

dr hab. inż. Piotr Zwierzykowski, prof. PP

4 maja 2020 r.

Politechnika Poznańska  
Wydział Informatyki i Telekomunikacji,  
Instytut Sieci Teleinformatycznych

## RECENZJA ROZPRAWY DOKTORSKIEJ

przygotowana dla Rady Wydziału Telekomunikacji, Informatyki i Elektrotechniki  
Uniwersytetu Technologiczno-Przyrodniczego w Bydgoszczy

Tytuł rozprawy:

**Effective solutions for high performance communication in the cloud**  
Efektywne rozwiązania dla wysokowydajnej komunikacji w chmurach obliczeniowych

Autor rozprawy:

**mgr inż. Sebastian Łaskawiec**

Promotorzy:

dr hab. inż. Michał Choraś, prof. UTP  
dr inż. Tomasz Andrysiak

Dziedzina:

nauki inżynieryjno-techniczne

Dyscyplina:

informatyka techniczna i telekomunikacja

### 1. Jakie zagadnienie naukowe jest rozpatrzone w pracy (teza rozprawy) i czy zostało ono dostatecznie jasno sformułowane przez autora?

Podstawowym celem, który postawił przed sobą Kandydat jest „wykazanie, że istnieje możliwość zwiększenia szybkości działania systemu uruchomionego w chmurze obliczeniowej. Podjęty przez Doktoranta problem jest problemem aktualnym i ważnym, zarówno pod względem teoretycznym, jak i praktycznym. Mgr inż. Sebastian Łaskawiec przyjął w rozprawie następującą tezę: „Istnieje możliwość poprawienia szybkości komunikacji, zdefiniowanej jako przepustowość lub opóźnienie, oraz obniżenie zapotrzebowania na pamięć przez serwer systemu typu „Data Grid” poprzez wykorzystanie nowych rozwiązań i algorytmów negocjacji protokołów oraz rozwiązań wykorzystujących technikę „Client Side Load Balancing” dla aplikacji uruchomionych w chmurze.”

W celu wykazania tezy rozprawy Kandydat sformułował szereg logicznie uporządkowanych zadań, które kolejno omówił w odpowiednich rozdziałach rozprawy doktorskiej. Zadania dotyczyły opracowania nowych metod izolowania danych pomiędzy użytkownikami systemu i zmniejszenia zapotrzebowania na pamięć serwera systemu typu „Data Grid”, propozycji nowych metod

udostępniania aplikacji uruchomionej w chmurze dla zewnętrznych aplikacji klienckich, zaprojektowania nowego rozwiązania do zmiany protokołów komunikacji dla aplikacji uruchomionych w chmurze oraz propozycji nowych rozwiązań do zarządzania konfiguracją aplikacji i odnajdywaniem błędów konfiguracyjnych.

W mojej opinii teza rozprawy została przez Autora dostatecznie jasno i precyzyjnie sformułowana.

## **2. Jaka jest przydatność rozprawy z punktu widzenia nauk technicznych, czy założenia przyjęte przez autora są uzasadnione?**

Chmury obliczeniowe są coraz częściej wykorzystywane w codziennej działalności wielu firm i organizacji na świecie. Sposób korzystania ze środowisk chmurowych w dużej mierze zależy od obszaru działalności firmy. Każda z nich ma określone wymagania i ograniczenia, które znajdują swój wyraz w sposobie wdrożenia i wykorzystania chmury obliczeniowej. Problemy podejmowane przez Doktoranta bardzo dobrze wpisują w te obszary, ponieważ dotyczą ważnych aspektów implementacyjnych i utrzymaniowych środowisk chmurowych.

Aktualnie w wielu ośrodkach naukowo-badawczych i akademickich prowadzone są intensywne badania, których celem jest zwiększenie szybkości, niezawodności i skalowalności środowisk chmurowych. Badania te obejmują zarówno aspekty implementacyjne, jak i teoretyczne.

Można zatem stwierdzić, że tematyka badań podejmowanych w rozprawie bardzo dobrze wpisuje w międzynarodowy nurt badań teoretycznych i aplikacyjnych prowadzonych w dziedzinie uprawianej przez Autora.

## **3. Czy autor rozwiązał postawione zagadnienie i czy użył właściwej do tego metody?**

Kandydat wykazał w rozprawie przy użyciu jakich metod i technik można poprawić szybkość komunikacji pomiędzy aplikacjami klienckimi, a aplikacjami uruchomionymi w chmurze oraz obniżyć zapotrzebowanie na pamięć przez serwer systemu typu „Data Grid”.

Cel rozprawy został osiągnięty w wyniku przeprowadzenia następujących działań:

- Zaproponowanie nowych metod izolowania danych pomiędzy użytkownikami systemu i zmniejszenie zapotrzebowania na pamięć serwera systemu typu „Data Grid”. Pomocniczym zadaniem było zaprojektowanie nowej metody wykorzystującej właściwości połączeń TCP do zidentyfikowania poszczególnych aplikacji klienckich oraz zaprojektowanie nowej metody dostępu z zewnątrz do usługi uruchomionej w chmurze.

- Zaproponowanie nowych metod udostępniania aplikacji uruchomionej w chmurze zewnętrznym aplikacjom klienckim w oparciu o technikę "Client Side Load Balancing".
- Zaprojektowanie nowego rozwiązania do zmiany protokołów komunikacyjnych dla aplikacji uruchomionych w chmurze. Zadanie to wymagało wykorzystania nowych metod przełączania komunikacji na protokoły binarne wykorzystujące to samo połączenie TCP.
- Zaproponowanie nowych rozwiązań do zarządzania konfiguracją aplikacji i wykrywaniem błędów konfiguracyjnych. Rozwiązanie to umożliwiło osiągnięcie lepszej wydajności, a także zmniejszenie ilości niezbędnej pamięci systemu.

Przyjęte przez Autora metody wpisują się w szeroki nurt metod wykorzystywanych do zwiększenia wydajności środowisk i protokołów chmurowych. Zaproponowane przez Doktoranta metody znalazły również zastosowanie w praktyce m.in. w projekcie Infinispan. Zostały także pozytywnie zweryfikowane przez pracowników firmy Red Hat oraz stały się przedmiotem trzech patentów.

Potwierdzona praktyczność zaproponowanych metod, technik i algorytmów oraz numeryczna ocena ich jakości zawarta w rozprawie, pozwalają na stwierdzenie, że Doktorant rozwiązał postawione zagadnienie i użył do jego rozwiązania właściwej metody.

#### **4. Jaki charakter ma rozprawa (teoretyczny, doświadczalny, konstrukcyjny), jaka jest jej pozycja w stosunku do stanu wiedzy czy poziomu techniki reprezentowanych przez literaturę światową?**

Rozprawa ma charakter doświadczalny z elementami konstrukcyjnymi.

Rozdziały 1 i 2 recenzowanej rozprawy mają charakter opisowy. W rozdziałach tych Kandydat zawarł wprowadzenie do pracy, przedstawił cel, tezę i zakres pracy. W rozdziale 3 Doktorant zawarł przegląd stanu badań. W rozdziale tym Autor opisuje m.in. metody równoważenia obciążenia w chmurze, metody dostępu do aplikacji klastrowych umieszczonych w chmurze i systemy ekspertskie.

Rozdział 4 zawiera propozycję rozwiązania problemu separacji danych różnych klientów kierowanych do serwerów w środowisku chmurowym. Na początku rozdziału Doktorant przyjął założenia, które proponowane rozwiązanie musi, lub powinno, spełniać. Autor założył, że propozycja musi zapewniać izolację danych przechowywanych w kontenerach oraz gwarantować autoryzację. Przyjął, że rozwiązanie powinno również zapewniać wsparcie dla komponentu „Reverse Proxy” i mieć jak najmniejszy wpływ na wydajność systemu (Tab.4.2). Do rozwiązania tego problemu Kandydat zaproponował wykorzystanie pola „hostname” z rozszerzenia SNI (*ang. Service Name Identification*) protokołu TLS (*ang. Transport Layer Security*), które zapewnia rozróżnienie i szyfrowanie

przesyłanych danych. Zaproponowane przez Autora rozwiązanie umożliwiłoby zapewnienie konfiguracji parametrów transportowych dla każdej aplikacji oddzielnie. Oznacza to możliwość dostosowania ilości pamięci przydzielanej do buforowania danych związanych z daną aplikacją do potrzeb tej aplikacji. Takie rozwiązanie pozwala z kolei na zmniejszenie wykorzystania zasobów pamięć serwera typu „Data Grid”. Rozwiązanie to można również wykorzystać do zapewnienia odpowiedniej konfiguracji systemu na podstawie parametrów uwierzytelniania i autoryzacji danej aplikacji. W pracy zaproponowano rozwiązanie implementacyjne realizujące wspomniane funkcje o nazwie „Multi-tenant Router”. Autor przedstawił również pseudokod algorytmu wykorzystywanego w implementacji oraz ocenę jego złożoności. Wyniki eksperymentów przedstawionych w pracy potwierdzają możliwość ograniczenia ilości wymaganej pamięci na użytkownika w systemie typu „Data Grid”, ale odbywa się to kosztem opóźnienia. Z przedstawionych wyników wynika, że rozwiązanie jest efektywne w przypadku dużej liczby użytkowników spełnia więc założenia przyjęte przez Autora rozprawy w ograniczonym zakresie.

W rozdziale 5 przedstawiono rozwiązanie zapewniające dostęp aplikacjom klienckim do konkretnego serwera pracującego w klastrze. Zgodnie z przyjętymi przez Autora założeniami rozwiązanie to musi spełniać wymagania dotyczące zapewnienia łączności (*ang. Connectivity*), znajomości topologii (*ang. Discoverability*) oraz powinno zapewniać możliwość automatyzacji i mieć ograniczony wpływ na obciążenie systemu (Tab.5.1). Kandydat początkowo podjął próbę realizacji tego celu stosując rozszerzenie SNI protokołu TLS, wykorzystujące do tego celu specjalny komponent, który podejmował decyzję o kierowaniu ruchu sieciowego. Na podstawie wyników przedstawionych przez Autora można stwierdzić, że próba ta spowodowała zmniejszenie wydajności systemu (Tab.5.2). Doktorant zrezygnował również z zaproponowania własnej implementacji komponentu Load Balancer, który jest stosowany w rozwiązaniach komercyjnych i może być wykorzystany do realizacji tych funkcji. *Uważam, że ta decyzja wymagałaby szerszego komentarza, ponieważ przesłanki implementacyjne nie są wystraszającym uzasadnieniem.* Ostatecznie Autor wykorzystał do tego celu technikę „client side load balancing”, która zapewnia dostęp do każdej instancji serwera poprzez odrębny adres IP. Początkowo rozwiązanie to korzystało z komponentu „binary proxy”, który umieszczono pomiędzy klientem, a serwerem w celu zagwarantowania dostarczenia wiadomości od klienta do działającej instancji serwera za pośrednictwem dostępnego dedykowanego serwerowi komponenta „Load Balancer”. W tym celu Doktorant zaproponował i zaimplementował nowy komponent systemu o nazwie „External IP Controller”, który zapewniał automatyzację procesu odkrywania nowych instancji serwera i tworzenia komponentów „binary proxy” oraz „Load Balancer” (dla wszystkich instancji serwera). Algorytm prezentujący działanie tego komponentu został zawarty w pracy. Badania przeprowadzone przez Autora wykazały jednak, że komponenty „binary proxy”

W rozdziale 7 opisano rozwiązanie do automatycznego wykrywania błędów konfiguracyjnych. W rozdziale tym zaproponowano system do automatycznego wykrywania niepoprawnej konfiguracji aplikacji uruchomionej w chmurze, w oparciu o inspekcję plików konfiguracyjnych i metryk aplikacji. Na początku rozdziału Autor przyjął, że zaproponowane rozwiązanie musi wykrywać typowe błędy konfiguracyjne oraz być łatwe do wyjaśnienia, a także powinno nie sprawiać większych trudności podczas implementacji (Tab. 7.3). Do rozwiązania problemu Kandydat zaproponował system będący odmianą systemu eksperckiego. Główna pętla sterująca systemu (nazwanego „Inteligentnym operatorem”) składa się z czterech kroków. W pierwszym kroku system zbiera metryki odpytując w tym celu bazę danych, a następnie buduje bazę wiedzy (łąząc konfigurację z metrykami). W kolejnym roku wykorzystuje klasyfikator oparty na drzewie decyzyjnym, a w ostatnim kroku wskazuje działania które należy podjąć. Proponowany system nie modyfikuje żadnych danych użytkownika, a decyzje podejmowane przez model zaimplementowany w rozwiązaniu można przedstawić w postaci grafu, a to znacznie ułatwia wyjaśnienie podjętej decyzji. Proponowany przez Kandydata model realizuje przyjęte przez Autora założenia. Rozwiązanie zostało przetestowane z wykorzystaniem dedykowanego środowiska testowego i założone cele zostały pomyślnie zweryfikowane. Warto podkreślić, że przedstawione przez Doktoranta rozwiązanie ma charakter unikalny, ponieważ istniejące narzędzia umożliwiają automatyczne wykrywanie błędów infrastruktury chmury, ale nie w aplikacjach. Propozycje zawarte w tej części pracy zostały również zawarte w trzech patentach.

Prowadzone przez Kandydata badania mają charakter oryginalny i stanowią bardzo dobre uzupełnienie szerokiego nurtu badań nad systemami chmurowymi prowadzonymi w wielu międzynarodowych ośrodkach badawczo-rozwojowych i akademickich.

## **5. Na czym polega oryginalność rozprawy, co stanowi samodzielny i oryginalny dorobek autora?**

Za najbardziej istotne i oryginalne wyniki Autora uważam:

- Zastosowanie metody wykorzystującej pole „hostname” rozszerzenia SNI protokołu TLS do współdzielenia serwera pomiędzy wielu użytkowników. Pozwoliło to na obniżenie zapotrzebowania na pamięć systemu typu „Data Grid”. *Rozwiązanie to znalazło praktyczne zastosowanie w projekcie Infinispan.*
- Zastosowanie techniki „Client Side Load Balancing” do dostępu do aplikacji uruchomionej w chmurze. *Również ten pomysł został wykorzystany w praktyce w implementacji złożonych funkcjonalności systemu Data Grid, takich, jak replikacja danych pomiędzy klastrami.*

- Zaproponowanie i zastosowanie techniki zmiany protokołów z wykorzystaniem rozszerzenia ALPN protokołu TLS oraz procedury HTTP/1.1 Upgrade oraz wykorzystanie binarnych protokołów komunikacyjnych. Umożliwiło to zwiększenie przepustowości połączenia pomiędzy serwerem i aplikacją kliencką. *Rozwiązanie to znalazło zastosowanie w wersji 10 projektu Infinispan.*
- Zaproponowanie system eksperckiego umożliwiającego wykrywanie błędów w konfiguracji aplikacji uruchomionych w chmurze. *System został pozytywnie zweryfikowany przez osoby utrzymujące projekty OpenSource oraz przez obsługę techniczną firmy Red Hat.*

Prezentowane w rozprawie wyniki badań stanowią samodzielny i oryginalny dorobek Autora.

## 6. Jakie są słabe strony rozprawy i jej główne wady?

Bezpośrednio w odniesieniu do rozważanych w rozprawie rozwiązań oraz prezentowanych wyników badań trudno jest sformułować poważniejsze zastrzeżenia merytoryczne. Doktorant od kilku lat z powodzeniem zajmuje się problematyką środowisk chmurowych o czym świadczą m.in. trzy patenty przytoczone w pracy. Wiele wyników przedstawionych w rozprawie zostało również opublikowanych w publikacjach, w których podlegały analizie i ocenie.

W pracy zabrakło mi jednak pewnego uogólnienia. Moim zdaniem ciekawe byłoby przedyskutowanie łącznego wpływu na szybkość komunikacji poszczególnych propozycji Doktoranta i ich wzajemnej korelacji. Skłoniło mnie to, to sformułowania kilku pytań:

- Czy proponowane rozwiązania można zastosować łącznie? Jeśli tak, to jaki jest sumaryczny zysk?  
A jeśli nie wszystkie elementy można wykorzystywać łącznie, to które można i jakie są warunki brzegowe dla takich rozwiązań.
- Dlaczego środowiska chmurowe w których porównywano wprowadzone modyfikacje były różne? Czy można je wprowadzić i przeprowadzić testy w jednolitym środowisku?
- Jak wpływa wielkość testowanego systemu na uzyskiwane charakterystyki? Czy można się spodziewać podobnych, lub takich samych, tendencji w większych systemach?
- Czy przedstawione rozwiązania warto i czy można zastosować w tradycyjnych środowiskach chmurowych?

- Gdyby pominąć aspekt kosztów (np. rozwiązań typu „Load Balancer” lub innych), to czy zaproponowane rozwiązania nadal można byłoby uznać za dające zysk w kategoriach przepustowości i pamięci?

Praca zawiera pewną liczbę błędów językowych, których można by uniknąć przy bardziej wnikliwej redakcji. Nie znalazłem natomiast w rozprawie błędów merytorycznych, a drobne zastrzeżenia lub wątpliwości, o których pisałem powyżej dotyczą raczej strony redakcyjnej pracy, lub mają charakter polemiczny i nie mogą mieć wpływu na ostateczną pozytywną ocenę pracy. Uważam, że recenzowana rozprawa zawiera wiele oryginalnych wyników i wnosi wartościowy wkład w rozwój informatyki technicznej i telekomunikacji.

## **7. Czy rozprawa świadczy o dostatecznej wiedzy autora i znajomości współczesnej literatury z zakresu dyscypliny naukowej, jakiej rozprawa dotyczy?**

Obszerna literatura przytoczona w pracy (spis publikacji zawiera 126 pozycji) świadczy o rozległej wiedzy i orientacji Autora w dziedzinie, którą uprawia. Zamieszczone pozycje z ostatnich lat (ponad 80% cytowanych prac zostało opublikowane po 2015 roku) potwierdzają, że Kandydat nie zajmuje się tematyką wyczerpaną, lecz przeciwnie, jest ona aktualna i inspirująca badawczo. O kompetencji Kandydata świadczy przywołanie w rozprawie 3 prac, w tym dwóch artykułów w czasopiśmie, których jest współautorem. Prace zostały opublikowane w dwóch międzynarodowych czasopiśmiech „Journal of Cloud Computing Research” oraz „Cluster Computing” (IF=1,851; 70 pkt.). Warto również podkreślić, że Doktorant jest współautorem 3 patentów dotyczących badań zawartych w pracy.

## **8. Czy autor wskazał umiejętność poprawnego i przekonywującego przedstawienia uzyskanych przez siebie wyników (zwięzłość, jasność, poprawność redakcyjna rozprawy)?**

Cel, zakres, podstawy metodologiczne, rezultaty osiągnięte w rezultacie badań i sformułowane wnioski zostały przedstawione w rozprawie najczęściej jasno i precyzyjnie. Kandydat wykazał, że skutecznie opanował metodologię konstrukcji modeli i logikę komponentów stosowanych w środowisku chmurowym. Posiada również dużą wiedzę dotyczącą technicznych i implementacyjnych aspektów środowisk chmurowych. Autor jasno przedstawił swój wkład do dziedziny badań, w której mieści się rozprawa. Układ, strona redakcyjna i terminologiczna rozprawy nie budzą większych zastrzeżeń.

## 9. Czy i jaka jest przydatność rozprawy dla gospodarki narodowej?

Chmury obliczeniowe są coraz częściej wykorzystywane w codziennej aktywności wielu firm i organizacji na całym świecie. Sposób korzystania ze środowisk chmurowych w dużej mierze zależy od dziedziny aktywności danej firmy. Każda z nich ma określone wymagania i ograniczenia, które znajdują swój wyraz w sposobie wdrożenia i wykorzystania chmury obliczeniowej. Wśród nich znajdują się problemy, których propozycje rozwiązania zawarł Kandydat w rozprawie.


Recenzowana praca podejmuje problemy, które bardzo dobrze wpisują się w ten szeroki nurt aktualnie prowadzonych badań. O praktycznym znaczeniu osiągnięć Doktoranta może świadczyć wykorzystanie rozwiązań przedstawionych w pracy w projektach OpenSource takich, jak JGroups, czy Infinispan. Stanowi to potwierdzenie aktualności i praktyczności rozwiązań zaproponowanych w rozprawie dotyczących systemów typu „Data Grid” i chmur obliczeniowych opartych o kontenery. Świadczą o niej również trzy patenty, które bezpośrednio związane są z tematyką pracy.

Proponowane przez Kandydata rozwiązania mogą wpłynąć na ułatwienie procesów uruchomienia i zarządzenia środowiskiem chmurowym opartym na kontenerach oraz zwiększać jego wydajność. Pozwala to na stwierdzenie, że rozprawa może mieć istotny wpływ na poprawę szybkości komunikacji dla aplikacji uruchomionych w chmurze, a tym samym przyczyni się do zwiększenia możliwości wykorzystania środowisk chmurowych, które są istotnym elementem wspomagającym pracę wielu obszarów gospodarki narodowej.

## 10. Czy rozprawa spełnia wymagania stawiane rozprawom doktorskim przez obowiązujące przepisy?

Biorąc pod uwagę wnioski zaprezentowane w poprzednich punktach i wymagania podane w Artykule 13 Ustawy z dnia 14 marca 2003 r. o stopniach naukowych i tytule naukowym oraz stopniach i tytule w zakresie sztuki (z późniejszymi zmianami) uważam, że **rozprawa doktorska** mgr inż. Sebastiana Łaskawca pt. „*Efektywne rozwiązania dla wysokowydajnej komunikacji w chmurach obliczeniowych*” **zawiera oryginalne rozwiązanie problemu naukowego** oraz dowodzi, że **Kandydat posiada ogólną wiedzę teoretyczną** w dyscyplinie informatyka techniczna i telekomunikacja i **posiada umiejętność samodzielnego prowadzenia pracy naukowej.**

Wnoszę o dopuszczenie rozprawy doktorskiej Pana mgr inż. Sebastiana Łaskawca pt. „Efektywne rozwiązania dla wysokowydajnej komunikacji w chmurach obliczeniowych” do publicznej obrony.



Podpis