Review of the PhD. thesis

Packet Classification Algorithms

submitted by

Ing. Viktor Puš

The main goal of the doctoral thesis submitted by Ing. Viktor Puš has been to design an algorithm for FPGA and ASIC hardware, able to support the IP packet classification for the links with the speed of 100 Gbps.

Chapters 2 till 5 are dedicated to related problems, solution methods and results of their applications. The Chapter 2 describes IP network technologies, which are supporting security by filtration of packet flows. Formal description of classification methods, based on a set of parameters, i.e. for IP packets, on their set of source and destination addresses, protocols and their ports used, are presented in the Chapter 3. The Chapter 4 is oriented to the overview of several algorithms designed for classifying IP packets. Finally, comparison of some known packet classification algorithms is presented in the Chapter 5.

The contribution of the research work of Ing. Viktor Puš has been the design of four new packet classification algorithms, presented in the text of the thesis:

> Perfect Hashing Crossproduct Algorithm (PHCA),
> Prefix Filtering Classification Algorithm (PFCA),
> Prefix Coloring Classification Algorithm (PCCA),
> Multi-Subset Prefix Coloring Classification Algorithm (MSPCCA).

All the algorithms detect the source and destination IP addresses, protocol type, and source and destination ports. They are tested on a small number of rule sets (synth1 - rules4) that differ mainly in number of protocols and in number of source and destination ports (Table 5.1) and in the number of LPM rules (rule1, rule4). It would be worth to present some reason for differences in these sets.

PHCA algorithm - Chapter 6

Although usage hashing function for LPM vector can generate a rule number candidate for some packets, which are not corresponding to the selected rule, comparison of selected candidate rule with the start packet fields requires only a simple additional step.

I would like to obtain a better explanation of the Algorithm 6.1. A method of selection of hash functions f_1 and f_2 is not described in the text, the small note saying that both functions f_1 and f_2 are based on a CRC32 should be better explained.

The sufficient number of vertices is obtained in the cycle rather quickly (assuming the resulting medium number of c around 2.5 in the tests). In the same number of steps hash functions f_1 and f_2 are selected as well. I certainly appreciate some description of these hash functions, which are taking LPM vector and correct rule number as parameters. Implementation of f_1 and f_2 is based on CRC (specifically CRC-32 is mentioned in the text), example in pg.45 seems that 3-bit CRC has been used. If f_1 and f_2 are really of a CRC type it would be interesting to discuss selection of generator polynomials and use of *don't care* bits.

PFCA algorithm - Chapter 7

The PFCA algorithm only a slightly modifies PHCA by reduction of the number of pseudorules by taking into account existence of universal conditions in rules. All small two-dimensional classifications' examples are giving preference to the first dimension and the chapter starts with source ID addresses as the primary fields for the generalization stage. Well, such an approach corresponds to the sequence of packet fields in all rule sets and in all selection algorithms (packet heads in the figures). Is it the best sequence of fields for the IP packet classification?

PCCA algorithm - Chapter 8

The PCCA algorithm is based on the idea of reduction pseudorule sets by coloring vertexes in rule sets and by using all prefixes in modified LPM (instead of the longer one) and selecting only these combinations which use the same color in Color Processing step.

The algorithm decreases the size of the memory required by vertex table (increased by memory required by LPM modules) in comparison to PHCA.

MSPCCA algorithm - Chapter 9

Finally, the MSPCCA algorithm combines approaches of the known MSCA algorithm and the PCCA designed by Ing. Puš and his colleagues. Bloom filters' module (utilized in MSCA) is placed between the Color Processing and Perfect Hash Table utilized in PHCA, PFCA and PCCA algorithms.

As a result of smaller number of the rules and pseudorules in comparison with all others algorithms (MCSA, PHCA, PFCA and PCCA) the algorithm requires the smaller memory for almost all tests (Figure 10.1).

The algorithms MSCA, PHCA, PFCA, PCCA and MSPCCA are compared in the second part of the last Chapter 10. The Table 10.2 and Figure 10.1 present memory required and Table 10.3 presents a specific ratio (bits per rule / megapackets per second). In any case, it would be better to present the information about the throughput and about the memory size for all these algorithms (although the memory request in kbits for algorithms are included in their chapters).

Since the classification algorithms have been designed for FPGA circuits, and the PHCA has been implemented on Xiling Virtex-5, it would be interesting to present an example of the mapping of algorithm(s) to FPGA circuit(s).

To conclude, although I had some notes to the text of the thesis, I can assert, that the thesis solves important and timely topics, it presents well the ability of its author, Ing. Viktor Puš, to work in the very interesting area of the packet classification required for security support in network communication, and that the results presented in the thesis could create a good area for the future research and implementation of the packet classification algorithms in FPGA hardware.

I can recommend the thesis, in the interpretation of the Law 111/98 in the Digest, for the defence and judge the candidate worthy to be awarded by the PhD. degree in Department of Computer Systems at the FIT BUT in Brno.

Prague, July 26th, 2012

doc. Ing. Jan Janeček, CSc.