

Enis Afgan

Josipa Hatzea 29

10110 Zagreb

Croatia

Phone: ++ 385 (99) 200 85 85

E-mail: afgane@gmail.com

Web: <http://www.usecloudman.org/enis/>

EDUCATION

UNIVERSITY OF ALABAMA AT BIRMINGHAM

Birmingham, AL

Ph.D. in Computer Science (GPA: 4.0 on 4.0 scale)

May 2009

Major Field of Study: Grid Computing

Dissertation:

Utility Driven Grid Scheduling Framework

Dissertation Advisor: Dr. Purushotham Bangalore

Committee Members: Dr. Brandon Eames, Dr. Elliot Lefkowitz,

Dr. Anthony Skjellum, Dr. Alan Sprague

UNIVERSITY OF ALABAMA AT BIRMINGHAM

Birmingham, AL

Bachelor of Science in Computer Science (Honors Program)

December 2003

Final Undergraduate GPA: 3.94 on 4.0 scale (*summa cum laude*)

PELL CITY HIGH SCHOOL

Pell City, AL

High School Diploma

May 1999

PUBLICATIONS

Journal Papers

Afgan E., Baker D., Coraor C., Goto H., Paul M., Makova K., Nekrutenko A., Taylor J., "Harnessing cloud computing with Galaxy Cloud", *Nature Biotechnology*, vol 29, 2011.

Afgan E., Baker D., Taylor J., "A Reference Model for Deploying Applications in Virtualized Environments," *Concurrency and Computation: Practice and Experience*, first published online, 2011.

Goto H., Dickins B., Afgan E., Paul I., Taylor J., Makova K., Nekrutenko A., and the Galaxy Team, "Dynamics of mitochondrial heteroplasmy in three families: A fully reproducible re-sequencing study", *Genome Biology*, vol. 12, issue R59, 2011.

Afgan E., Bangalore P., Skala T., "Scheduling and Planning Job Execution of Loosely Coupled Applications," *Journal of Supercomputing*, Online First, Feb 2011.

Afgan E., Bangalore P., Skala K., "Application Information Services for Distributed Computing Environments", *Journal of Future Generation Compute Systems*, Volume 27, Issue, 2, p. 173-181, 2011.

Afgan E., Baker D., Coraor N., Chapman B., Nekrutenko A., Taylor J., "Galaxy CloudMan: Delivering Cloud Compute Clusters," *BMC Bioinformatics*, Vol 11, Issue 12, 2010.

Afgan E., Bangalore P., "Exploiting Performance Characterization of BLAST in the Grid", *The Journal of Cluster Computing*, Volume 13, Issue 4, pp. 385-395, 2010.

Afgan E., Bangalore P., "Dynamic BLAST – a Grid Enabled BLAST", *The International Journal of Computer Science and Network Security*, Volume 9, Issue 4, pp. 149-157, 2009.

Book Chapters

Afgan E., J. Goecks, D. Baker, N. Coraor, the Galaxy Team, A. Nekrutenko, and J. Taylor, "Galaxy - a Gateway to Tools in e-Science," *Guide to e-Science: Next Generation Scientific Research and Discovery*, K. Yang, Ed., ed: Springer, 2011, p. 145-177.

Afgan E., Bangalore P., Gray J., "A Domain-Specific Language for Describing Grid Applications", *Software Applications: Concepts, Methodologies, Tools, and Applications*, Ed. Pierre F. Tiako, March, 2009. (derived from the original publication)

Afgan E., Bangalore P., "A framework to assist users with planning and scheduling jobs on the grid", Udoh E., Wang F. (Eds), *Handbook of Research on Grid Technologies and Utility Computing: Concepts for Managing Large-Scale Applications*, pp. 22-31, May 2009.

Afgan E., Bangalore P., Gray J., "Configuring Grid Applications from Higher Level Domain-Specific Languages", *Designing Software-Intensive Systems: Methods and Principles*, Ed. Pierre F. Tiako, pp. 402-439, July, 2008.

Conference Papers

Afgan E., Bangalore P., Duncan D., "GridAtlas - A Grid Application and Resource Configuration Repository and Discovery Service", *Int. Conference on Cluster Computing*, New Orleans, LA, pp.1-10, August 31-September 4, 2009.

Bangalore P., Afgan E., "G-BLAST: A Grid Service for BLAST", *Int. Conference on Grid Computing and Applications (GCA) 2008*, Las Vegas, NV, pp. 264-270, July 14-17, 2008.

Afgan E., Bangalore P., "Application Specification Language (ASL) - A Language for Describing Applications in Grid Computing", *Int. Conference on Grid Service Engineering and Management (GSEM) 2007*, Leipzig, Germany, pp. 24-38, September 24-26, 2007.

Afgan E., Bangalore P., "Performance Characterization of BLAST on the Grid", *IEEE 7th Int. Symposium on Bioinformatics & Bioengineering (BIBE) 2007*, Boston, MA, pp. 1394-1398, October 14-17, 2007.

Afgan E., Sathyanarayana P., Bangalore P., "Dynamic Task Distribution in the Grid for BLAST", *IEEE Int. Conference on Granular Computing (GrC) 2006*, Atlanta, GA, pp. 554- 557, May 10-12, 2006.

Afgan E., Jones W., "Design, Development and Usage of a Generic Job Submission Grid Service", *44th ACM Southeast Conference*, Melbourne, FL, pp. 738 – 739, March 10-12, 2006.

Afgan E., Velusamy V., Bangalore P., "Grid Resource Broker with Application Profiling and Benchmarking", *European Grid Conference (EGC) 2005*, Amsterdam, Netherlands, pp. 691-701, February 14-16, 2005.

Afgan E.: "Role of the Resource Broker in the Grid". *42nd Annual ACM Southeast Conference*, Huntsville, AL, pp. 299 – 300, April 2-3, 2004.

Refereed Workshop Papers

Afgan E., Bangalore P., "Embarrassingly Parallel Jobs Are Not Embarrassingly Easy to Schedule on the Grid", *SC08 Int. Conference for High Performance, Networking, Storage and Analysis - Workshop on Many-Task Computing on Grids and Supercomputers*, Austin, TX, pp. 1-10, November 15-21, 2008.

Afgan E., Bangalore P., "Experiences with developing and deploying Dynamic BLAST", *Mardi Gras Conference 2008 - Workshop on Grid-Enabling Applications*, New Orleans, LA, pp. 1-10, January 31-February 2, 2008.

Afgan E., Gray J., Bangalore P., "Using Domain-Specific Modeling to Generate User Interfaces for Wizards", *Int. Conference on Model Driven Engineering Languages and Systems (MODELS) - Workshop on Model Driven Development of Advanced User Interfaces (MDDAUI) 2007*, Nashville, TN, September 30-October 5, 2007.

Afgan E., Bangalore P., "Computation Cost in Grid Computing Environments", *Int. Conference on Software Engineering (ICSE) - First International Workshop on the Economics of Software and Computation 2007*, Minneapolis, MN, May 20, 2007.

Invited Talks

Title: "Experiences and Techniques for Grid-Enabling Applications", Clemson University, Clemson, SC, January 25, 2008.

Title: "Razvoj grid aplikacija – Dynamic BLAST iskustvo", Institut Ruder Boskovic, Zagreb, Croatia, December 18, 2007.

Talks

Title: "CloudMan: Galaxy on the Cloud"
Galaxy Community Conference, Lunteren, the Netherlands, May 26, 2011.

Title: "Dynamically Scalable, Accessible Analysis for High-Throughput Sequence Data"
Bio-IT World, Boston, MA, April 13, 2011.

Title: "NGS Analyses with Galaxy on the Cloud"
Intelligent Systems for Molecular Biology (ISMB), Boston, MA, July 12, 2010. (live demo)

Title: "Deploying Galaxy on the Cloud"
Bioinformatics Open Source Conference (BOSC), Boston, MA, July 9, 2010.

Title: "A Framework for Efficient Execution of Bioinformatics Applications across the Grid,"
MidSouth Computational Biology and Bioinformatics Society (MCBIOS), Starkville, MS, February 21, 2009.

Best Oral Presentation award

Title: "UABgrid: Practice and Experience,"
Open Grid Forum (OGF) 22, Boston, MA, February 26, 2008.

Title: "Language for Describing Application Software in Grid Computing,"
Alabama Academy of Science 2007 at the Tuskegee University, Tuskegee, AL, February 28-March 2, 2007.

Title: "UABgrid Dynamic BLAST: Searching Nucleotide and Protein Databases Using SURAGrid,"
Internet 2 Meeting - Fall 2006, Chicago, IL, December 3-7, 2006.

Title: "Resource Brokering in a Grid Computing Environment,"
ACM Mid-Southeastern Conference, Gatlinburg, TN, November 11-12, 2004.

Title: "Role of the Resource Broker in the Grid,"
Austin Peay State University, Gatlinburg, TN, November 21-22, 2003.

Title: "Adaptive Web Based Resource Broker for the Grid,"
University of Montevallo, Montevallo, AL, March 17-20, 2004.

Abstracts

Afgan E., the Galaxy Team, Nekrutenko A., Taylor J., "Enabling NGS Analysis with(out) the Infrastructure", Bioinformatics Open Source Conference (BOSC 2011), Vienna, July 17-19, 2011.

Afgan E., Baker D., Coraor N., The Galaxy Team, Nekrutenko A., Taylor J., "Deploying Galaxy on the Cloud", 18th International Conference on Intelligent Systems and Molecular Biology (ISMB) 2010, Boston, MA, July 9-13, 2010.

Halappanavar M., Robinson J.P., Afgan E., Yafchak M.F., Bangalore P., "A common application platform for the SURAGrid (CAP)", Mardi Gras Conference 2008 - Workshop on Grid-Enabling Applications, New Orleans, LA, January 31-February 2, 2008.

Afgan E., Bangalore P., "Dynamic BLAST - An Approach to Dynamic Grid Application Development", *GlobusWORLD 2006*, Washington, D.C., September 11-15, 2006.

Afgan E., Bangalore P., "Extensible Resource Broker for the Globus Toolkit", *GlobusWORLD 2005*, Boston, MA, February 7-11, 2005.

Technical Reports

Afgan E., Bangalore P., "Application Specification Language (ASL)", Technical Report, UABCIS-TR-2007-0123-1, Collaborative Computing Laboratory, University of Alabama at Birmingham, Birmingham, AL, January 23, 2007.

Afgan E., Bangalore P., "Effective Utilization of the Grid with the Grid Application Deployment Environment (GADE)", Technical Report, UABCIS-TR-2005-0601-1, Collaborative Computing Laboratory, University of Alabama at Birmingham, Birmingham, AL, June 1, 2005.

Posters

Afgan E., Baker D., The Galaxy Team, Nekrutenko A., Taylor J., "The Elastic Analysis with Galaxy on the Cloud," Beyond the Genome, Boston, MA, Oct 11-13, 2010.

Afgan E., "Assisting users with planning and scheduling jobs on the grid," IEEE International Parallel & Distributed Processing Symposium (IPDPS) 2008 TCPP PhD Forum, Miami, FL April 14-18, 2008.

Afgan E., Bangalore P., "Dynamic BLAST – an Effective and Efficient BLAST Wrapper for the Grid," IEEE 7th International Symposium on Bioinformatics & Bioengineering (IEEE BIBE 2007), Boston, MA, October 14-17, 2007.

ACADEMIC SERVICE

January 2004 – May 2009

Spring 2009

Class Instructor: CS 333 - UNIX Operating System Fundamentals - I had complete responsibility for the class thought.

Lab Instructor and Primary TA: CS 201 – Introduction to Object Oriented Programming

Fall 2008

Lab Instructor and TA: CS 201 – Introduction to Object Oriented Programming

Primary TA: CS 633/733 –Grid Computing

Fall 2006

Class Instructor: CS 333 - UNIX Operating System Fundamentals - I had complete responsibility for the class thought.

Primary TA: CS 431 - Distributed Computing

Primary TA: CS 432 - Parallel Computing

Secondary TA: CS 101 - Computing Fundamentals

May 2006

Class Instructor: CS 101 - Computing Fundamentals – I had complete responsibility for the class thought.

Fall 2005

Primary TA: CS 101 - Computing Fundamentals

Secondary TA: CS 440 - Operating Systems

Fall 2004

Primary TA: CS 101 - Computing Fundamentals

Secondary TA: CS 350 - Automata and Formal Language Theory

HONORS AND AWARDS

Outstanding Graduate Student at Doctoral Level in the CIS department (May 2009)

Best Oral Presentation award at the Mid-South Computational Biology and Bioinformatics Society (MCBIOS) 2009 (February 2009)

Travel award for IEEE International Parallel & Distributed Processing Symposium (IPDPS) 2008 , which was awarded only to top student submission selected to present their work at PhD Forum (April 2008)

Student travel grant for International Conference on Software Engineering (ICSE) 2007 (May 2007)

1st place UAB Tennis Intramurals Tournament - Intermediate Category (Spring 2006)

2nd place at UAB Graduate Student Research Days (March 2006)

Passed Ph.D. qualifying exam with distinction (January 2005)

2nd Place at Doctoral Level at the ACM Mid-Southeast Conference (November 2004)

Graduated *Summa cum Laude* (December 2003)

Graduated with *Honors in Computer Science* (December 2003)

Phi Kappa Phi Honor Society

Nominee for UAB International Scholar and Student Services Academic Excellence Award (Fall 2002)

Departmental Award for the Department of Computer and Information Sciences (Fall 2001)

Phi Eta Sigma Honor Society (Fall 2000)

Golden Key National Honor Society (Spring 2000)

EXPERIENCE

Research Scientist: Institute Ruđer Bošković (IRB), CIR

September 2011 – present

Director of the CloudMan open source project (www.usecloudman.org)

Involved in writing grants and projects

Postdoctoral Research Fellow: Emory University, Taylor Lab*August 2009 – August 2011*

In charge of enabling the Galaxy application and its internal execution model to be automatically and transparently deployed on cloud computing resources (www.galaxyproject.org/cloud)

Devised a reference model for deploying applications into virtualized environments

Developed an autonomously deployable and automatically scalable system for the Galaxy application called Galaxy CloudMan

Developed methods for automated deployment of applications and required dependencies onto distributed computational infrastructures

Setup a cloud computing infrastructure based on Euacalyptus for use by the Galaxy project.

Research Assistant: University of Alabama at Birmingham, CCL group*January 2004 – August 2009*

Involved in use, enablement and advancement of grid related technologies: Globus Toolkit, GridWay, DRMAA, GridSim, Grid services, OGF, OGCE, GridSphere

Introduced new concepts into the field of grid application execution and scheduling enabling shorter turnaround times of submitted jobs (up to 50%) through adoption of principles in Artificial Intelligence, Statistics, and Application Optimizations

Experience in grid enablement and execution optimization of domain specific applications such as BLAST and R.

Assisted in setup and configuration of campus wide grid (UABgrid) and regional grid (SURAGrid), dealing with software stack installation and access method enablement through portal customization and portlet creation (e.g., OGCE).

Taught a junior level class on UNIX principles. Broad experience in Grid Computing and Distributed Systems through a range of teaching and grading assignments.

Designed, managed and guided graduate and undergraduate students in software development: Web portal development using web service technologies with a back-end supporting grid application registration and information retrieval as well as creation of automated grid job submission interfaces.

PROJECTS

Galaxy ObjectStore – a Galaxy instance needs to manage terabytes of data, which has grown to be an infrastructure management substantial problem. The devised and developed ObjectStore is a middleware layer that abstracts the backend hardware storage from the application layer using the storage. The end result of the ObjectStore is that the storage back end can be replaced (and even swapped) for several storage options, for example AWS Simple Storage Service (S3), thus isolating the application from having to manage the support infrastructure. I developed the ObjectStore with its support for local file systems, a cached file system, and S3.

mi-deployment – I devised and developed a process for automated infrastructure and configuration management. The project enables seamless deployment of complex infrastructures and configurations, reproducibility of the environment and follows the deployment model supported by DevOps movement. The project has been in use by the Galaxy Project and largely embedded into the CloudBioLinux project.

Galaxy CloudMan – design, portability, and efficiency study as well as implementation of autonomous web-based application for seamless use of cloud infrastructures. I had to resolve issues regarding application data persistence, infrastructure and application management, and cloud usability transparency. Specifically, I worked with Amazon Web Services (AWS) and Eucalyptus infrastructures and accompanying tools (e.g., boto, euca2ools, RabbitMQ, SGE).

OptionView, a Grid Metascheduler Implementation – design and implementation of a grid metascheduler that advances user's interaction method across grid environments on individual job basis. Derived methods and

tools for effective planning and execution of application jobs on real-world grid resources resulting in significant alteration and improvement to user experience and Quality of Service (QoS).

Application Information Services (AIS) – devised and developed a set of core grid services. AIS enable collection and retrieval of relevant application- and resource-specific information through a well-defined format in a known access location and enables realization of application- and user-oriented metascheduling.

Application Performance Database (AppDB) – design and implementation of a repository for collecting and storing application level execution characteristics of previous application executions on the grid, which enables development of more efficient job scheduling policies and mechanisms.

GridAtlas – global architecture design and development of a service that hides and automates the process of keeping track of installation properties of any one application across grid resources. Availability of this service enables automatic job submission to grid resources by tools such as GridWay.

Application Specification Language (ASL) – designed and created an XML language used to describe functionality and options of individual applications in heterogeneous grid environments. The language enables grid schedulers and job submission tools to automatically learn about application preferences and thus enable application oriented scheduling and job submission. Also developed a meta-modeling tool to ease composition of ASL documents.

Java implementation of Dynamic BLAST – a multi-threaded, master-worker, grid enabled wrapper for NCBI BLAST that leverages resource heterogeneity to reduce job execution time. The application focuses on parameterization of individual tasks to best match capabilities of the application and the resource resulting in job runtime reduction of up to 50% through 40% resource utilization increase.

Metamodel construction and C++ implementation of code generator – automatic generation of code for wizards from various composed instance models under the constraints of the metamodel

Java based implementations using reflective technologies, namely AspectJ, Javassist and OpenJava, to create a basic application debugger.

C++ implementations of MPI based parallel matrix-matrix multiplication algorithms and LU decomposition – Cannon, Fox, and Broadcast-broadcast algorithms for matrix-matrix multiplication

Java implementation of generic Resource Broker – using fuzzy logic to perform grid application-specific resource selection making use of software design patterns

Adaptation of OGCE grid portal using Java to reflect needs of applications developed for UABGrid

SOFTWARE SKILLS

<i>Operating Systems</i>	Linux, OS X
<i>Programming Languages</i>	Python, Java, AspectJ, C++
<i>Grid/Cloud Technologies</i>	Boto, AWS, euca2ools, GridWay, DRMAA, Globus Toolkit
<i>Markup Languages</i>	HTML, Mako
<i>Web Technologies</i>	Django, Google App Engine

SERVICE RECORD

Member and one of the founders of the Green Initiative at UAB, Summer 2007 – Spring 2009

Student volunteer for International Conference on Software Engineering (ICSE) May 2007

Senator for the Graduate Student Association at UAB	<i>Fall 2004 – Summer 2007</i>
Ambassador for the student chapter of ACM at UAB	<i>Spring 2005 – Spring 2007</i>
Student volunteer at Supercomputing 2005 (SC'05)	<i>November 2005</i>
President for the student chapter of ACM at UAB	<i>Spring 2004 – Spring 2005</i>
Webmaster for the student chapter of ACM at UAB	<i>Fall 2002 – Fall 2003</i>
Participated at the regional ACM Programming Contest in Daytona, FL	<i>November 2003</i>

RELATED COURSEWORK

Grid Computing, Parallel Computing, Internetworking, Reflective and Adaptive Systems, Software Engineering, Research Methods, Computer Systems, Database Systems, Numerical Computing, Automata, Language and Computation, Artificial Intelligence, Operating Systems, Calculus, Linear Algebra

MEMBERSHIPS

International Society for Computational Biology (ISCB)	<i>2010 - present</i>
Mid-South Computational Biology and Bioinformatics Society (MCBIOS)	<i>2009</i>
Society for Industrial and Applied Mathematics (SIAM)	<i>2009</i>
Global Grid Forum (GGF) / Open Grid Forum (OGF)	<i>2006, 2008</i>
Association for Computing Machinery (ACM)	<i>2003 – 2009</i>
IEEE Computing Society	<i>2004 – 2009</i>

MISCELLANEOUS

Foreign Languages

Fluent in Croatian (including other regional languages), German

Hold Wilderness First Responder (WFR) Medical Certification (certified January 2009)

UAB Outdoor Pursuits Trip Leader (2005 – 2009)

Volunteer

Volunteer at UAB CIS High School Programming Contest (2005 through 2007)

Student Volunteer at International Conference on Software Engineering (ICSE 2007) (May 2007)

Volunteer at K-12 Computer Science Workshop at UAB CIS Department (July 2006)

Student Volunteer at Super Computing 2005 (SC'05) (November 2005)

Volunteer for Unbridled Joy Special Equestrian Program at Alabama Institute for Deaf and Blind (AIDB) at Talladega, AL. I worked in stables and horse management tasks as well as worked with riders with disabilities (2003).

Hobbies

Photography, running, sailing, tennis, windsurfing, hiking, snow skiing, water skiing, mountain biking