

ANALÝZA DYNAMICKÝCH VLASTNOSTÍ SMĚROVACÍCH TABULEK

Jiří Matoušek

Výpočetní technika a informatika, 3. ročník, prezenční studium

Školitel: Jan Kořenek

Fakulta informačních technologií, Vysoké učení technické v Brně

Božetěchova 1/2, 612 66 Brno

imatousek@fit.vutbr.cz

Abstrakt. Současné požadavky kladené na směrování paketů v páteřních sítích vynucují akceleraci této operace v hardware s použitím paměťově efektivních technik reprezentace odpovídající sady prefixů. Ačkoliv je dynamická povaha směrovacích informací významným faktorem ovlivňujícím paměťovou reprezentaci sady prefixů, doposud nebyly dynamické vlastnosti směrování v páteřních sítích studovány z pohledu změn ve směrovací tabulce. Tento příspěvek obsahuje chybějící analýzu, která byla provedena jako první krok k návrhu a implementaci systému dynamického přidělování paměti na čipu FPGA pro potřeby reprezentace sady prefixů ze směrovací tabulky. V příspěvku je ukázáno, že směrovací tabulka obsahuje nezanedbatelné množství dlouhodobě stabilních záznamů. Změny ve směrovací tabulce pak připadají především na aktualizace záznamů, které jsou většinou provedeny za méně než 2 hodiny. V rámci analýzy bylo také ukázáno, že záznamy odstraňované ze směrovací tabulky byly v 70 % případů přidány před méně než 24 hodinami.

Klíčová slova. Směrovací tabulka, dynamické vlastnosti směrování, páteřní síť, LPM.

1 Úvod

Neustále se zvyšující množství dat přenášených přes počítačové sítě má přímý vliv na nárůst podporovaných přenosových rychlostí. Například pro Ethernet je již standardizována přenosová rychlost 100 Gb/s [1]. Počet koncových zařízení připojených k Internetu navíc dosahuje řádu miliard a nadále rychle narůstá. Oba tyto trendy se přitom přímo dotýkají jedné ze základních operací v počítačových sítích – směrování paketů. Směrovací tabulky pro páteřní síť dosahují velikosti přes 500 tisíc IPv4 a téměř 18 tisíc IPv6 záznamů (viz [2]), na základě kterých je při podpoře přenosové rychlosti 100 Gb/s nutné učinit rozhodnutí o směrování paketu za 6,72 ns. Směrování paketu již tedy není možné provádět v software, ale tato operace musí být v páteřních sítích implementována v hardware.

V rámci své disertační práce na téma *Využití rekonfigurovatelných obvodů v oblasti počítačových sítí* se zabývám použitím technologie FPGA pro implementaci operace vyhledání nejdelšího shodného prefixu (anglicky *longest prefix match*, LPM), která představuje výpočetně nejnáročnější součást procesu směrování paketů. Typická implementace operace LPM pro vysokorychlostní síť je založena na stromové datové struktuře kódující prohledávanou množinu prefixů. Kvůli požadavku na rychlé zpracování je vyhledávání v této stromové struktuře nejčastěji implementováno v hardware ve formě zřetězené linky, jejíž jednotlivé stupně zajišťují vyhledávání na různých hladinách stromové struktury. Pro uložení od-

povídajících hladin stromové struktury je každému stupni zřetězené linky přiřazen samostatný paměťový blok.

Implementace s využitím zřetězené linky zajišťuje dostatečnou rychlost vyhledávání, která je však limitována rychlostí přístupu do paměti sloužící k uložení stromové struktury. V [3] bylo ukázáno, že při vhodném zakódování množiny prefixů je možné k jejímu uložení využít dostupnou distribuovanou paměť na čipu FPGA a dosáhnout tak dostatečné rychlosti vyhledání pro podporu přenosové rychlosti 170 Gb/s.

Vzhledem k dynamické povaze množiny prefixů se však paměťové nároky v jednotlivých stupních zřetězené linky v čase mění. S ohledem na omezené množství dostupné paměti na čipu FPGA se proto jeví jako vhodné zajistit dostatečné množství paměti v jednotlivých stupních zřetězené linky prostřednictvím dynamického přidělování paměti. Tento příspěvek se tudíž zabývá analýzou dynamických vlastností směrovacích tabulek v páteřních sítích, která by následně měla být rozšířena o pohled na vztah změn ve směrovacích tabulkách a paměťových nároků v jednotlivých stupních zřetězené linky. Na základě těchto analýz by pak měla být navržena architektura umožňující dynamické přidělování paměti na čipu FPGA.

Struktura příspěvku je následující. V kapitole 2 jsou představeny příbuzné práce a je identifikován prostor pro provedenou analýzu. Následně jsou v kapitole 3 popsána data, na nichž byla analýza prováděna, a způsob jejich předzpracování. Popis vlastní analýzy a jejích výsledků je obsahem kapitoly 4. Provedená analýza je také v rámci kapitoly 5 zasazena do kontextu dalších cílů disertační práce. Příspěvek a dosažené výsledky jsou shrnuty v kapitole 6.

2 Příbuzné práce

Existuje mnoho prací zabývajících se dynamickou povahou směrování paketů z pohledu koncových síťových zařízení. Komplexní analýzu aktualizací směrovacích informací zasílaných mezi páteřními směrovači však najdeme pouze v [4] a [5].

Původní analýza [4] provedená na datech z roku 1996 je založená na sledování posloupnosti aktualizací zasílaných protokolem BGP pro danou dvojici (prefix, směrovač). Identifikované posloupnosti přidání a odebrání prefixu jsou klasifikovány do tří kategorií: 1) aktualizace spojené se změnou směrování, 2) aktualizace spojené se změnou směrovací politiky a 3) patologické aktualizace směrovacích informací. Z výsledků analýzy vyplývá, že 99 % BGP aktualizací zasílaných mezi páteřními směrovači spadá do kategorie patologických aktualizací. Tato práce také přináší pohled na některé kvantitativní vlastnosti zasílaných aktualizací směrovacích informací: a) množství zasílaných aktualizací závisí na zátěži sítě a kopíruje její denní, týdenní a roční vzory, b) aktualizace pro danou dvojici (prefix, autonomní systém) jsou registrovány převážně s periodou 30 a 60 s a c) 35-100 % dvojic (prefix, autonomní systém) je aktualizováno alespoň jedenkrát za den, přičemž medián je 50 %.

Revize závěrů původní analýzy po 10 letech byla publikována v [5]. V rámci této práce byly uplatněny stejné metody analýzy aktualizací směrovacích informací jako v [4], avšak u některých posloupností přidání a odebrání prefixu došlo k jejich zpřesnění. Díky tomu bylo možné aktualizace směrovacích informací přesněji klasifikovat do výše uvedených 3 kategorií. Základním zjištěním revidované analýzy je skutečnost, že z pohledu BGP aktualizací je Internet „zaměstnanější“ (legitimní aktualizace spojené se změnou směrování či směrovací politiky představují 84 % všech aktualizací) a „zdravější“ (podíl patologických aktualizací je jen 16 %). Ostatní zjištění revidované analýzy potvrzují výsledky z před desíti let a pouze u periody aktualizací pro danou dvojici (prefix, autonomní systém) se kromě významného podílu periody 30 s objevuje také značné množství aktualizací s periodou větší než 8 hodin.

Obě uvedené práce se zabývají především klasifikací aktualizací směrovacích informací a také některými jejich kvantitativními vlastnostmi. Z pohledu implementace operace LPM je však mnohem podstatnější vliv přijatých BGP aktualizací na samotnou směrovací tabulku, která se nemění s přijetím každé aktualizace. Navíc ne všechny změny směrovací tabulky znamenají změnu prefixu uloženého v tabulce,

a tudíž datové struktury reprezentující množinu prefixů pro potřeby operace LPM. Tento příspěvek se proto zabývá analýzou vlivu aktualizací směrovacích informací na směrovací tabulku, přičemž jejich vliv na reprezentaci prefixové sady pro potřeby operace LPM ponechává k rozpracování do budoucna.

3 Vstupní data a jejich předzpracování

Pro provedenou analýzu byla vybrána data z kolektoru RRC14 umístěného v Palo Alto (USA). Přístup k použitým datům je umožněn prostřednictvím služby RIS (*Routing Information Service*) raw data [6]. Použitá data pocházejí z období 1.10.2013 až 28.2.2014. V tomto období se počet záznamů ve směrovací tabulce zvýšil o 3,2 % z původních 493 219 až na 508 881 záznamů.

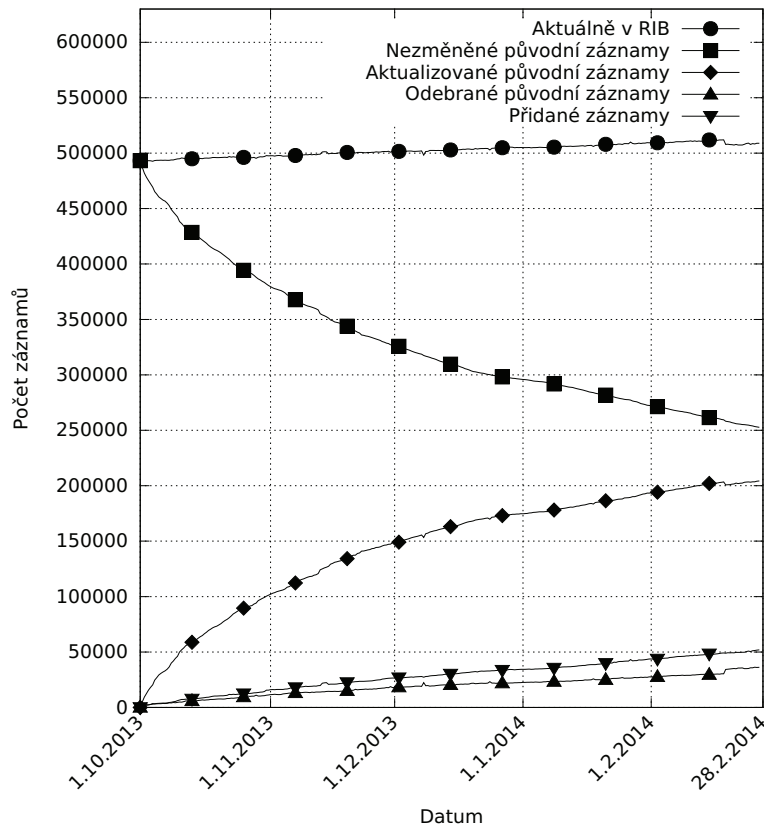
Na kolektoru RRC14 jsou data ze směrovacích tabulek k dispozici ve formě obrazů celých směrovacích tabulek s 8hodinovými rozestupy a také v podobě souhrnu všech BGP zpráv v intervalech po 5 minutách. Pro provedenou analýzu byla použita data v obou formátech. Celkový obraz směrovací tabulky sloužil jako výchozí stav, na který byly postupně aplikovány aktualizace (přidání či odebrání záznamu) obsažené v souhrnu BGP zpráv v jednotlivých 5minutových intervalech. Je však třeba poznamenat, že ne každé přidání či odebrání záznamu předepsané v BGP zprávě se projevilo jako přidání či odebrání záznamu z rekonstruované směrovací tabulky. Identifikátorem záznamu je totiž kromě prefixu IP adresy také identifikace zdroje této informace. Nový záznam byl tedy zaveden pouze v případě, že pro daný prefix nebyla přítomná směrovací informace ze žádného jiného zdroje. Odebrání záznamu pak nastávalo pouze v případě, že už pro daný prefix neexistovala směrovací informace ze žádného dalšího zdroje.

Protože záznamy v obrazu směrovací tabulky obsahují informaci o času přidání a BGP zprávy si s sebou také nesou informaci o času přijetí, je možné výše uvedeným postupem zrekonstruovat stav směrovací tabulky v libovolném okamžiku uvažovaného časového intervalu (s rozlišením na sekundy).

4 Analýza dynamických vlastností směrovacích tabulek

První částí analýzy bylo sledování vývoje směrovací tabulky (anglicky *routing information base*, *RIB*) v celém uvažovaném intervalu. Výsledky sledování jsou zobrazeny v grafu na obrázku 1, ze kterého je patrný nárůst velikosti směrovací tabulky (viz „Aktuálně v RIB“) způsobený větším počtem přidávaných záznamů než odebraných původních záznamů. Z pohledu vytyčeného cíle analýzy je však nejpodstatnější informací celkové množství změněných záznamů. Oproti stavu na začátku října 2013 se na konci února 2014 vyskytovalo ve směrovací tabulce přibližně 50 000 nově přidávaných záznamů a více než 36 000 původních záznamů bylo odebráno. Největší podíl změn (přes 200 000) však připadá na aktualizace původních záznamů, které reprezentují jednu či více posloupností odebrání a následného navrácení záznamu do směrovací tabulky. Z grafu je také na poklesu počtu nezměněných původních záznamů patrné, že během 5 měsíců došlo k obměně (odebrání či aktualizaci) téměř poloviny původních záznamů směrovací tabulky.

Po zjištění trendů vývoje změn ve směrovací tabulce v průběhu celého sledovaného období byly tyto trendy sledovány v průběhu jednoho dne. Histogramy na obrázku 2a zobrazují počty změn v jednotlivých částech dne v průměru za 5 měsíců sledovaného období. Z uvedeného grafu je patrné větší množství změn záznamů ve směrovací tabulce během pracovních hodin. Uvedené průběhy také potvrzují zjištění z první části analýzy, že největší podíl změn připadá na aktualizace záznamů. Hlavním zjištěním je však skutečnost, že aktualizace (tj. odebrání a následné navrácení záznamu do směrovací tabulky) často proběhne za méně než 2 hodiny. Tuto vlastnost lze využít při implementaci operace LPM, konkrétně při aktualizaci prefixu v datové struktuře. Vzhledem k předpokládanému opětovnému přidání odebíraného prefixu během následujících 2 hodin není nutné jeho skutečné odebrání, ale je možné jej realizovat nastavením příznaku neplatnosti, který bude s opětovným přidáním prefixu vynulován.



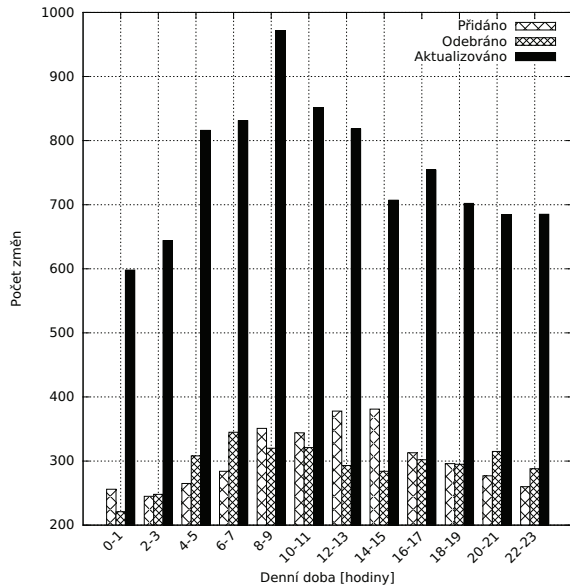
Obrázek 1: Vývoj směrovací tabulky v období 5 měsíců

Poslední část analýzy byla věnována sledování délky výskytu záznamu ve směrovací tabulce. Výsledky tohoto pozorování, zobrazené na obrázku 2b, se vztahují pouze na záznamy, které byly ze směrovací tabulky odstraněny, a tudíž u nich byla známa celá délka jejich výskytu v RIB. Na rozdíl od předcházejících částí analýzy byl každý výskyt opakovaně přidávaného a odebíraného prefixu započítán samostatně a jeden prefix tudíž mohl přispět k výsledkům analýzy několika různými hodnotami délky výskytu v RIB. Graf na obrázku 2b představuje kumulativní funkci délky výskytu záznamu v RIB a udává podíl záznamů, které byly ve směrovací tabulce přítomny uvedenou či kratší dobu. Z grafu lze tedy například vyčíst, že 70 % odstraněných záznamů se ve směrovací tabulce vyskytovalo méně než 1 den. V rámci sledovaného 5měsíčního období byl nejdelší výskyt záznamu 284 dnů a v průměru byly záznamy ze směrovací tabulky odstraňovány za 9 dnů a 4 hodiny. Medián délky výskytu záznamu v RIB byl ale 35 minut a lze tedy konstatovat, že více než polovina odebraných záznamů stráví ve směrovací tabulce méně než 1 hodinu.

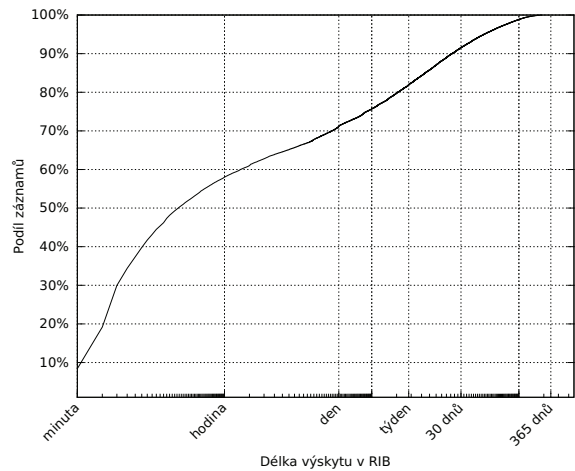
5 Cíle disertační práce

Aktuální situace v páteřních sítích klade vysoké požadavky na implementaci operace směrování paketů. Kvůli podpoře vysokých přenosových rychlostí je nutné akcelarovat směrování paketů v hardware. Velké množství záznamů ve směrovacích tabulkách také vynucuje využití speciálních paměťově efektivních reprezentací sady prefixů ze směrovací tabulky a efektivní nakládání s přidělenou pamětí během provádění aktualizací směrovacích informací.

V rámci své disertační práce se zabývám možnostmi akcelerace operace LPM s využitím technologie FPGA. Požadovanou rychlost zpracování lze poměrně snadno zajistit implementací operace LPM



(a) Změny směrovací tabulky během dne



(b) Délka výskytu záznamů ve směrovací tabulce

Obrázek 2: Vývoj směrovací tabulky v různých časových intervalech

zřetězenou linkou. Problematickým je však dostatečně rychlý přístup do paměti. V první části disertační práce jsem se proto věnoval analýze stávajících LPM algoritmů, především z hlediska jejich paměťové náročnosti při reprezentaci prefixových sad z aktuálních směrovacích tabulek páteřních směrovačů. Na základě této analýzy jsem následně navrhl novou reprezentaci prefixových sad a hardwarovou architekturu pro její zpracování. Tato reprezentace umožňuje uložit kompletní prefixové sady z páteřních směrovacích tabulek v rychlé paměti na FPGA čipu. Navržená hardwarová architektura podporuje propustnost přes 170 Gb/s. Výsledky dosažené v rámci první části disertační práce jsem publikoval v [3].

Dynamickou povahu směrovací tabulky je nutné reflektovat při jejím uložení v paměti pro potřeby operace LPM. Vzhledem k použití zřetězené linky a oddělených paměťových bloků v jejich jednotlivých stupních by statické přidělení rezervní paměti mohlo vést k jejímu neefektivnímu využití. Proto jsem se v druhé části disertační práce nejprve zaměřil na analýzu dynamických vlastností směrovacích tabulek, jejíž popis je obsahem tohoto příspěvku. Provedená analýza navazuje na podobné práce v této oblasti, přičemž posouvá předmět zájmu z aktualizací vyměňovaných mezi směrovači na samotnou směrovací tabulku. Znalosti získané touto analýzou a sada skriptů vytvořená při jejím sestavování následně poslouží k bližšímu pohledu na vztah změn ve směrovacích tabulkách a paměťových nároků v jednotlivých stupních zřetězené linky. Aktualizace směrovací tabulky budou transformovány na aktualizace stromové datové struktury reprezentující sadu prefixů a na jednotlivých hladinách datové struktury budou sledovány změny v paměťových nárocích během provádění aktualizací. S touto znalostí se pak budu věnovat návrhu a implementaci systému dynamického přidělování paměti na čipu FPGA, jehož vytvoření je hlavním cílem druhé části disertační práce.

6 Závěr

Zatímco se existující studie dynamických vlastností směrování v páteřních sítích zaměřovaly především na analýzu aktualizací zasílaných mezi páteřními směrovači a jejich klasifikaci do kategorií odpovídajících příčinám těchto změn, žádná práce se nevěnovala vlivu aktualizací na směrovací tabulku z pohledu implementace směrování v páteřním směrovači. Tento příspěvek proto představuje analýzu poskytující chybějící pohled na dynamickou povahu směrování v páteřních sítích.

Provedená analýza ukázala, že obměna poloviny směrovací tabulky trvá více než 5 měsíců a nezanedbatelná část záznamů je tudíž dlouhodobě stabilních. Hlavní příčinou změn ve směrovací tabulce jsou aktualizace stávajících záznamů (jejich odebrání a opětovné přidání), u nichž se při bližším pohledu ukázalo, že velmi často proběhnou za méně než 2 hodiny. Tato vlastnost by mohla být využita při implementaci aktualizací záznamů, například pomocí příznaku platnosti. V rámci analýzy bylo také ukázáno, že se 70 % odstraňovaných záznamů vyskytovalo ve směrovací tabulce méně než 1 den.

Výsledky analýzy představené v tomto příspěvku budou doplněny o studii vztahu změn ve směrovacích tabulkách a paměťových nároků v jednotlivých stupních zřetěžené linky implementující operaci LPM. Tyto informace následně poslouží pro návrh a implementaci systému dynamické alokace paměti na čipu FPGA v rámci disertační práce *Využití rekonfigurovatelných obvodů v oblasti počítačových sítí*.

Poděkování

Tato práce byla podpořena Evropským fondem regionálního rozvoje (ERDF) v rámci projektu Centra excelence IT4Innovations (CZ.1.05/1.1.00/02.0070) a dále projektem Architektury paralelních a vestavěných počítačových systémů (FIT-S-14-2297).

Reference

- [1] IEEE Computer Society: Part 3: Carrier Sense Multiple Access with Collision Detection (CSMA/CD) Access Method and Physical Layer Specifications; Amendment 4: Media Access Control Parameters, Physical Layers, and Management Parameters for 40 Gb/s and 100 Gb/s Operation. IEEE std 802.3ba-2010, June 2010. ISBN 978-0-7381-6322-2.
- [2] (2014, Jun.) IPv6 / IPv4 Comparative Statistics. [Online]. Available: <http://bgp.potaroo.net/v6/v6rpt.html>
- [3] J. Matoušek, M. Skačan, and J. Kořenek: Memory Efficient IP Lookup in 100 Gbps Networks. In 23rd International Conference on Field Programmable Logic and Applications (FPL'13), Porto: IEEE Circuits and Systems Society, 2013. ISBN 978-1-4799-0004-6.
- [4] C. Labovitz, R. G. Malan, and F. Jahanian: Internet Routing Instability. IEEE/ACM Transactions on Networking, vol. 6, no. 5, pp. 515–528, October 1998, ISSN 1063-6692.
- [5] J. Li, M. Guidero, Z. Wu, E. Purpus, and T. Ehrenkranz: BGP Routing Dynamics Revisited. ACM SIGCOMM Computer Communication Review, vol. 37, no. 2, pp. 5–16, April 2007, ISSN 0146-4833.
- [6] (2014, Jun.) RIS Raw Data – RIPE Network Coordination Centre. [Online]. Available: <http://www.ripe.net/data-tools/stats/ris/ris-raw-data>