

НАСТАВНО-НАУЧНОМ ВЕЋУ

Предмет: Подобност теме и кандидата Владимира Славнића за израду докторске дисертације

Одлуком бр. 5033/09-1 од 11. јула 2016. године именовани смо за чланове Комисије за оцену подобности теме и кандидата Владимира Славнића за израду докторске дисертације и научне заснованости теме **„Примена дистрибуираних рачунарских инфраструктура за рачунарство високих перформанси”** (енг. **“Application of distributed computing infrastructures to high performance computing”¹**).

На основу материјала приложеног уз Захтев кандидата, Комисија подноси следећи

ИЗВЕШТАЈ

1. Подаци о кандидату

1.1. Биографски подаци

Владимир Славнић је рођен 11. августа 1981. године у Смедеревској Паланци. Основну школу ”Свети Сава”, као и Гимназију, завршио је у Великој Плани. Основне студије је похађао на Електротехничком факултету Универзитета у Београду на смеру Рачунарска техника и информатика у периоду од 2000. до 2007. године, када је дипломирао са радом под називом ”Имплементација наставничких сервиса на мобилним телефонима”. На основу овога стекао је титулу мастер инжењер електротехнике и рачунарства. Докторске студије је уписао 2009. године на модулу Софтверско инжењерство.

1.2. Стечено научноистраживачко искуство

Кандидат Владимир Славнић је уписао докторке студије школске 2009/2010. године. По истеку законског рока за завршетак докторских академских студија, на захтев студента, одобрено је продужење рока за завршетак ових студија за још два семестра, сагласно члану 92. став 4 Статута Универзитета у Београду.

У оквиру докторских студија, до подношења захтева за одобрење докторске дисертације, кандидат Владимир Славнић је положио све испите и успешно завршио све обавезе предвиђене планом и програмом докторских студија. Испити које је кандидат положио су:

¹ Дисертација ће бити писана на енглеском језику

1. Виртуелна стварност (оцена 10, ЕСПБ: 9)
2. Медицинска информатика (оцена 10, ЕСПБ: 9)
3. Скенирајућа микроскопија у нанонауци и нанотехнологији (оцена 10, ЕСПБ: 9)
4. Организација система дискова (оцена 10, ЕСПБ: 9)
5. Сигурност и заштита рачунарских система (оцена 9, ЕСПБ: 9)
6. Одабрана поглавља из рачунарске графике (оцена 10, ЕСПБ: 9)
7. Интернет програмирање (оцена 10, ЕСПБ: 6)
8. Функционална анализа и примене (оцена 10, ЕСПБ: 9)
9. Увод у научни рад (оцена 10, ЕСПБ: 6)
10. Специјалне функције (оцена 10, ЕСПБ: 9)

и испунио следеће обавезе:

1. Научно стручни рад (ЕСПБ: 3),
2. Студијски истраживачки рад I (ЕСПБ: 15),
3. Студијски истраживачки рад II (ЕСПБ: 15),

остваривши укупно 120 ЕСПБ и просечну оцену 9,90. Кандидат је такође успешно положио испит о подобности теме и кандидата (докторски испит) пред комисијом у саставу:

1. др Захарије Радивојевић, доцент, Универзитет у Београду – Електротехнички факултет
2. др Антун Балаж, научни саветник, Универзитет у Београду – Институт за физику у Београду
3. др Јелена Радовановић, редовни професор, Универзитет у Београду – Електротехнички факултет

Јавно усмено излагање је одржано 15. августа 2016. године при чему је добио оцену задовољно.

Владимир Славнић је у периоду од 2008. до 2010. године био ангажован у Лабораторији за примену рачунара у науци Института за физику у Београду као сарадник на FP6 пројектима SEE-GRID-SCI (SEE-GRID eInfrastructure for regional eScience) и EGEE-III (Enabling Grids for E-science). Запослен је у Институту за физику у Београду као истраживач приправник у Лабораторији за примену рачунара у науци од почетка 2011. године. Звање истраживач сарадник стекао је у септембру 2011. године. Поред пројекта основних истраживања OH171017 којим руководи др Антун Балаж, Владимир Славнић је био ангажован на низу FP7 пројеката у оквиру међународне сарадње: EGI-InSPIRE (European Grid Infrastructure), HP-SEE (High-Performance Computing Infrastructure for South East Europe's Research Communities), agINFRA (Data infrastructure to support agricultural scientific communities), SemaGrow (Data Intensive Techniques to Boost the Real-Time Performance of Global Agricultural Data Infrastructures), као и на серији PRACE IP пројеката (Partnership for Advanced Computing in Europe): PRACE-1IP, PRACE-2IP и PRACE-3IP. Тренутно је ангажован на Horizon 2020 пројекту VI-SEEM (Virtual Research Environment (VRE) for interdisciplinary communities in Southeast Europe and the Eastern Mediterranean). Објавио је 1 рад у међународном часопису изузетних вредности категорије M21a, 1 рад у врхунском међународном часопису категорије M21, 2 рада у тематским зборницима водећег међународног значаја категорије M13, 1 предавање по позиву на међународном скупу штампано у изводу категорије M32, 5 саопштења са међународних конференција категорије M33, 2 саопштења са међународних скупова штампана у изводу категорије M34 и 2 саопштења са скупова националног значаја категорије M63.

Радови у међународним часописима изузетних вредности (M21a):

1. Hybrid OpenMP/MPI Programs for Solving the Time-dependent Gross-Pitaevskii Equation in a Fully Anisotropic Trap
B. Satarić, **V. Slavnić**, A. Belić, A. Balaž, P. Muruganandam, and S. K. Adhikari
Comput. Phys. Commun. **200**, 411 (2016)
doi: 10.1016/j.cpc.2015.12.006
ISSN: 0010-4655
IF 2014: 3.112

Радови у врхунским међународним часописима (M21):

1. Development of Grid E-Infrastructure in South-Eastern Europe
A. Balaž, O. Prnjat, D. Vudragović, **V. Slavnić**, I. Liabotis, E. Atanassov, B. Jakimovski, and M. Savić
J. Grid Comput. **9**, 135 (2011)
doi: 10.1007/s10723-011-9185-0
ISSN: 1570-7873
IF 2011: 1.310

Радови у тематским зборницима водећег међународног значаја (M13)

1. Implementation and Benchmarking of New FFT Libraries in Quantum ESPRESSO
D. Stanković, P. Jovanović, A. Jović, **V. Slavnić**, D. Vudragović, and A. Balaž
High-Performance Computing Infrastructure for South East Europe's Research Communities, Results of the HP-SEE User Forum 2012
Springer Book Series on Modeling and Optimization in Science and Technologies, **2**, 155 (2014)
doi: 10.1007/978-3-319-01520-0_19
ISBN: 978-3-319-01519-4
2. An Analysis of FFTW and FFTE Performance
M. Nikolić, A. Jović, J. Jakić, **V. Slavnić**, and A. Balaž
High-Performance Computing Infrastructure for South East Europe's Research Communities, Results of the HP-SEE User Forum 2012
Springer Book Series on Modeling and Optimization in Science and Technologies, **2**, 163 (2014)
doi: 10.1007/978-3-319-01520-0_20
ISBN: 978-3-319-01519-4

Саопштења са међународних скупова штампана у целини (M33)

1. SPEEDUP - optimization and porting of path integral MC Code to new computing architectures
V. Slavnić, A. Balaž, D. Stojiljković, A. Belić, and A. Bogojević
Proceedings of the XXII International Symposium on Nuclear Electronics & Computing, Dubna, Russia, 257 (2010)
2. Serbian Participation in Grid Computing Projects
D. Vudragović, A. Balaž, **V. Slavnić**, and A. Belić
Proceedings of the XXII International Symposium on Nuclear Electronics & Computing, Dubna, Russia, 286 (2010)

3. Operational Grid Tools Developed at SCL
V. Slavnić, B. Acković, D. Vudragović, A. Balaž, and A. Belić
 Proceedings of the SEE-GRID-SCI User Forum 2009, Istanbul, Turkey, 9–10 December 2009, 117 (2009)
 ISBN: 978-975-403-510-0
4. Grid Site Monitoring Tools Developed and Used at SCL
V. Slavnić, B. Acković, D. Vudragović, A. Balaž, A. Belić, and M. Savić
 Proceedings of the SEE-GRID-SCI User Forum 2009, Istanbul, Turkey, 9–10 December 2009, 123 (2009)
 ISBN: 978-975-403-510-0
5. Optimization and Porting of the Path Integral Monte Carlo SPEEDUP Code to New Computing Architectures
V. Slavnić, A. Balaž, D. Stojiljković, A. Belić, and A. Bogojević
 Proceedings of the SEE-GRID-SCI User Forum 2009, Istanbul, Turkey, 9–10 December 2009, 133 (2009)
 ISBN: 978-975-403-510-0

Предавање по позиву са међународног скупа штампано у изводу (M32)

1. Overview of Grid and High Performance Computing activities in Serbia
V. Slavnić
 Book of Abstracts, RO-LCG 2014 Conference, "Grid, Cloud and High-Performance Computing in Physics Research", 3-5 November 2014, Bucharest-Magurele, Romania, 15 (2014)
 ISBN: 978-973-0-17842-5

Саопштења са међународних скупова штампана у изводу (M34)

1. AgINFRA - a Data Infrastructure to Support Agricultural Scientific Communities: Promoting Data Sharing and Development of Trust in Agricultural Sciences
 D. Vudragović, **V. Slavnić**, I. Spasojević, V. Nedović, and A. Balaž
 Book of Abstracts, 12th Congress of Nutrition, 31 October 2012, Belgrade, Serbia, 422 (2012)
2. Range and Sensitivities of 2-[(Carboxymethyl) sulfanyl]-4-oxo-4-arylbutanoic Acids Property Spaces. Part 2. Multidimensional Free Energy Landscapes
 B. J. Drakulić, A. Pedretti, M. Zloh, **V. Slavnić**, I. O. Juranić and M. M. Dabović
 Abstract Book of 18th European Symposium on Quantitative Structure-Activity Relationships, Rhodes, Greece, 278 (2010)

Саопштења са скупова националног значаја штампана у целини (M63)

1. WMSMON - GLite WMS Monitoring Tool
 D. Vudragović, **V. Slavnić**, A. Balaž, and A. Belić
 Proceedings of the MIPRO 2009 Conference, Opatija, Croatia, 25–29 May 2009, GSV-02, 239-243 (2009)

2. DWARF - The Framework for Authorized YUM/APT Repositories Management
D. Vudragović, A. Balaž, V. Slavnić, and A. Belić
Proceedings, INFOTEH 2009 Conference, Jahorina, Bosnia and Herzegovina, 18–20 March
2009, E-V-8, 721-725 (2009)

1.3. Оцена подобности кандидата за рад на предложеној теми

На основу научноистраживачке и стручне активности и постигнутих резултата може се закључити да је Владимир Славнић подобан кандидат за израду докторске дисертације на тему: „Примена дистрибуираних рачунарских инфраструктура за рачунарство високих перформанси” (енг. “Application of distributed computing infrastructures to high performance computing”). Број објављених радова и положен докторски испит, као и наведене истраживачке активности у области предложене теме докторске дисертације, показују спремност за самосталан рад и посвећеност датој теми што представља несумњив аргумент за одобрење за рад на поменутој теми.

2. Предмет и циљ истраживања

Предмет истраживања кандидата су дистрибуиране Грид рачунарске инфраструктуре и електронске инфраструктуре за рачунарство високих перформанси (High-performance computing, HPC), као и њихова примена у решавању савремених научних проблема кроз паралелне нумеричке симулације. Модерна наука се у великој мери ослања на рачунарске ресурсе, посебно на Грид и HPC инфраструктуре, како би се дошло до научних пробоја и значајних открића. Кандидат ће испитивати особине оваквих инфраструктура као и програмске алате и парадигме на којима се заснивају, радиће на развијању оперативних алата за управљање оваквим рачунарским ресурсима као и на алатима за њихово надгледање (мониторинг). Кандидат ће такође радити на паралелизовању и оптимизовању нумеричких симулација које ће се извршавати на Грид и HPC инфраструктурама и истражити напредне технике за визуелизацију и обраду података који се добијају извршавањем симулација. Кроз ово ће такође бити истраживана ефикасност електронских инфраструктура, али са корисничке тачке гледишта.

Циљеви истраживања у оквиру ове дисертације ће бити следећи:

- ☐ Унапређење Грид рачунарских инфраструктура ради њиховог ефикаснијег коришћења од стране корисника, посебно креирање окружења које ће омогућити надгледање инстанци система за дистрибуирано управљање извршавањем нумеричких симулација (Workload Management System, WMS), једног од централних Грид сервиса, кључног за функционисање ових инфраструктура. Такође, циљ кандидата биће и развијање низа операционих алата који ће омогућити ефикасну инсталацију, конфигурацију и одржавање Грид рачунарских кластера, као и Линукс кластера опште намене.
- ☐ Испитивање могућности интеграције дистрибуираних рачунарских ресурса у сервисе који би били на располагању научним заједницама. Истражиће се могућности развоја сервиса који ће искористити предности оваквих инфраструктура, тј. могућност истовременог извршавања великог броја корисничких апликација као и приступа значајним складишним ресурсима, а у оквиру тога кандидат ће радити и на прилагођавању (porting) различитих програма на Грид инфраструктуру и развити одговарајућа окружења за контролу њиховог извршавања од стране корисника. Овако развијени сервиси ће бити примењени у међународним пројектима у којима учествује Лабораторија за примену рачунара у науци Института за физику у Београду, где је кандидат запослен.

- Имплементација и оптимизација различитих нумеричких алгоритама за решавање сложених нумеричких проблема на HPC рачунарским инфраструктурама. У оквиру развоја научних апликација Лабораторије за примену рачунара у науци Института за физику у Београду кандидат ће радити на паралелизацији алгорита за решавање нелинеарних парцијалних диференцијалних једначина (Грос–Питаевски једначина, односно нелинеарна Шредингерова једначина, у оквиру GP-SCL апликације), као и на имплементацији и паралелизацији Монте Карло алгорита за решавање функционалних интеграла за израчунавање квантних амплитуда прелаза (апликација SPEEDUP). SPEEDUP апликација ће такође бити прилагођена различитим рачунарским архитектурама (Intel x86_64, IBM POWER, IBM PowerXCell). Како се брза Фуријеова трансформација (FFT) користи прилично често у научним апликацијама, кандидат ће такође истраживати перформансе различитих имплементација FFT нумеричких библиотека и могућност њихове интеграције у постојеће научне апликације. Ово ће конкретно бити урађено у значајном и широко коришћеном програмском пакету за рачунање особина електронских структура и моделовање материјала, Quantum Espresso (QE), чије извршавање значајно зависи од рачунања Фуријеове трансформације.
- Проучавање и примена напредних метода визуелизације и анализе података добијених у нумеричким симулацијама. Кандидат ће истражити савремене визуелизационе алате и *in situ* метод приказа и обраде података на рачунарским кластерима. То истраживање ће затим бити примењено на конкретним нумеричким симулацијама у оквиру пројеката из FP7 и Horizon 2020 програма (Partnership for Advanced Computing in Europe, PRACE).

Грид инфраструктуре пружају значајне истраживачке ресурсе хиљадама научника широм света, али су са друге стране прилично комплексне и још увек недовољно ефикасне. Оптимална искоришћеност кластера који су део Грид инфраструктуре у великој мери зависи од ефикасности Workload Management System (WMS) инстанци, једног од централних сервиса који дистрибуира корисничке апликације на извршавање на конкретне рачунарске кластере, и кандидат ће у оквиру свог истраживачког рада развити окружење које ће за циљ имати мониторинг рада WMS инстанци и које ће на време указати администраторима кластера на неправилан рад неке од многобројних компоненти сервиса. Кандидат ће такође проучавати управљање Грид рачунарским ресурсима и радити на развоју алата за његово унапређивање, као што су методи за паралелно извршавање команди и пренос података на велики број сервера истовремено, креирање и управљање репозиторијумима пакета за Линукс оперативни систем (нпр. RPM пакета), генерисање скупова Грид корисника на кластерима, итд. Таква Грид инфраструктура ће се користити у оквиру SemaGrow FP7 пројекта где ће бити потребно искористити Грид инфраструктуру за интеграцију хетерогених репозиторијума пољопривредних података (XML, NetCDF и др.) што ће омогућити доступност тих података кроз јединствену SPARQL приступну тачку. Истраживачким радом у оквиру своје дисертације, а такође и учествовањем у европским Грид пројектима (SEE-GRID-SCI, EGI-inSPIRE) кандидат ће такође радити и на координацији рада рачунарских кластера у Србији и на унапређењу националне Грид инфраструктуре у оквиру AEGIS (Academic and Educational Grid Initiative of Serbia).

Појам рачунарства високих перформанси је одувек био значајан због потребе да се кориснички програми извршавају што је могуће брже, а такође хардверска захтевност и комплексност програма непрекидно расте. HPC заједно чине брзи рачунарски системи, софтверски алати, програмски језици и паралелне програмске парадигме. HPC је област која се бурно развија и чије истраживање и примена добијених решења значајно доприноси напретку науке и индустрије. Сложеност хардвера и његова хетерогеност (постојање вишејезгарних процесора са више нивоа кеша, у последње време и већег броја различитих

врста акцелератора као што су GPU, Intel Xeon Phi, Maxeler DFE, више типова мрежне инфраструктуре између сервера, као што су Ethernet, InfiniBand, итд.) отежавају убрзавање рада софтвера и потребно је одабрати одговарајуће технике паралелизације и оптимизације за сваку појединачну комбинацију софтвера и архитектуре на којој ће се он извршавати да би се добило жељено убрзање рада апликације и прихватљиво скалирање на циљаном броју процесора/сервера на којима ће се апликација извршавати. Циљ оптимизације апликација је потпуно искоришћавање доступних рачунарских ресурса, елиминисање „уских грла“ која се могу појавити у различитим деловима кода, а све то имајући у виду хардверску имплементацију и архитектуру процесора посматране платформе. Кандидат ће применом различитих техника паралелизовати и оптимизовати горе поменуте апликације развијане у Лабораторији за примену рачунара у науци Института за физику у Београду (SPEEDUP, GP-SCL), упоредити перформансе различитих имплементација FFT библиотека и интегрисати их у постојеће научне апликације (QE), и на тај начин омогућити ефикасније коришћење расположивих инфраструктура и добијање нових научних резултата.

У последње време је присутан изузетан раст, како квалитета, тако и квантитета података које научне заједнице широм света генеришу у нумеричким симулацијама, прикупљају у експериментима, добијају са различитих сензора, друштвених мрежа, из медицинских установа, итд. Упис резултата нумеричких симулација у фајлове на дисковима и њихово читање је већ дуго познато као типично „уско грло“ у рачунарству високих перформанси, а ово је посебно тачно за визуелизационе алате јер савремене симулације на садашњим суперрачунарима као излаз производе податке чија се величина мери и у петабајтима и где се подаци свакако не могу приказати без неке посебне технике филтрирања која смањује резолуцију или количину података које би требало да обрађује алат за визуелизацију. Један од савремених приступа је *in situ* метод визуелизације који омогућује спрегу рачунања и визуелизације. *In situ* визуелизационе библиотеке омогућавају кориснику да се директно повеже на симулацију која се извршава, да прегледа податке, извршава нумеричке упите, филтрира податке и креира графички излаз користећи исте ресурсе на којима се извршава и симулација, а све то док симулација и даље траје. Такође, постоји могућност и аутоматизоване анализе и визуелизације података док се симулација извршава. Овај метод је повезан са растућим потребама за симулације које се скалирају на великом броју процесора, док у исто време елиминише потребу за уписом велике количине података на складишни систем кластера. Штавише, он дозвољава и управљање симулацијом, тј. симулација може бити привремено заустављена док се анализирају тренутно добијени подаци или се покретати корак по корак. Из ових разлога *in situ* визуелизација постаје све више интересантна научницима широм света. Имплементацијом најмодернијих (state-of-art) визуелизационих техника које нуде савремени визуелизациони алати омогућила би се ефикаснија анализа и обрада великих количина података које нумеричке симулације генеришу у току извршавања и то је једна од тема коју ће кандидат истражити у свом раду и применити на конкретним примерима.

Литература:

1. I. Foster and C. Kesselman, *The Grid: Blueprint for a New Computing Infrastructure* (Morgan Kauffman, San Francisco, 1999).
2. I. Foster, *Designing and building parallel programs* (Addison-Wesley Reading, 1995).
3. Tiziana Ferrari and Luciano Gaido, *Resources and Services of the EGEE Production Infrastructure*, *J. Grid Comput.* **9**, 119 (2011).
4. A. Balaž, O. Prnjat, D. Vudragović, V. Slavnić, I. Liabotis, E. Atanassov, B. Jakimovski, and M. Savić, *Development of Grid E-Infrastructure in South-Eastern Europe*, *J. Grid Comput.* **9**, 135 (2011).

5. F. Dalfovo, S. Giorgini, L.P. Pitaevskii, and S. Stringari, Theory of Bose–Einstein condensation in trapped gases, *Rev. Mod. Phys.* 71, 463 (1999).
6. X. Antoine, W. Bao, and C. Besse, Computational methods for the dynamics of the nonlinear Schrödinger/Gross–Pitaevskii equations, *Comput. Phys. Commun.* 184, 2621 (2013).
7. A. Balaž, I. Vidanović, D. Stojiljković, D. Vudragović, A. Belić, and A. Bogojević, SPEEDUP Code for Calculation of Transition Amplitudes Via the Effective Action Approach, *Commun. Comput. Phys.* 11, 739 (2012).
8. L. E. Young-S., D. Vudragović, P. Muruganandam, S. K. Adhikari, and A. Balaž, OpenMP Fortran and C programs for solving the time-dependent Gross-Pitaevskii equation in an anisotropic trap, *Comput. Phys. Commun.* 204, 209 (2016).
9. V. Lončar, A. Balaž, A. Bogojević, S. Škrbić, P. Muruganandam, and S. K. Adhikari, CUDA Programs for Solving the Time-dependent Dipolar Gross-Pitaevskii Equation in an Anisotropic Trap, *Comput. Phys. Commun.* 200, 406 (2016).
10. B. Satarić, V. Slavnić, A. Belić, A. Balaž, P. Muruganandam, and S. K. Adhikari, Hybrid OpenMP/MPI Programs for Solving the Time-dependent Gross-Pitaevskii Equation in a Fully Anisotropic Trap, *Comput. Phys. Commun.* 200, 411 (2016).
11. D. Vudragović, I. Vidanović, A. Balaž, P. Muruganandam, and S. K. Adhikari, C Programs for Solving the Time-dependent Gross-Pitaevskii Equation in a Fully Anisotropic Trap, *Comput. Phys. Commun.* 183, 202 (2012).

3. Полазне хипотезе

Хипотезе које ће бити испитиване и провераване у овом истраживању су следеће:

1. Грид инфраструктуре су неефикасне због своје комплексности. Исправан рад централних сервиса као што је WMS који дистрибуира апликације на извршавање на кластере који су део инфраструктуре је круцијалан за њено ефикасно функционисање. Мониторинг инстанци овог сервиса може помоћи да се утврди где су „уска грла“, да би се то онда искористило за оптимизовање рада целе Грид инфраструктуре.
2. Постоји низ операција које се често извршавају на Линукс кластерима и које се могу поједноставити и тако олакшати и аутоматизовати рад администратора: једноставније паралелно извршавање команди и трансфера фајлова на свим серверима у кластеру (којих може бити и више стотина), креирање и управљање репозиторијумима пакета, генерисање скупова (pools) Грид корисника, итд. Развојем алата који извршавају те операције може се повећати ефикасност кластера, а тиме и ефикасност инфраструктуре којима они припадају.
3. Грид инфраструктуре научним заједницама пружају могућност коришћења значајних рачунарских и складишних ресурса. Да би се ти ресурси ефикасно искористили потребно је развити сервисе доступне научним заједницама који би у позадини искористили Грид за потребна рачунања (могућност истовременог слања хиљада апликација на извршавање) и складиштења података (расположиви терабајти простора на Грид складишним ресурсима).
4. Време извршавања нумеричких симулација је понекад веома дуго, што често може да буде критичан проблем у науци. То време се може смањити прилагођавањем симулације архитектуре на којој се извршава, оптимизацијом њене структуре, другачијим избором структура података итд. У случају да је чак и након оптимизације време извршавања серијске верзије симулације сувише велико, решење је у њеној паралелизацији. Одговарајућом паралелизацијом и оптимизацијом (што укључује и правилан избор одговарајућих нумеричких библиотека и компајлера, као и опција за компајлирање) може се вишеструко убрзати извршавање нумеричких симулација и повећати њихова скалабилност, а тиме и ефикасност рада целе инфраструктуре.

- Истраживачи немају могућност анализе и приказа података које генерише нумеричка симулација док је она у току. Такође, постоји велика хардверска (I/O) захтевност (уписи на и читање са диска) обраде и визуелизације велике количине података која се обично извршава на једном рачунару, док се симулација извршава на великом броју сервера у кластеру. Имплементацијом *in situ* визуелизације омогућава се коришћење истих ресурса и за рачунање и за обраду и приказ података, као и интеракција корисника са симулацијом, док се сама симулација и даље извршава.

4. Научне методе истраживања

У току израде дисертације резултати ће бити добијени различитим методама и биће коришћено више различитих технологија у зависности од посматраног проблема са циљем одабира најефикаснијег приступа.

У случају унапређивања Грид инфраструктуре користиће се IEEE POSIX стандарди да би развијени алати за управљање Грид кластерима и корисничким функцијама на њима били компатибилни са свим оперативним системима из шире породице UNIX. У случају WMS мониторинга ће се, користећи спецификације Грид софтвера (gLite), утврдити компоненте овог сервиса које је потребно надгледати, а затим одредити оптимална архитектура окружења које ће се развити (колектор-агент). Уз BASH POSIX скриптинг, као додатне технологије потребне за развој, биће коришћени и Apache веб сервер, PHP као и RRD (Round Robin Database) за креирање дела окружења за интеракцију са корисницима (front-end). Апликација за креирање и управљање репозиторијумима пакета је замишљена као веб апликација модуларне структуре. Биће реализована уз помоћ Apache веб сервера, PHP и BASH скрипти, а биће коришћен HTTP са SSL технологијом да би се омогућило коришћење дигиталних сертификата као методе аутентификације.

У случају имплементације Грид инфраструктуре у оквиру сервиса за интеграцију хетерогених репозиторијума података, кандидат ће морати да се упозна са елементима семантичког веба, а уз Грид, биће потребно користити и низ других постојећих технологија: RESTful веб сервисе, CouchDB базу, као и триплстор (triplestore) базе. Кандидат ће развити неопходни RESTful интерфејс за интеракцију са сервисом, а постојеће корисничке апликације за конверзију података у RDF формат и складиштење тако креираних података у триплстор базу ће бити прилагођене извршавању на Грид инфраструктури.

У случају оптимизације и паралелизације HPC апликација кандидат ће морати да евалуира постојеће перформансе имплементираних алгоритама и нумеричких библиотека, да развије унапређене верзије апликација и да затим утврди добитак у перформансама (кроз времена извршавања и добијено скалирање апликација).

Имплементација алгорита за рачунање функционалних интеграла уз помоћ Монте Карло метода у виду SPEEDUP апликације писане у програмском језику C ће бити оптимизована и прилагођена извршавању на више различитих платформи, користећи различите начине паралелизације (MPI, Posix Threads, и др.). Како се SPEEDUP код ослања на Монте Карло алгоритам за решавање функционалних интеграла, потребно је користити ефикасан паралелан генератор псеудослучајних бројева и то ће бити SPRNG библиотека (Scalable Parallel Random Number Generator), која може да генерише некорелисане случајне бројеве користећи униформну расподелу густине вероватноће у форми погодној за паралелне симулације. Уз стандардне GNU C компајлере доступне на свим архитектурама, биће коришћени и посебни компајлери, специфични за одређене архитектуре: Intel ICC компајлер намењен Intel-овим процесорима, IBM-ов XL C/C++ компајлер за POWER платформу и IBM

XL C/C++ Compiler for Multicore Acceleration специјализован за Cell архитектуру, а на свакој платформи биће потребно и одредити оптималне опције за компајлирање апликације.

У случају GP-SCL апликација писаних у програмском језику C за решавање нелинеарних парцијалних диференцијалних једначина, почетне OpenMP верзије кода (за пропагацију у реалном и имагинарном времену) ће бити хибридизоване, тј. кандидат ће развити паралелне верзије које ће моћи да се извршавају на више сервера. Користиће хибридни приступ: MPI (Message Passing Interface) програмска парадигма ће бити комбинована са постојећим OpenMP кодом да би се добио флексибилни паралелизам у оквиру дистрибуираног/дељеног меморијског модела и код који ће моћи да се извршава на различитим рачунарским системима, са добром скалабилношћу на великом броју процесора. Ефикасност и скалабилност развијених верзија кода ће бити тестиране на расположивим рачунарским ресурсима.

Код истраживања утицаја FFT имплементација на перформансе апликација биће развијени тест програми који ће користити посматране библиотеке (FFTW и FFTE) и перформансе и скалабилност FFT у три димензије ће бити тестиране на расположивим суперрачунарима (JUGENE у Немачкој, CURIE у Француској). Затим ће у апликацију Quantum Espresso бити имплементирана подршка за FFTE, као и за верзију FFTW библиотеке која подржава рад са нитима (threads). Перформансе тако добијеног програма ће бити анализирани и упоређени са перформансама програма када се користе већ подржане библиотеке за FFT (на пример, интерна FFTW библиотека која се дистрибуира са самом QE апликацијом).

У случају in situ визуелизације, кандидат ће анализирати постојеће најсавременије (state-of-the-art) апликације за визуелизацију (VisIt и ParaView), њихове библиотеке за in situ визуелизацију, као и њихове могућности. Након тога ће се одабрати in situ приступ имплементирати у нумеричку симулацију BrainCore, једну од апликација PRACE Preparatory Access пројекта „Visualization of output from large-scale brain simulation“ истраживача шведског Краљевског института за технологију (KTH). BrainCore је експериментални код написан у C++ програмском језику који симулира апстрактни Поасонов пулсирајући модел неокортикалне мреже мозга и користи Message Passing Interface (MPI) библиотеку за паралелизацију. In situ методом ће се визуелизовати активност неурона на поједностављеној верзији BrainCore апликације користећи рачунарске кластере који су на располагању кандидату.

5. Очекивани научни допринос

Очекивани научни допринос се може поделити у следеће ставке:

- ☐ Унапређење мониторинга и оптимизација рада дистрибуираних рачунарских инфраструктура, као и унапређење управљања оваквим системима.
- ☐ Имплементација окружења за интеграцију хетерогених репозиторијума података на Грид инфраструктури у оквиру SemaGrow пројекта ће олакшати приступ различитим пољопривредним подацима и метаподацима заједници која се бави пољопривредним наукама.
- ☐ Монте Карло апликација SPEEDUP за решавање функционалних интеграла која се користи за проучавање квантомеханичких система и истраживање глобалних и локалних особина Бозе-Ајнштајн кондензације биће оптимизирана, паралелизована и прилагођена извршавању на новим рачунарским архитектурама, чиме ће бити омогућено њено коришћење на ширем скупу рачунарских кластера и суперрачунара.
- ☐ Развој хибридних MPI/OpenMP GP-SCL кодова за рачунање стационарних и нестационарних решења временски зависне Грос-Питаевски једначине у три

просторне димензије ће омогућити извршавање тих кодова на већем броју модерних рачунарских архитектура и проучавање Бозе-Ајнштајн кондензата на ултраниским температурама са много више детаља (са много већом резолуцијом) него што је то до сада био случај.

- Интеграција ефикасних FFT библиотека у Quantum Espresso програмски пакет за рачунање електронских структура и моделовање материјала ће омогућити ефикасније извршавање овог пакета на већем броју HPC кластера од чега ће корист имати научна заједница која се бави науком о материјалима.
- Развијена методологија тестирања поменутих апликација и нумеричких библиотека представљаће основу за одлучивање о стратегији коришћења рачунарских ресурса за њихово извршавање. Такође, ова методологија може да представља основу за тестирање ових или других апликација на другим рачунарским архитектурама.
- Елиминисање I/O-а као „уског грла“ у процесу обраде и визуелизације података које генеришу нумеричке симулације. Имплементација *in situ* начина визуелизације ће омогућити спрегу рачунања и приказа и обраде података кроз коришћење истих рачунарских ресурса, а истраживачима дати контролу над извршавањем симулације и визуелизацију података док се симулација извршава. На тај начин ће се брже и ефикасније долазити до научних резултата. Имплементација *in situ* приступа у случају BrainCorg апликације ће омогућити јаснији увид у активност неокортикалне мреже мозга.

6. План истраживања и структура рада

Кандидат ће испитивати и упоредити особине две врсте модерних рачунарских инфраструктура, географски дистрибуиране Грид рачунарске инфраструктуре и електронске инфраструктуре за рачунарство високих перформанси и радиће на развијању оперативних алата за управљање оваквим рачунарским ресурсима, као и на алатима за њихово надгледање. Кандидат ће такође радити на паралелизовању и оптимизовању нумеричких симулација које ће се извршавати на Грид и HPC инфраструктурама и истражити напредне технике за визуелизацију и обраду података који се добијају извршавањем симулација. Ово ће представљати алтернативан начин за мерење ефикасности поменутих електронских инфраструктура, са корисничке тачке гледишта.

На самом почетку кандидат ће се упознати са литературом која се тиче технологија које ће се користити током истраживања и које ће сукцесивно морати да савлада: дистрибуираним рачунарским инфраструктурама, паралелним програмским парадигмама, Монте Карло методом, функционалним интегралима, елементима семантичког веба итд. Такође, кандидат ће све време морати да ради на унапређивању својих рачунарских вештина да би ефикасно решавао постављене нумеричке и остале проблеме. Кандидат ће се паралелно бавити развојем и унапређивањем Грид инфраструктуре и решавањем HPC проблема.

У првој фази истраживања кандидат ће се бавити унапређивањем Грид инфраструктуре, елиминисањем „уских грла“ насталих због неоптималног функционисања инстанци WMS сервиса, као и унапређивањем оперативности Линукс кластера. На основу одабране архитектуре система развиће окружење за мониторинг инстанци централног Грид WMS сервиса коришћењем Линукс демона написаних у BASH скриптинг језику, Apache сервера, PHP језика и RRD базе. Коришћењем сличних технологија (Apache, PHP, HTTPS) ће бити развијена и апликација за креирање и управљање репозиторијумима Линукс пакета, а њени модули ће бити реализовани BASH скриптама. Алати за ефикасније администрирање Грид Линукс кластера ће такође бити развијани користећи скрипте написане за BASH. Тако добијене апликације и алати ће затим бити тестирани у продукционом Грид окружењу.

Паралелно са Грид истраживањем, кандидат ће радити на имплементацији Монте Карло алгоритма за решавање функционалних интеграла. Након упознавања са расположивим рачунарским архитектурама и компајлерима кандидат ће креирати паралелне верзије кода користећи MPI програмску парадигму, као и POSIX нити. У случају IBM Cell процесора биће потребно осмислити одговарајући приступ јер се његова хетерогена архитектура прилично разликује од осталих платформи. Затим ће бити потребно оптимизовати добијене кодове и компајлирати их на одговарајући начин (на основу стеченог знања о специфичностима доступних компајлера). Тако развијене верзије апликације ће затим бити тестиране, а времена извршавања и постигнута скалирања упоређена, и на основу тога биће истраживан утицај особина различитих архитектура и одабраних техника паралелизације и оптимизације на добијене перформансе.

У другој фази истраживања кандидат ће се бавити утицајем FFT нумеричких библиотека на перформансе HPC апликација. Он ће одредити и упоредити ефикасност доступних FFT библиотека у случају тродимензионалне Фуријеове трансформације на HPC рачунарским инфраструктурама. Након тога ће се упознати са FFT имплементацијом у Quantum Espresso програмском пакету за рачунање електронских структура и моделовање материјала и након тога интегрисати одабране FFT библиотеке у апликацију. Тако добијену апликацију ће затим тестирати и измерити њене перформансе уз коришћење различитих верзија FFT библиотека, а након тога анализирати утицај библиотека на време извршавања и скалирање апликације. У овој фази истраживања кандидат ће такође радити на *in situ* техници визуелизације података која ће, између осталог, омогућити и елиминисање уписа и читање података са диска (I/O) као „уског грла“ у процесу обраде и визуелизације података. Биће евалуирани доступни алати за приказ података и њихове *in situ* библиотеке, а затим ће одабрана библиотека бити имплементирана у BrainCore нумеричку симулацију и тестирана таква метода визуелизације и обраде података извршавањем инструментализоване апликације на HPC инфраструктури. У следећој фази кандидат ће радити на имплементацији окружења за интеграцију хетерогених репозиторијума пољопривредних података који ће бити конвертовани у RDF формат и затим постати доступни истраживачима кроз јединствену SPARQL приступну тачку. Кандидат ће развити и дизајнирати одговарајући RESTful интерфејс, евалуирати и одабрати имплементацију триплстор базе, а затим и прилагодити извршавању на Грид инфраструктури одговарајуће компоненте за конвертовање пољопривредних података и метаподатака из различитих формата у RDF и смештање тако добијених података у триплстор базу.

У последњој фази истраживања кандидат ће на основу иницијалне верзије GP-SCL кодова за решавање временски зависне нелинеарне Грос-Питаевски једначине у три димензије (који су паралелизовани OpenMP методом) развити хибридную MPI/OpenMP верзију која ће моћи да се извршава на кластерима, уз бољу скалабилност апликације. Перформансе тако добијених верзија кода (за пропагацију и у имагинарном и у реалном времену) ће затим бити тестиране на доступној HPC инфраструктури и дискутовани добијени резултати.

7. Закључак и предлог

Наведени научни резултати кандидата Владимира Славнића из области предложене теме докторске дисертације потврђују значај и актуелност предмета истраживања. Резултати имају изражену нумеричку компоненту и за њихову реализацију су били потребни значајни рачунарски ресурси доступни на Институту за физику у Београду у Лабораторији за примену рачунара у науци. Након увида у наведене резултате досадашњег рада кандидата, Комисија је донела одлуку да је он показао изражену способност за научноистраживачки рад и за реализацију планираних истраживања. Одлука је донета имајући у виду његове изузетне резултате постигнуте током истраживања на докторским студијама.

Такође, Комисија је мишљења да је предложена тема докторске дисертације „Примена дистрибуираних рачунарских инфраструктура за рачунарство високих перформанси” (енг. “Application of distributed computing infrastructures to high performance computing”) подобна у смислу наведених циљева рада, очекиваних резултата и научних доприноса, имајући у виду актуелна истраживања у одговарајућој области и могућност њене практичне примене.

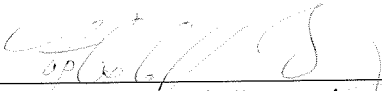
Комисија констатује да предложена тема припада области Рачунарска техника и информатика за коју је Електротехнички факултет у Београду матичан и за ментора рада се предлаже др Зоран Јовановић, редован професор, који је запослена на Електротехничком факултету у Београду. Због тога је уз овај извештај приложен и списак радова проф. др Зорана Јовановића, који га квалификују за ментора. Имајући у виду да је истраживање кандидата делом вршено у оквиру међународних пројеката, дисертација ће бити писана на енглеском језику.

Узимајући у обзир чињенице наведене у овом извештају, Комисија предлаже Наставно-научном већу Електротехничког факултета у Београду да прихвати тему докторске дисертације „Примена дистрибуираних рачунарских инфраструктура за рачунарство високих перформанси” (енг. “Application of distributed computing infrastructures to high performance computing”) и да кандидату Владимиру Славнићу омогући приступ њеној изради.

Београд, 17. август 2016. године

ЧЛАНОВИ КОМИСИЈЕ


др Зоран Јовановић, редовни професор
Универзитет у Београду – Електротехнички факултет


др Захарије Радивојевић, доцент
Универзитет у Београду – Електротехнички факултет


др Антун Балаж, научни саветник
Универзитет у Београду – Институт за физику у Београду


др Јелена Радовановић, редовни професор
Универзитет у Београду – Електротехнички факултет

