú odložené v paralelnej FIFO pamäti. Celá konfigurácia procesnej linky je realizovaná cez špeciálnu jednotku, ktorá vie atomicky spravovať pravidlá priamo v pamäti vyhľadávacej jednotky. SDM firmvér je teda realizovaný piatimi jednotkami:

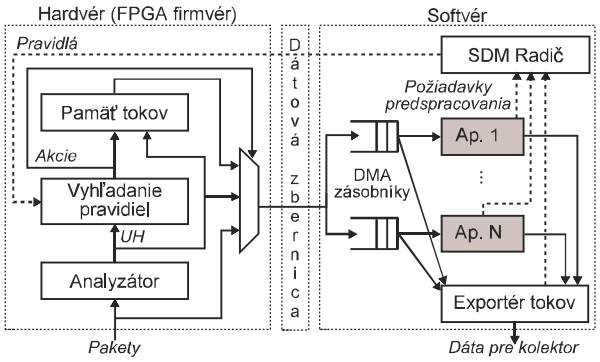
Analyzátor extrahuje zaujímavé informácie z hlavičiek paketov, najmä polia identifikujúce sieťový tok (IP adresy, čísla portov a protokol). Navyše je štruktúra analyzátora modulárna a umožňuje jednoduché pridanie ďalších analyzačných modulov (A1..An). Podrobnejšie analyzátor popisujem v [3, 4].

Hľadanie pravidiel s cieľom prideliť akciu (inštrukciu spracovania) každému paketu na základe identifikátoru toku a množiny softvérom nakonfigurovaných pravidiel. Efektívna implementácia je možná napríklad špeciálnou haš tabuľkou s kukučím hašovaním ako ukazujem v [5]. Ku tomu potrebné haš funkcie je ďalej možné v FPGA efektívne realizovať pomocou CRC ako uvádzam v [6, 7].

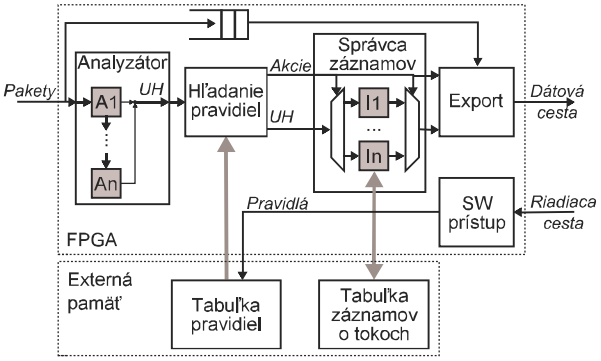
Správca záznamov spravuje stavové záznamy v tabuľke tokov. Stará sa hlavne o aktualizáciu ich hodnôt pomocou inštrukcií podľa paketom patriacich akcií. Každá akcia nesie okrem inštrukčného kódu aj adresu záznamu, na ktorú sa má aplikovať. Pri aktualizácii inštrukcie pracujú s aktuálnou hodnotou záznamu z pamäte aj s dátami z UH. Okrem aktualizačných inštrukcií podporuje jednotka aj špeciálnu inštrukciu exportovania (a nulovania) zvoleného záznamu, ktorá je iniciovaná na konci toku alebo v pravidelných intervaloch. Správcu záznamov je možné jednoducho rozširovať o nové inštrukčné moduly (I1..In). Túto problematiku podrobnejšie rozoberám napríklad v [8].

Export sa stará o smerovanie dát v správnom formáte a správnym softvérovým kanálom.

SW prístup je hlavným prístupovým bodom k SDM firmvéru zo strany softvéru. Zaisťuje správu pravidiel o tokoch a iniciuje export záznamov o tokoch na základe príkazov od SDM radiča.



Obr. 3: Konceptuálna schéma SDM systému

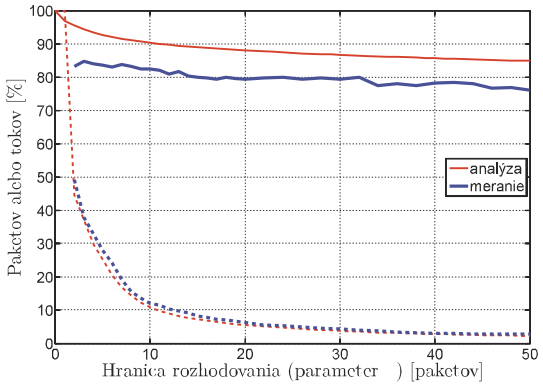


Obr. 4: Podrobnejšia schéma SDM firmvéru

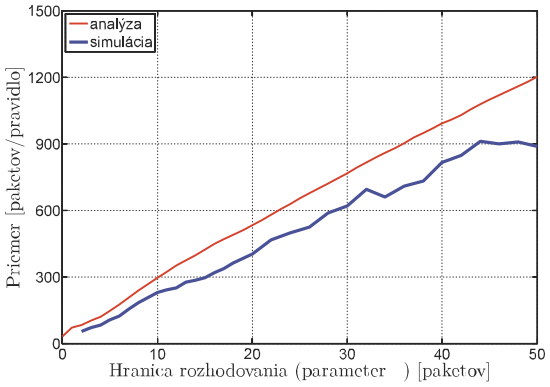
4 Výsledky

Navrhnutý SDM systém je implementovaný. Pritom je použitá akceleračná PCI Express karta s FPGA čipom rodiny Virtex-7 v troch variantoch sieťových rozhraní: 8 10 GbE, 2 40 GbE a 1 100GbE. Vo všetkých troch prípadoch pripadá na samotné funkčne jadro SDM len necelá štvrtina zdrojov firmvéru, ktorý celkovo zaberá necelú polovicu zdrojov čipu. Výkonnosť vytvorenej realizácie SDM je ďalej otestovaná s ohľadom na dosiahnuteľný stupeň akcelerácie.

Prvým testom je meranie percenta paketov, ktoré je SDM firmvér schopný spracovať na základe softvérom za behu vytvorených pravidiel o tokoch. Využité je pri tom pravidlo rozpoznania ťažkých tokov predstavené na konci sekcie 2. Výsledky testu sú zanesené v grafe na Obr. 5, ktorý ukazuje závislosti medzi hodnotou parametru a časťou tokov považovaných za ťažké (prerušovaná čiara) a paketov ich výberom pokrytých (plná čiara). Je vidno, že s rastúcou hranicou rozhodovania rapídne klesá podiel vybraný tokov, ale podiel nimi pokrytých paketov klesá len pozvoľna. To vedie na rast priemerného počtu paketov pokrytých jedným pravidlom (zisk zo zavedenia pravidla), ako ukazuje aj graf na Obr. 6. V grafoch tiež vidno rozdiel medzi analyticky a meraním zistenou efektivitou systému. Rozdiel (5 až 10 % paketov) je spôsobený istým časovým oneskorením zavádzania pravidiel ako reakcie na prvé pakety toku v reálnom systéme, ktoré nie je pri analytickom vyhodnotení brané do úvahy.



Obr. 5: Zachytené percento paketov alebo tokov



Obr. 6: Počet zachytených paketov na pravidlo

Ďalšie testy efektivity akcelerácie SDM sú realizované pre reálnejšie prípady nasadenia systému a ich výsledky sú zanesené v tabuľke 1. Testované je nasadenie SDM na akceleráciu piatich rôznych prípadov: (1) základné NetFlow monitorovanie tokov [1], (2) detektor skenovania portov, (3) detektor útoku Hartbleed na HTTPS protokol, (4) podrobná analýza aplikačného protokolu HTTP a (5) základné monitorovanie tokov obohatené o podrobnú analýzu HTTP. Hodnoty zanesené do tabuľky sú dvojakého typu – podiel využitia podporovaných typov hardvérového predspracovania a objem redukovaného dátového toku do softvéru v jednotlivých prípadoch nasadenia. Všeobecne vidno, že aplikácie zamerané na podrobnejšiu analýzu špecifických dát (2, 3, 4) využívajú hlavne filtrovanie. Naproti tomu, aplikácie vyžadujúce štatistické informácie o všetkých paketoch na linke (1) využívajú hlavne agregovanie. Nakoniec aplikácie nepracujúce priamo s dátami paketov (1, 2) používajú do istej miery aj extrakciu. Z posledných dvoch stĺpcov tabuľky vidno, že dosiahnutá redukcia záťaže softvéru oproti prípadu bez použitia SDM je relatívne vysoká – väčšinou ide o redukciu počtu paketov aspoň päťkrát a bajtov ešte viac.

Prípad použitia	**HW akcia [% paketov]**				**HW akcia [% bajtov]**				**SW záťaž [%]**	
	∅	**Ex**	**Ag**	**Fi**	**∅**	**Ex**	**Ag**	**Fi**	**Paketov**	**Bajtov**
NetFlow	–	20.55	79.45	–	–	12.03	87.97	–	20.66	0.98
Port sken	–	17.54	–	82.46	–	10.35	–	89.65	17.54	0.86
Heartbleed	4.91	–	–	95.09	3.77	–	–	96.23	4.91	3.77
HTTP	22.82	–	–	77.18	27.82	–	–	72.18	22.82	27.82
HTTP+NetFlow	23.34	10.56	66.10	–	28.50	3.63	68.87	–	34.02	29.00

Tabuľka 1: Využitie podporovaných typov hardvérového predspracovania v rôznych prípadoch použitia

5 Stav a ciele dizertačnej práce

Príspevok predstavil súčasný trend zvyšovania prenosových rýchlostí v počítačových sieťach vedúci na nutnosť výkonnejších monitorovacích a bezpečnostných systémov. Práve touto oblasťou sa zaoberám v rámci dizertačnej práce, kde som navrhol realizoval a základne testoval práve popísaný unikátny koncept flexibilnej akcelerácie monitorovania označený SDM. Zatiaľ čo konkurenčné postupy akcelerácie monitorovania sa spoliehajú buď na čisto hardvérové riešenia, ktorým chýba flexibilita alebo na čisto softvérové riešenia, ktorým zase chýba výkonnosť, predstavený koncept SDM predstavuje cestu vhodného spojenia hardvéru a softvéru pri zachovaní ich výhod a limitovaní ich nedostatkov. Základný koncept SDM ako je popísaný v tomto príspevku bol už publikovaný na IEEE konferencii INFOCOM [9] a prezentovaný na viacerých sieťových konferenciách (napr. *IETF Meeting* či *TERENA Networking*

Conference). Okrem toho boli publikované aj riešenia viacerých špecifických častí systému, ako sú odkazované priamo z textu príspevku. Aktuálne sa tiež o SDM pripravuje článok na vyžiadanie do časopisu *IEEE Transactions on Computers*. Prototyp systému je taktiež aktuálne v testovacom režime nasadený na produkčnej siete združenia CESNET a očakáva sa jeho skoré produkčné nasadenie. O SDM prejavila záujem aj komerčná firma Invea-Tech, ktorá ho chce zaradiť do svojho portfólia produktov.

V rámci ďalšieho smerovania dizertačnej práce sa chcem v priebehu druhého ročníka zaoberať hlavne experimentmi s akceleráciou rôznych aplikácií z oblasti monitorovania a bezpečnosti pomocou SDM na produkčnej sieti ako aj jeho ďalším rozširovaním a vylepšovaním. Pričom výsledky tohto snaženia plánujem priebežne publikovať. Nakoniec v treťom ročníku by som sa zameral na skonsolidovanie všetkých získaných výsledkov a spísanie finálneho textu dizertačnej práce.

6 Záver

Príspevok ukazuje návrh a implementáciu nového konceptu (systému) flexibilnej akcelerácie monitorovania vysokorýchlostných počítačových sietí. Uvádza tiež vybrané výsledky analýzy a testovania výkonnosti na dátach z reálnej siete, ktoré ukazujú, že vytvorený systém je schopný napomôcť monitorovaniu aplikačných protokolov na rýchlostiach liniek až do 100 Gb/s. Prezentované výsledky sú dosiahnuté v rámci dizertačnej práce na tému *Softwarově řízené monitorování síťového provozu*, ktorej ďalším pokračovaním bude prehlbovanie experimentálnych výsledkov z nasadenia na reálnej sieti a ďalšie vylepšovanie vlastností predstaveného konceptu SDM.

Poďakovanie

Príspevok vznikol čiastočne za podpory projektu VUT v Brne FIT-S-14-2297, projektu Centra excelencie IT4Innovations CZ.1.05/1.1.00/02.0070 a výskumného zámeru MSM 0021630528. Prezentovaná práca je tiež súčasťou projektu MŠMT "Velká infrastruktura CESNET" s číslom LM2010005.

Literatúra

[1] B. Claise: Cisco Systems NetFlow Services Export Version 9, RFC 3954, IETF, 2004

[2] B. Claise, B. Trammell, and P. Aitken: Specification of the IP Flow Information Export (IPFIX) Protocol for the Exchange of Flow Information, RFC 7011, IETF, 2013

[3] L. Kekely, V. Puš and J. Kořenek: Design Methodology of Configurable High Performance Packet Parser for FPGA, International Symposium on Design and Diagnostics of Electronic Circuits & Systems, IEEE, 2014

[4] L. Kekely, V. Puš and J. Kořenek: Low-Latency Modular Packet Header Parser for FPGA, Symposium on Architectures for Networking and Communications Systems, ACM, 2012, ISBN 978-1-4503-1685-9

[5] L. Kekely, M. Žádník, J. Matoušek and J. Kořenek: Fast Lookup for Dynamic Packet Filtering in FPGA, International Symposium on Design and Diagnostics of Electronic Circuits & Systems, IEEE, 2014, ISBN 978-1-4799-4558-0

[6] L. Kekely, T. Závodník and V. Puš: CRC based hashing in FPGA using DSP blocks, International Symposium on Design and Diagnostics of Electronic Circuits & Systems, IEEE, 2014

[7] L. Kekely, T. Závodník and V. Puš: Using DSP blocks to compute CRC hash in FPGA, International Symposium on Field-Programmable Gate Arrays, ACM, 2014, ISBN 978-1-4503-2671-1

[8] L. Kekely, V. Puš, P. Benáček and J. Kořenek: Trade-offs and Progressive Adoption of FPGA Acceleration in Network Traffic Monitoring, International Conference on Field Programmable Logic and Applications, IEEE, 2014

[9] L. Kekely, V. Puš, and J. Kořenek: Software Defined Monitoring of Application Protocols, The 33rd Annual IEEE International Conference on Computer Communications, IEEE, 2014, ISBN 978-1-4799-3360-0