

RECENZJA

rozprawy doktorskiej Pana mgra inż. Macieja Czekaja
pt. „**Hardware acceleration of traffic classifiers
for high throughput Ethernet**”

opracowana na zlecenie

Rady Dyscypliny Informatyki Technicznej i Telekomunikacji
Akademii Górniczo-Hutniczej z dnia 18.10.2021 roku

1. Ocena wyboru tematu rozprawy

Tematyka rozprawy doktorskiej związana jest ze sprzętowymi klasyfikatorami ruchu sieciowego umożliwiającymi klasyfikację danych jednostek protokołowych dla różnych warstw w modelu ISO/OSI. Klasyfikacja ruchu i strumieni danych umożliwia rozdzielanie ruchu przychodzącego np. w dużych centrach przetwarzania danych lub w chmurach obliczeniowych. Zagadnienia związane z oceną i klasyfikacją ruchu sieciowego mają bardzo duże znaczenie praktyczne. Problematyka klasyfikacji ruchu sieciowego jest znana od wielu lat. Zostało w tym zakresie zaproponowanych i wdrożonych wiele rozwiązań, które były adekwatne do potrzeb rynku.

W ciągu ostatnich dwóch lat w czasie pandemii COVID-19 ocenia się, że ruch sieciowy wzrósł od 35% do 40%. Zmiany w przyzwyczajeniach i sposobie pracy dużej części społeczeństwa mogą być trwałe i zapotrzebowanie na szybko dostępne usługi sieciowe znacząco wzrośnie w najbliższych latach. Kolejnym aspektem są trwające obecnie przekształcenia związane ze światową energetyką i przejściem wytwarzania energii z paliw kopalnianych na odnawialne źródła energii. Wiąże się to obecnie ze zmianami technologicznymi i podniesieniem cen prądu, co ma bardzo duży wpływ na koszty działania i rentowność centrów obliczeniowych. Ponadto w ostatnich latach odnotowuje się bardzo znaczący wzrost cen półprzewodników. Związane jest to z ogromnym popytem na tzw. elektronikę oraz ze wzrostem cen surowców w tym głównie metali ziem rzadkich.

Wskazane trzy aspekty tj. wzrost zapotrzebowania na usługi sieciowe, wzrost cen energii i wzrost cen półprzewodników wpływają znacząco na potrzebę zmian w podejściu do projektowania urządzeń sieciowych. Obecnie wytwarzane urządzenia powinny być tanie w

Sekretariat Rady Dyscypliny
Informatyka Techniczna i Telekomunikacja

1

09-12-2021

WIK
data wpływu

produkcji, wydajne i energooszczędne. Powoduje to potrzebę opracowania nowej klasy rozwiązań spełniającej stawiane wymagania.

Praca doktorska napisana została w języku angielskim. Temat rozprawy: „*Hardware acceleration of traffic classifiers for high throughput Ethernet*”, (pol. „*Sprzętowa akceleracja klasyfikatorów ruchu w sieci Ethernet o wysokiej przepustowości*”) dotyczy problematyki budowy i oceny klasyfikatorów ruchu w sieciach komputerowych ze szczególnym uwzględnieniem zagadnień związanych z zastosowaniem wydajnych, tanich i energooszczędnych technologii.

Tematykę rozprawy uważam więc za aktualną i nowoczesną. Szerokie spektrum problemów, które pojawiły się w trakcie realizacji pracy, jak również aktualność tematyki oraz jej duża ważność praktyczna gwarantują, że zainicjowane w pracy badania mogą być kontynuowane w przyszłości.

2. Hipoteza badawcza rozprawy

Praca Pana mgr inż. Macieja Czekaja jest wynikiem studiów prowadzonych przez Doktoranta nad zagadnieniem sprzętowej akceleracji klasyfikatorów ruchu sieciowego, a w szczególności znaczeniem optymalizacji pamięci klasyfikatorów pakietów z wykorzystaniem pamięci podręcznej.

Przedmiotem badań zaprezentowanych w pracy było przeanalizowanie dostępnych wybranych logów ruchu sieciowego, przeprowadzenie badań symulacyjnych umożliwiających zaproponowanie odpowiedniego rozwiązania, a następnie wykonanie projektu fragmentu urządzenia sieciowego.

W pracy, na stronie 13, postawiona została teza badawcza:

The optimal use of cache memory in packet classifiers leads to significant performance improvements and power savings. (pol. *Optymalne wykorzystanie pamięci podręcznej w klasyfikatorach pakietów prowadzi do znacznej poprawy wydajności i oszczędności energii.*)

Cel pracy i problem badawczy nie zostały jasno sformułowane.

Zagadnienia naukowe, jakie Autor postawił sobie do rozwiązania, zostały sformułowane w sposób mało precyzyjny i nie obejmują one całości zagadnień poruszanych i rozwiązanych w ramach rozprawy. Brak bezpośredniego wskazania celu pracy lub problemu badawczego utrudnił ocenę pracy.

Dodatkowo samo użycie słowa *teza*, w przypadku przedstawionego w pracy na stronie 13 zdania nie jest trafne. Mamy tu do czynienia raczej z hipotezą pracy.

3. Ocena merytoryczna rozprawy

Przedstawiona rozprawa została przygotowana jako praca doktorska. Rozprawa składa się z siedmiu rozdziałów, słownika, listy akronimów oraz spisu literatury. Ogólna objętość rozprawy to 110 stron.

Recenzowaną rozprawę zaliczam do grupy prac badawczo-projektowych. Nie ulega wątpliwości, że Doktorant osiągnął wysoki stopień opanowania teorii i praktyki w zakresie analizy i klasyfikacji ruchu sieciowego, budowy klasyfikatorów sieciowych i ich projektowania.

W rozdziale pierwszym pracy przedstawione zostało uzasadnienie, wskazujące na potrzebę zajęcia się wybraną tematyką oraz sformułowana została hipoteza pracy.

Rozdział drugi oraz trzeci stanowią przegląd literatury. W rozdziale drugim omówione zostały zagadnienia procesu przetwarzania danych w sieciach telekomunikacyjnych i centrach danych. Omówione zostały zagadnienia protokołów sieciowych, jak również sprzętowych rozwiązań wspierających obsługę. W rozdziale trzecim przedstawione i ocenione zostały zagadnienia dotyczące klasyfikacji ruchu, a w szczególności klasyfikacji pakietów, segmentów i datagramów przesyłanych w sieci (protokoły warstwy sieciowej i transportowej). Opisane zostały zagadnienia związane z zastosowaniem pamięci typu TCAM i SRAM w sprzętowych klasyfikatorach ruchu sieciowego. Zaprezentowana została również wstępna koncepcja wykorzystania pamięci podręcznej typu SRAM umożliwiającej ograniczenie wielkości drogiej w zastosowaniu i energochłonnej pamięci TCAM.

Z naukowego punktu widzenia wartościowym rozdziałem jest rozdział czwarty, w którym Autor przedstawił badania nad próbkami ruchu sieciowego uzyskanymi z ogólnie dostępnych logów. Wyniki badań wykazały, że zastosowanie pamięci podręcznej może dać dobre rezultaty. Dla potwierdzenia tej tezy opracowany został program symulacyjny i wykonane zostały badania, które potwierdziły zasadność stosowania pamięci podręcznej oraz wskazujące optymalne parametry dla rozwiązania sprzętowego.

Rozdział piąty zawiera projekt i szczegóły konstrukcji wybranych elementów inteligentnej karty sieciowej Kornik Ethernet 400Gb/s. Ponieważ doktorat prowadzony był jako tzw. *doktorat wdrożeniowy* fundowany przez firmę, w której pracuje Autor pracy, dlatego też

wyniki prac zostały zawarte w projekcie planowanej do drożenia karty Kornik. Sam projekt wykonany przez Autora uznaję za prace inżynierskie zrealizowane na wysokim poziomie. Niemniej pokazano w tym rozdziale zużycie zasobów sprzętowych oraz energii dla różnej wielkości pamięci podręcznej. Głównym atutem naukowym tego rozdziału jest więc luźne powiązanie sprzętowej implementacji oraz wyników badawczych teoretycznych otrzymanych w rozdziale czwartym i szóstym. Szkoda, że w pracy nie powiązано bezpośrednio wyników rozdziału czwartego i piątego i na przykład nie przeprowadzono dyskusji na temat optymalnej wielkości pamięci podręcznej z uwzględnieniem zużycia energii czy też równoległej pracy modułów sprzętowych (maksymalizacja przepustowości przy ograniczonych zasobach sprzętowych). Niemniej ze względu na zmienny charakter ruchu sieciowego uzyskanie takich wyników nie byłoby do końca miarodajne.

Kolejny rozdział, szósty, również jest wartościowy z naukowego punktu widzenia. Poświęcony został zagadnieniom zastosowania pamięci podręcznej w obsłudze strumieni danych zawierających pakiety IP po defragmentacji. W celu zamodelowania strumieni danych zawierających odpowiednią ilość pakietów po defragmentacji, Autor pracy opracował program symulacyjny. Należy tu podkreślić, że wykonanie wskazanego programu jest własnym osiągnięciem Doktoranta, ponieważ nie są opisywane w literaturze, ani dostępne na rynku symulatory posiadające wskazane funkcjonalności. Badania wykonane z wykorzystaniem opracowanego stanowiska badawczego wskazują, że rozważana w rozdziale 4 organizacja pamięci podręcznej również może być wykorzystana i może przyspieszyć obsługę strumieni sieciowych zawierających pakiety IP po defragmentacji.

Pracę kończy rozdział siódmy, który stanowi podsumowanie rozprawy. W rozdziale tym zawarto również wskazania na dalsze możliwe kierunki badań.

Załączony obszerny wykaz cytowanej literatury (133 pozycje) dobrze odzwierciedla stan badań prowadzonych we wskazanym zakresie na świecie i w Polsce. Dodatkowo zawartość rozdziału drugiego oraz trzeciego wskazują, że Autor posiada bardzo dużą wiedzę z zakresu prowadzonych badań w dyscyplinie Informatyka techniczna i telekomunikacja.

Na podstawie przedstawionego skrótowego omówienia treści całej rozprawy należy odnotować, że Autor efektywnie rozwiązał postawiony problem z zastosowaniem właściwych metod badawczych.

Oceniając pracę pragnę podkreślić, iż jest wartościowa z punktu widzenia pogłębienia wiedzy w dyscyplinie. Wnosi ona także oryginalny wkład naukowy i potwierdza wysokie

kwalifikacje Autora rozprawy. Należy również podkreślić duże praktyczne znaczenie opracowanych rozwiązań, stwarzające istotne potencjalne możliwości ich zastosowania.

Do najważniejszych osiągnięć uzyskanych w rozprawie można zaliczyć:

- wykonanie przez Autora analizy ruchu sieciowego i opracowanie z wykorzystaniem badań symulacyjnych efektywnego systemu wykorzystującego pamięć podręczną wspierającą proces klasyfikacji pakietów IP,
- wykonanie symulatora pozwalającego na prowadzenia badań nad strumieniami pakietów IP zawierających zadaną ilość pakietów po defragmentacji,
- wykonanie badań wskazujących na aplikowalność opracowanych z zakresu pamięci podręcznej rozwiązań do przetwarzania strumieni sieciowych zawierających pakiety IP po defragmentacji.

4. Uwagi ogólne i szczegółowe

Praca została napisana poprawnym językiem. Pewnym utrudnieniem w czytaniu pracy jest zastosowanie bardzo dużej liczby skrótów oraz posługiwanie się pojęciami niezdefiniowanymi wcześniej, które mają wiele definicji w obrębie dziedziny.

Przedstawione poniżej uwagi ogólne i szczegółowe nie obniżają mojej pozytywnej oceny pracy.

W trakcie zapoznawania się z treścią pracy nasunęło mi się kilka pytań i uwag dyskusyjnych, do których prosiłbym o komentarz ze strony Doktoranta.

1. Proszę o sformułowanie celu pracy i celów szczegółowych odzwierciedlających wykonane w ramach rozprawy prace.
2. Proszę o sformułowanie problemu badawczego rozwiązywanego w pracy.
3. Czy opracowane w ramach rozprawy rozwiązania wykorzystujące pamięć podręczną będą wspierały nowe protokoły sieciowe takie jak np. QUIC lub SCTP?
4. W rozdziale 6.3.2 opisana została szczegółowa architektura stacji roboczej, na której przeprowadzone zostały badania symulacyjne. W jaki sposób opisana architektura wybranego komputera wspiera wykonywanie badań symulacyjnych? Czy badania takie można przeprowadzić na komputerach z procesorami z architekturą CISC? Na ile wydłużyłby się wtedy czas symulacji? Jaki był czas wykonania pojedynczego badania symulacyjnego na wskazanej w rozprawie stacji roboczej.



5. W rozdziale 6.3.1 wskazane zostało, że zamiast stosowania polityki LRU (ang. Least Recently Used) do obsługi pamięci podręcznej zastosowana została polityka pseudo-LRU. Proszę o określenie na ile wskazana zmiana wpływa na przyspieszenia działania zaproponowanych rozwiązań.

W pracy zauważyłem dodatkowo drobne błędy lub pomyłki językowe i logiczne:

- bardzo liczne błędy literowe w streszczeniu pracy w języku polskim,
- str. 17, zamiana kolejności rysunków 2.2 z 2.3,
- str. 35, wzór 3.1 i 3.2, w obu wzorach inaczej definiowany jest znak przypisania „=”, zmieniając notację matematyczną należy opisać notację, z której się korzysta,
- rysunki 5.3, 5.5, 5.7, 5.8 są nieczytelne,
- str. 38, rozdział 3.5 zaczyna się od samotnej litery „e”,
- str.39 pośrodku znajduje się wyliczenie nie przypisane do żadnego zdania.

5. Konkluzja recenzji

Wyniki rozważań zawarte w rozprawie upoważniają do stwierdzenia, iż została potwierdzona hipoteza pracy. Dodatkowo potwierdzam, że Doktorant umiejętnie korzysta z najnowszej literatury. Praca stanowi samodzielne rozwiązanie przez Autora szeregu zagadnień naukowych przy użyciu nowoczesnych metod badawczych. W związku z tym oceniam rozprawę pozytywnie.

Stwierdzam, że rozprawa pt. *„Hardware acceleration of traffic classifiers for high throughput Ethernet”* autorstwa Pana mgr inż. Macieja Czekaja spełnia wszystkie wymagania stawiane pracom doktorskim przez Ustawę o Tytule Naukowym i Stopniach Naukowych oraz o Stopniach i Tytule w zakresie sztuki z dnia 14 marca 2003 r. (Dz. U. Nr. 65, poz. 595) oraz Rozporządzenia MNiSW z dnia 19 stycznia 2018 roku. W związku z powyższym stawiam wniosek o przyjęcie przedstawionej pracy, jako rozprawy doktorskiej w dziedzinie nauk technicznych, dyscyplinie Informatyka techniczna i telekomunikacja i dopuszczenie jej Autora Pana mgr inż. Macieja Czekaja do publicznej obrony pracy.

