



prof. dr hab. inż. Tadeusz Czachórski
Instytut Informatyki, Wydział AEI Politechniki Śląskiej

Gliwice, 8 wrzesień 2013 r.

Recenzja pracy doktorskiej

Autor rozprawy: mgr inż. Błażej Adamczyk

Tytuł rozprawy: *Wirtualizacja łączny z wykorzystaniem monitora maszyn wirtualnych*

Promotor rozprawy: dr hab. inż. Andrzej Chydziański, profesor nadzwyczajny Politechniki Śląskiej

Kontekst, cel i zakres rozprawy.

Praca dotyczy ważnego i bardzo aktualnego problemu podniesienia efektywności pracy Internetu poprzez wirtualizację jego struktury - tworzenie ponad rzeczywistą fizyczną topologią sieci i jej rzeczywistymi zasobami wielu innych sieci o niezależnej strukturze logicznej. Zasoby sieci są łączone poprzez dzielenie dostępnego pasma na poszczególne kanały, z których każdy jest niezależny od pozostałych, oraz każdy może być przypisany do komputera lub urządzenia w czasie rzeczywistym. Każdy kanał jest niezależnie zabezpieczony. Każdy użytkownik posiada współdzielony dostęp do wszystkich zasobów sieciowych z pojedynczego komputera. Ułatwia to operatorom jednoczesne świadczenie usług pomiędzy grupami użytkowników typu jeden - do - jeden, jeden - do - wielu, itp., oddzielenie od siebie wielu grup użytkowników, co podnosi także bezpieczeństwo pracy, ułatwia zapewnienie gwarancji jakości przesyłu danych, jak również usuwanie błędów i usterek. Praca nawiązuje do rozwiązań zaproponowanych w ramach polskiego projektu Inżynieria Internetu Przyszłości (IIP) i powstałej tam doświadczalnej sieci zwanej Systemem IIP. Węzły wirtualne są tam realizowane m.in. za pomocą monitora maszyn wirtualnych Xen. Są to pełne maszyny wirtualne, dzięki czemu można wykorzystać wewnątrz każdego węzłów dowolne systemy operacyjne. To z

kolei pozwala na zastosowanie już istniejących protokołów i mechanizmów transmisji danych oraz łatwa implementację nowych technologii poprzez modyfikacje lub rozszerzenie wielu dostępnych modułów obecnych systemów operacyjnych, w szczególności systemu Linux. Węzły wirtualne są uruchomione na jednym węźle fizycznym, a więc współdzielą wszystkie zasoby tego węzła. Ważne jest współdzielenie interfejsów sieciowych, czyli właśnie wirtualizacja łączy. W Systemie IIP wirtualizacja łączy jest osiągana poprzez zastosowanie dwóch mechanizmów: klasyfikacji pakietów na wejściu węzła fizycznego, oraz szeregowania pakietów na jego wyjściu.

Teza pracy mówi, że możliwe jest uzyskanie łączy wirtualnych o wysokiej wydajności, wysokiej wzajemnej izolacji wydajnościowej oraz dających gwarancje jakości usług przy pomocy (polonista powiedziałby "za pomocą" !) monitora maszyn wirtualnych. Teza ta jest uwiarygodniona eksperymentalnie.

Struktura pracy Praca liczy 126 stron i zawiera 6 rozdziałów.

Rozdział 1 to wprowadzenie, *rozdział 2* zawiera przegląd podstawowych zagadnień związanych z tematyką wirtualizacji łączy, widzianych z perspektywy Internetu Przyszłości. Omówiono zalety wirtualizacji i powody, dla których jest ona kluczowym elementem wielu architektur sieci przyszłości. Przedstawiono System IIP, t.j. propozycję architektury Internetu Przyszłości zrealizowaną w ramach wspomnianego projektu IIP, w szczególności zastosowany sposób wirtualizacji łączy i algorytm szeregowania pakietów zapewniający pełną izolację wydajnościową. Algorytm ten wykorzystuje N odrębnych buforów, po jednym dla każdego z N łączy wirtualnych i działa w następująco: po opuszczeniu interfejsu wirtualnego ramka zostaje skierowana do bufora odpowiadającego danemu łączy wirtualnemu – w buforach tworzą się kolejki. Wszystkie kolejki są cyklicznie obsługiwane przez łączy fizyczne. Ramki z pierwszej kolejki są transmitowane przez stały okres czasu W_1 , następnie ramki z drugiej kolejki są transmitowane przez stały czas W_2 , itd. Po zakończeniu ostatniego w cyklu okresu transmisji (czas W_N), ponownie obsługiwana jest pierwsza kolejka przez czas W_1 i cykl powtarza się. Transmisja ostatniej ramki w obrębie każdej fazy W_i może się okazać niemożliwa, bo jest za mało czasu na jej całkowitą obsługę (podział ramek nie jest dozwolony). Ramka musi wtedy pozostać w buforze do kolejnej fazy transmisji danego łączy wirtualnego, natomiast łączy fizyczne pozostaje niewykorzystane do końca trwającej fazy, co powoduje pewną utatę przepustowości łączy fizycznego na koniec każdej

fazy transmisji. Ponadto, nawet w przypadku braku ruchu jednego łącza wirtualnego (pusty bufor), algorytm szeregowania oczekuje na zakończenie trwającej fazy nie pozwalając na transmisję innym łączom wirtualnym. Algorytm zapewnia pełną izolację wydajnościową pomiędzy łączami wirtualnymi. Natomiast nie zapewnia ciągłości pracy: czasami łącze wyjściowe nie wysyła ramek, mimo że istnieją ramki oczekujące na transmisję.

Rozdział 3 zawiera dokładną analizę przedstawionego algorytmu szeregowania. Przedstawiony model ma postać wielowymiarowego łańcucha Markowa, i uwzględnia założenia algorytmu: jest N oddzielnych strumieni wejściowych Poissona o intensywnościach $\lambda_1, \dots, \lambda_N$. Każdy strumień wejściowy posiada własny rozkład rozmiaru ramek, D_i . N buforów o rozmiarach b_1, \dots, b_N przechowuje ramki z odpowiednich strumieni wejściowych (N odrębnych kolejek). Jeżeli w momencie przyścia kolejnej ramki odpowiadający jej bufor jest pełny, zostaje ona odrzucona i tracona. Łącze wyjściowe ma stałą przepustowość C bit/s. Kolejki są obsługiwane przez to łącze w sposób cykliczny tak, że każda kolejka otrzymuje stały czas obsługi. Czas obsługi ramki jest proporcjonalny do jej rozmiaru i odwrotnie proporcjonalny do przepustowości łącza. Jeżeli transmisja ramki jest przerwana na końcu czasu W_i , cała ramka zostaje zatrzymana w buforze, a jej transmisja jest ponowiona w następnym cyklu. Wewnątrz każdej kolejki obowiązuje dyscyplina FIFO.

Przy wykorzystaniu modelu określono prawdopodobieństwa przejść między stanami tego łańcucha, uzyskano m.in. współczynnik strat łącza wirtualnego, transformatę Laplace'a łącznego rozkładu długości kolejki i rozmiaru transmitowanej ramki w czasie t , transformatę Laplace'a średniej liczby strat w przedziale $(0, t)$, średnią przepustowość i wprowadzane średnie opóźnienie. Wykonane przykłady numeryczne pokazują, że zachowanie algorytmu, mimo jego prostej budowy, jest złożone i nie można się tutaj opierać na intuicji.

Przedstawiony opis matematyczny tego modelu i sposób jego rozwiązania dowodzą dużej kultury matematycznej Doktoranta i znajomości teorii kolejek.

W *rozdziale 4* są przedstawione dostępne monitory maszyn wirtualnych, opisano metody wirtualizacji systemów komputerowych, i porównano monitory maszyn wirtualnych pod względem izolacji, stabilności, skalowalności i niezawodności. Szczegółowo omówiono stosowany przez Autora monitor maszyn wirtualnych Xen. Omówiono problemy związane z wirtualizacją i sposoby ich rozwiązywania.

Rozdział 5 opisuje implementację łączy wirtualnych w monitorze Xen oraz przeprowadzone testy - opisana jest implementacja klasyfikatora oraz algorytm szeregujący i ich mechanizmy konfiguracji. Ostatnia sekcja zawiera wyniki testów potwierdzające poprawną implementację opisanych mechanizmów.

Rozdział 6 to podsumowanie.

Ocena rozprawy.

W pracy zaprojektowano, przeanalizowano i implementowano mechanizm zapewniający izolację wydajnościową w monitorze maszyn wirtualnych. Sama implementacja nie była łatwa, ponieważ: (i) algorytm szeregowania operuje na bardzo krótkich fazach czasowych (rzędu mikrosekund). W systemie zwirtualizowanym osiągnięcie przerwai o tak wysokiej rozdzielczości wymaga implementacji nowych mechanizmów zegarowych; (ii) implementacja sterowników parawirtualnych Xen nie posiada mechanizmów kontroli interfejsów wirtualnych, dlatego wymagana jest ich modyfikacja w celu zapewnienia pełnej izolacji wydajnościowej; (iii) mechanizm wirtualizacji łączy Xen wprowadza dodatkową kolejkę pomiędzy sterownikami. Wymagana jest więc synchronizacja tych sterowników w celu uniknięcia kolejkowania pakietów, aby nie zakłócać pracy algorytmu szeregowania drugiego poziomu; (iv) wirtualizacja znacząco wpływa na wydajność systemu, konieczna okazała się przemyślana konfiguracja wszystkich mechanizmów sieciowych, aby uzyskać jak najlepszą wydajność interfejsów.

Wyniki testów dowodzą, że implementacja łączy wirtualnych w Xen jest zgodna z założeniami, działa poprawnie z różnymi aplikacjami internetowymi i może być wykorzystana w praktyce. Dlatego tezę rozprawy można uznać za udowodnioną w praktyce.

Do najważniejszych wyników szczegółowych uzyskanych w tej pracy można zaliczyć:

- dokładną analizę matematyczną algorytmu szeregowania,
- dostosowanie monitora maszyn wirtualnych Xen do wirtualizacji łączy z izolacją wydajnościową,
- implementacje mechanizmów wirtualizacji łączy w środowisku Xen, w szczególności implementację mechanizmu szeregowania pakietów zapewniającego pełną izolację wydajnościową, implementację mechanizmu klasyfikacji pakietów zgodnie z ich nagłówkiem, wprowadzonym w ramach IIP, implementację innych mechanizmów wymaganych dla poprawnego działania łączy wirtualnych (nowych

mechanizmów przerwań czasowych, mechanizmu konfiguracji algorytmu szeregowania, mechanizmu automatycznej konfiguracji systemu dla zwiększenia wydajności).

Praca jest napisana starannie i równie starannie przygotowana pod względem edytorskim. Bibliografia zawiera dobrze dobrane 84 pozycje. Wyniki zostały przedstawione w 9 publikacjach, w większości w języku angielskim, w tym w takich czasopiśmie jak *Software: Practice and Experience*, *International Journal On Advances in Networks and Services*, *Theoretical and Applied Informatics*, *Applied Mathematical Modelling*, *Mathematical Problems in Engineering*.

Uwagi szczegółowe.

Kilka drobnych uwag:

- w rozdziale drugim przeprowadzono analizę algorytmu szeregującego zakładając losowe (wykładnicze) odstępy czasu pomiędzy ramkami na wejściu (strumienie Poissona). Jednak pomiary rzeczywistego urządzenia przedstawione w rozdziale 5 były wykonywane dla ruchu wejściowego typu CBR (Constant Bit Rate), czyli dla stałego odstępu czasu pomiędzy ramkami. Jak może wpływać to na ocenę wyników, dlaczego pomiarów nie wykonano również dla losowego odstępu czasu pomiędzy ramkami?
- przedstawiony model algorytmu szeregującego zakłada, że rozmiary buforów wejściowych są wyrażone w ramkach, a nie w bajtach. Przy zmiennym rozmiarze ramki wymaga to zmiennej ilości pamięci na bufor. Jakie błędy może wprowadzać to założenie, bardzo często zresztą robione w modelach kolejkowych – wykorzystanie modeli $M/M/1/N$, $M/G/1/N$, itp. ? Wiem, że wprowadzenie wyrażonej w bajtach pojemności bufora to dodatkowa trudność, do opisu takiego opisu zmiernają prace prof. Tikhonenki, np. O. Tikhonenko, *Queueing Systems with Common Buffer: A Theoretical Treatment*, in: 18th Conference, CN 2011. Ustron, Poland. June 14-18, 2011. Proceedings, Series: Communications in Computer and Information Science, Vol. 160. w: A. Kwicień et al. (Eds.), 1st Edition., 2011, XII, pp. 61-69. Prosiłbym o krótki komentarz w czasie obrony. Jestem też zdania, że bardziej realistyczne założenia co do strumienia wejściowego i rozmiaru pakietów, stałej wyrażonej w bajtach pojemności bufora można by uwzględnić w modelu z wykorzystaniem metody aproksymacji dyfuzyjnej, choć wyniki będą

oczywiście przybliżone w sensie matematycznym. Nie jest to uwaga krytyczna, a raczej wskazanie alternatywnego podejścia,

- dlaczego rozdzielczość i dokładność nowych mechanizmów zegarowych (Tabela 4.5 str 79) została pokazana tylko dla czterech ustawień zegara tzn. $150 \mu s$, $100 \mu s$, $50 \mu s$ i $10 \mu s$? Lepiej byłoby pokazać wyniki dla większej liczby ustawień,
- niektóre przedstawione listingi są zbędne, np. Listing 4.1 str 82, który przedstawia stałe typu "define" można moim zdaniem pominąć,
- we wzorze na obciążenie ρ (ostatni wzór na str. 21) w mianowniku powinno być CW, a nie CV,
- w trzeciej linii wzoru 3.27 (str. 30) brakuje wyrażenia $f(s)$, które powinno się znaleźć przed wyrażeniem z_{b-k} .

Uwagi te nie mają wpływu na moją wysoką ocenę pracy.

Wniosek końcowy.

Podsumowując, uważam, że rozprawa doktorska mgr inż. Błażeja Adamczyka spełnia warunki stawiane rozprawom doktorskim przez ustawę o stopniu i tytułach naukowych. Autor wykazał się dużą znajomością zagadnień związanych z działaniem Internetu, potrafił zaproponować, zastosować i szczegółowo zbadać oryginalną metodę zwiększenia jego efektywności, która być może znajdzie zastosowanie w praktyce. Praca zawiera interesujący opis matematyczny, wymagający dobrego wprowadzenia w teorię kolejek, jak również wyniki doświadczalne, wymagające dużych umiejętności praktycznych i pieczołowitego przygotowania stanowiska badawczego. Wyniki są oryginalne i mają znaczenie poznawcze. Wnioskuje o przyjęcie tej pracy jako rozprawy doktorskiej i dopuszczenie jej do publicznej obrony.



T. Chalicki