

Enis Afgan

359 Homeland Southway, Apt. 3A
Baltimore, MD 21212

USA

Phone: ++ 1 678 561 3647

E-mail: enis.afgan@gmail.com

Web: <http://cloudman.irb.hr/enis/>

<https://www.linkedin.com/in/afgane>

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

Stewart, Craig A, David Y Hancock, Therese Miller, Jeremy Fischer, R Lee Liming, George Turner, John Michael Lowe, et al. "Jetstream: A Novel Cloud System for Science." In *Contemporary High Performance Computing*, 189–222. CRC Press, 2019.

Jalili, Vahid, Enis Afgan, James Taylor, and Jeremy Goecks. "Cloud Bursting Galaxy: Federated Identity and Access Management." *Bioinformatics*, Jun 2019.

Hancock, David Y, Craig A Stewart, Matthew Vaughn, Jeremy Fischer, John Michael Lowe, George Turner, Tyson L Swetnam, et al. "Jetstream—Early Operations Performance, Adoption, and Impacts." *Concurrency and Computation: Practice and Experience* 31, no. 16 (2019): e4683.

Afgan, E., Lonie, A., Taylor, J., Goonasekera, N., "CloudLaunch: Discover and Deploy Cloud Applications", *Future Generation Computer Systems (FGCS)*, June 2018.

Hancock, D., Stewart, C., Vaughn, M., Fischer, J., Lowe, M., Turner, G., Swetnam, T., Chafin, T., Afgan, E., Pierce, M., Snap-Childs, W., "Jetstream—Early Operations Performance, Adoption, and Impacts", *Concurrency and Computation: Practice and Experience (CCPE)*, Dec 2017, DOI: 10.1002/cpe.4683.

Afgan, E., Baker, D., Batut, B., Beek, M., Bouvier, D., Čech, M., Chilton, J., Clements, D., Coraor, N., Grüning, B., Guerler, A., Hillman-Jackson, J., Hiltemann, S., Jalili, V., Rasche, H., Soranzo, N., Goecks, J., Taylor, J., Nekrutenko, A., Blankenberg, D., "The Galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2018 update", *Nucleic Acids Research (NAR)*, DOI: 10.1093/nar/gky379, May 2018.

Kim, B, Ali, T, Lijeron, C, Afgan, E, Krampis, K, "Bio-Docklets: virtualization containers for single-step execution of NGS pipelines", *Gigascience*, Vol 6, Issue 8, 10.1093/gigascience/gix048, June 2017.

Forer, L., Afgan, E., Weissensteiner, H., Davidovic, D., Specht, G., Kronenberg, F., Schoenherr, S., "Cloudflow - enabling faster biomedical pipelines with MapReduce and Spark", *Scalable Computing: Practice and Experience (SCPE)*, Vol 17, Issue 2, pp. 103-114, June 2016.

Afgan, E., Baker, D., Beek, M., Blankenberg, D., Bouvier, D., Čech, M., Chilton, J., Clements, D., Coraor, N., Eberhard, C., Grüning, B., Guerler, A., Hillman-Jackson, J., Kuster, G., Rasche, E., Soranzo, N., Turaga, N., Taylor, J., Nekrutenko, A., Goecks, J., "The Galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2016 update", *Nucleic Acids Research (NAR)*, May 2016.

Skala, K., Davidović, D., Afgan, E., Sović, I., Šojat, Z., "Scalable Distributed Computing Hierarchy: Cloud, Fog and Dew Computing", *Open Journal of Cloud Computing (OJCC)*, Vol 2, Issue 1, pp. 16-24, Dec 2015.

Afgan, E., Sloggett, C., Goonasekera, N., Makunin, I., Benson, D., Crowe, M., Gladman, S., Kowsar, Y., Pheasant, M., Horst, R., Lonie, A., "Genomics Virtual Laboratory: A Practical Bioinformatics Workbench for the Cloud", *Plos ONE*, Oct 2015.

Afgan, E., Coraor, N., Chilton, J., Baker, D., Taylor J., "Enabling Cloud Bursting for Life Sciences within Galaxy", *Concurrency and Computation: Practice and Experience*, Vol 27, Issue 16, 2015.

Möller, S., Afgan, E., Banck, M., Bonnal, R., Booth, T., Chilton, J., Cock, P., Gumbel, M., Harris, N., Holland, R., Kalaš, M., Kaján, L., Kibukawa, E., Powel, D., Prins P., Quinn, J., Sallou, O., Strozzi, F., Seemann, T., Sloggett, C., Soiland-Reyes, S., Spooner, W., Steinbiss, S., Tille, A., Travis, A., Guimera R., Katayama, T., Chapman, B., "Community-driven development for computational biology at Sprints, Hackathons and Codefests", *BMC Bioinformatics*, Vol 15, Issue 7, 2014.

Leo, S., Pireddu, L., Cuccuru, G., Lianas, L., Soranzo, N., Afgan, E., Zanetti, G., "BioBlend.objects: metacomputing with Galaxy", *Bioinformatics*, Vol 30, Issue 19, 2014.

Blankenberg, D., Kuster, G.V., Bouvier E., Baker, B., Afgan, E., Stoler N., Rebolledo-Jaramillo B., the Galaxy Team, Taylor, T., Nekrutenko, A., "Dissemination of scientific software with Galaxy ToolShed: A case of mitochondrial RNA-modification", *Genome Biology*, Vol 15, Issue 2, 2014.

Sloggett, C., Goonasekera, N, Afgan, E., "BioBlend: automating pipeline analyses within Galaxy and CloudMan", *Bioinformatics*, Vol 29, Issue 13, 2013.

Möller S., Afgan E., Banck M., Cock PJA, Kalas M., Kajan Laszlo, Prins P., Quinn J., Sallou O., Strozzi F., Seemann T., Tille A., Guimera RV, Katayama T., Chapman B., "Sprints, Hackathons and Codefests as community gluons in computational biology", *EMBnet.journal*, Vol 19, Issue B, Oct 2013.

Afgan, E., Chapman, B., Taylor, J., "CloudMan as a platform for tool, data, and analysis distribution", *BMC Bioinformatics*, Vol 13, 2012.

Afgan, E., Chapman, B., Jadan, M., Franke, V., Taylor, J., "Using Cloud Computing Infrastructure with CloudBioLinux, CloudMan, and Galaxy", *Current Protocols in Bioinformatics*, Vol 38, 2012.

Afgan E., Baker D., Nekrutenko A., Taylor J., "A Reference Model for Deploying Applications in Virtualized Environments," *Concurrency and Computation: Practice and Experience*, Vol 24, Issue 12, 2012.

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.

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*, Vol 59, Issue 3, pp. 1431-1454, Feb 2011.

Afgan E., Bangalore P., Skala K., "Application Information Services for Distributed Computing Environments", *Journal of Future Generation Compute Systems*, Vol 27, Issue 2, pp. 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*, Vol 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*, Vol 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*, Ed. K. Yang, Springer, Chapter 6, pp. 145-177, April 2011.

Afgan E., Bangalore P., "A framework to assist users with planning and scheduling jobs on the grid", *Handbook of Research on Grid Technologies and Utility Computing: Concepts for Managing Large-Scale Applications*, Eds. Udoh E., Wang F., Chapter 3, pp. 22-31, May 2009.

Afgan E., Bangalore P., Gray J., "A Domain-Specific Language for Describing Grid Applications", *Software Applications: Concepts, Methodologies, Tools, and Applications*, Ed. Pierre Tiako, Idea Group, Chapter 13, pp. 402-438, March 2009.

Conference Papers

Afgan, E., Jalili, V., Goonasekera, N., Taylor, J., Goecks, J., "Federated Galaxy: Biomedical Computing at the Frontier", *IEEE International Conference on Cloud Computing (IEEE CLOUD 2018)*, San Francisco, CA, July 2018.

Afgan, E., Christie, M., Goonasekera, N., "CloudLaunch as a Gateway for Discovering and Launching Cloud Applications", *Gateways 2017*, Ann Arbor, MI, Oct 2017.

Goonasekera, N., Lonie, A., Taylor, J., Afgan, E., "CloudBridge – a Simple Cross-Cloud Python Library", *XSEDE 16*, Miami, FL, July 2016.

Afgan, E., Lonie, A., Taylor, J., Skala, K., Goonasekera, N., "Architectural Models for Deploying and Running Virtual Laboratories in the Cloud", *39th International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO)*, Opatija, Croatia, IEEE, June 2016.

Forer, L., Afgan, E., Weissensteiner, H., Davidović, D., Specht, G., Kronenberg, F., Schönherr, S., “Cloudflow – A Framework for MapReduce Pipeline Development in Biomedical Research”, 38th *International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO)*, Opatija, Croatia, IEEE, May 2015.

Afgan, E., Krampis, K., Goonasekera, N., Skala, K., Taylor, J., “Building and Provisioning Bioinformatics Environments on Public and Private Clouds”, 38th *International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO)*, Opatija, Croatia, IEEE, May 2015.

Forer, L., Lipić, T., Schönherr, S., Weißensteiner, H., Davidović, D., Kronenberg, F., Afgan, E. “Delivering bioinformatics MapReduce applications in the cloud”, 37th *International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO)*, Opatija, Croatia, IEEE, May 2014.

Kowsar, Y., Afgan, E., “Support for data-intensive computing with CloudMan”, 36th *International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO)*, Opatija, Croatia, IEEE, May 2013.

Booth, T., Bica, M.; Hyun Soon Gweon; Field, D.; Afgan, E., “Bio-Linux as a tool for bioinformatics training”, 12th *International Conference on Bioinformatics & Bioengineering (BIBE)*, Larnaca, Cyprus, IEEE, November 2012.

Afgan, E., Skala, K., Davidović, D., Lipić, T., Sović, I., “CloudMan as a tool execution framework for the cloud”, 35th *International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO)*, Opatija, Croatia, IEEE, 2012.

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”, 7th *Int. Symposium on Bioinformatics & Bioengineering (BIBE) 2007*, Boston, MA, pp. 1394-1398, IEEE, 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

Pierce, M., Miller, M., Brookes, E., Wong, M., Liu, Y., Afgan, E., Gesing, S., Dahan, M., Marru, S., Walker, T., "Towards a Science Gateway Reference Architecture", *10th International Workshop on Science Gateways (IWSG)*, Edinburgh, Scotland, June 2018.

Afgan, E., Baker, D., Chilton, J., Coraor, N., Taylor, J., "Galaxy Cluster to Cloud - Genomics at Scale", *SC14 - Int. Conference for High Performance, Networking, Storage and Analysis - GCE: The 9th Gateway Computing Environments Workshop*, New Orleans, LA, pp. 47-50, November 2014.

Lipić, T., Skala, K., Afgan, E., "Deciphering Big Data Stacks: An Overview of Big Data Tools", *SC14 - Int. Conference for High Performance, Networking, Storage and Analysis - Workshop on Big Data Analytics: Challenges and Opportunities (BDAC-14)*, New Orleans, LA, November 2014.

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

"Horizontal Scaling with Galaxy", Galaxy Africa conference, Cape Town, South Africa, April 2018.

"Delivering Data Analysis Environments on Multiple Clouds", Converged IT & the Cloud conference, San Francisco, Feb 2018.

"A cloud-agnostic stack", University of Alabama at Birmingham, Birmingham, AL, Nov 2017.

"The Pulse of Cloud Computing, with Bioinformatics as an Example", University of Colombo, Colombo, Sri Lanka, February 2018.

"HPC/Cloud based Business in Bioinformatics for SMEs", HPC for SME Forum, MIPRO 2016, Opatija, Croatia, June 2016.

"Building and provisioning genomics platforms on the world's clouds", University of Heidelberg, Heidelberg, Germany, April 2016.

"Democratizing the ability to analyze genomic data", Medical University of Innsbruck, Innsbruck, Austria, February 2015.

"Panel discussion at The 9th Gateway Computing Environments Workshop", SC14 Int. Conference for High Performance, Networking, Storage and Analysis, New Orleans, LA, November 2014.

"Setting up and administering Galaxy - panel discussion", Galaxy Australasia Workshop (GAW), Melbourne, Australia, March 2014.

"Cloud Bio-Linux", Center for Ecology and Hydrology (CEH), Wallington, UK, February, 2014.

"Making Genomics Accessible: Practice & Experience", Penn State University, State College, PA, November 2013.

"Genomics Virtual Laboratory", Australian Physiological Society (AuPS), Sydney, Australia, December 4, 2012.

"Uber Dojo: Advanced Black Belt Event for Tools & Data in the Cloud", NeCTAR & ANDS, Melbourne, Australia, September 24, 2012.

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

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

Organized workshops

Afgan, E., Beek, M., Gruening, B., "The Galaxy Docker Project", training workshop at the Galaxy Community Conference, Bloomington, IN, July 2016.

Afgan, E., Goonasekera, N., Turaga, N., "Get your own Galaxy within minutes", training workshop at the Galaxy Community Conference, Bloomington, IN, July 2016.

Afgan, E., Coraor, E., "How to use Galaxy Ansible Playbooks", training workshop at the Galaxy Community Conference, Bloomington, IN, July 2016.

Afgan, E., Goonasekera, N., "Galaxy on the Cloud: build it", training workshop at the Galaxy Community Conference, Bloomington, IN, July 2016.

Afgan, E., Davidović, D., Čubrić, J., Skala, K., "Bioinformatics Methods in Genomics", Rudjer Boskovic Institute (RBI), Zagreb, Croatia, March 2015.

Krampis, N., Afgan, E., Sanka R., Chapman, B., "Scriptable Bioinformatics Cloud Infrastructures with Cloud BioLinux, CloudMan & Galaxy", Galaxy Community Conference (GCC), Baltimore, MD, July 2014.

Afgan, E., Baker, D., "Running Galaxy on the Cloud", Galaxy Community Conference (GCC), Oslo, Norway, July 2013. □

Nekrutenko, A., Afgan, E., "Biomedical Data Analysis with Galaxy", European Human Genetics Conference (EHGC) 2013, Paris, France, June 2013.

Afgan, E., Sloggett, C., Lonie, A., Pheasant, M., "Genomics Virtual Laboratory Workshop", eResearch Australasia, Sydney, Australia, December, 2012. □

Afgan, E., Baker, D., "Galaxy CloudMan", Galaxy Community Conference (GCC), Chicago, IL, July 2012.

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 SURAggrid,"
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., Heydarian, Mo., "Resource planning on the Cloud: exploring the scalability spectrum", *Galaxy Australasia Meeting, Melbourne, Australia*, Feb 2017.

Afgan, E., Chilton, J., Baker, D., Coraor, N., Taylor, J., "Genomics Across Clouds with Galaxy", *NSF Cloud Workshop: Experimental Support for Cloud Computing*, Arlington, VA, December 2014.

Afgan, E., "CloudMan Project", *Galaxy Australasia Workshop (GAW)*, Melbourne, Australia, March 2014.

Sloggett, C., Goonasekera, N., Afgan, E., "BioBlend - Enabling Pipeline Dreams", *Bioinformatics Open Source Conference (BOSC)*, Berlin, Germany, July, 2013.

Kowsar, Y., Afgan, E., "Towards Enabling Big Data and Federated Computing in the Cloud", *Bioinformatics Open Source Conference (BOSC)*, Berlin, Germany, July, 2013.

Afgan, E., Sloggett, C., Lonie, A., Pheasant, M., "Establishing a National Genomics Virtual Laboratory with Galaxy CloudMan", *Galaxy Community Conference (GCC)*, Chicago, IL, July 2012. □

Afgan, E., Chapman, B., Krampis, K., Taylor, J., “Zero to a Bioinformatics Analysis Platform in Four Minutes”, *Bioinformatics Open Source Conference (BOSC)*, Long Beach, CA, July, 2012.

Afgan E., the Galaxy Team, Nekrutenko A., Taylor J., “Enabling NGS Analysis with(out) the Infrastructure”, *Bioinformatics Open Source Conference (BOSC)*, Vienna, July, 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)*, Boston, MA, July, 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, 2006.

Afgan E., Bangalore P., “Extensible Resource Broker for the Globus Toolkit”, *GlobusWORLD 2005*, Boston, MA, February, 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., Chapman, B., Krampis, N., Taylor, J., “Zero to a Bioinformatics Analysis Platform in Four Minutes”, *Bioinformatics Open Source Conference (BOSC)*, Long Beach, CA, July, 2012.

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)*, TCPP PhD Forum, Miami, FL April 14-18, 2008.

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

AWARDED GRANTS

“An Integrative and Collaborative Genomic Data Analysis Service with Galaxy”

National Human Genome Research Institute (NHGRI), National Institutes of Health (NIH)

09/2019 – 08/2020 (PI), Project number: 1R41HG010982-01

We are exploring new models for sustainability of the Galaxy project and this is a collaborative project with the Johns Hopkins University via the NIH Small Business Technology Transfer (STTR) program to provide commercial services supporting the Galaxy application.

“Implementing the Genomic Data Science Analysis, Visualization, and Informatics Lab-Space (ANVIL)”

National Human Genome Research Institute (NHGRI), National Institutes of Health (NIH)

09/2018 – 06/2023 (Co-I), Project number: 1U24HG010263-01

This large and collaborative project focuses on delivering a managed, cloud-based platform for a number of NHGRI's datasets (e.g., CCDG, CMG, eMERGE) and some of the most popular analysis applications (e.g., Galaxy, Bioconductor, Jupyter, RStudio).

"Securing Science Gateway Cyberinfrastructure with Custos"

Office of Advanced Cyberinfrastructure (OAC), National Science Foundation (NSF)

08/2018 – 07/2021 (Co-PI), Project number: 1840003

The aim of this project is to design, implement, and maintain a software service that improves science gateways' management of users' identities, secrets, and groups and sharing permissions on digital research objects created by science gateway users.

Jan 2014 - Dec 2015 (PI, 50%)

"Scalable Big Data Bioinformatics Analysis in the Cloud"

MZOS (Croatian Ministry of Science, Education and Sports)

€ 20,000

Jun 2011 - May 2015 (100%)

"Application Information Services for Distributed Computing Environments"

Marie Curie Actions - Incoming International Fellowship (IIF), European Commission FP7

€100,000

EXPERIENCE

Associate Research Scientist: Johns Hopkins University, Department of Biology, Taylor Lab

July 2014 – present

A member of the Galaxy Project (www.galaxyproject.org) leading the efforts around provisioning of the Galaxy application ecosystem across distributed compute resources in an integrated fashion. More specifically, working on harmonizing use of cloud and dedicated compute clusters to provide a seamless platform for executing bioinformatics jobs. In the process, dealing with cloud resource provisioning (AWS, OpenStack, iPlant), setup automation (Ansible), data locality and container technologies (Docker).

Research Scientist: University of Melbourne, Victorian Life Sciences Computation Initiative (VLSCI)

February 2012 – July 2014

Technical lead on the Genomics Virtual Laboratory (GVL) project. Outcomes (architectural concepts, software tools) of this project have been adopted and independently applied to two other comparable projects.

In charge of establishing computationally functional aspects of the GVL; namely, coordinating utilization of the Australian national Research Cloud built on OpenStack cloud middleware and bioinformatics applications running on top. This also involved work on automating the process of building, deploying, and configuring a comprehensive bioinformatics workbench (100+ bioinformatics tools, Galaxy application, cluster-on-the-cloud) to make use of the Research cloud.

Involved in numerous training activities teaching cloud computing and programming for the cloud concepts to bioinformaticians and software engineers alike.

Research Scientist: Ruđer Bošković Institute (RBI), Center for Computing and Informatics (CIR)

September 2011 – July 2014

Director of the CloudMan open source project (www.usecloudman.org); responsible for devising, engineering, building, and maintaining the system that powers *Galaxy on the Cloud* while extending it with novel computational models from the domain of Big Data to create a versatile computational platform for

the cloud. The effort involved coordination of objectives between three actively involved and geographically distributed groups and consolidating those into a cohesive and functional product.

Involved in formulating and writing grant proposals for FP7, Horizon 2020, and Croatian EU Structural Funds programs.

Help with annual VIS-DC conference and workshop organization as part of the MIPRO conference series.

Contributed to the outreach activities of popularizing cloud computing and bioinformatics within RBI and the region via a series of training sessions.

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 and engineered a system called Galaxy CloudMan that autonomously deploys and automatically scales cloud resources required to setup and manage the Galaxy application

Devised a reference model for deploying arbitrary applications into virtualized environments

Developed procedures for automated deployment of dozens of bioinformatics applications and required dependencies onto distributed computational infrastructures, including cloud resources

Setup a cloud computing infrastructure based on the Eucalyptus project 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 (SURAggrid), 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

Genomics Virtual Lab (GVL) – GVL project is a combination of a scalable compute infrastructure, workflow platforms and community resources for Australian genomics researchers. GVL comprises: a workflow management system based on the Galaxy framework, a bioinformatics toolkit (for command-line users) based on CloudBioLinux, and a visualisation service based on the UCSC Genome Browser, all implemented on the Australian national research cloud (NeCTAR). GVL is developing set of tutorials and exemplar workflows targeted at common high throughput genomics tasks. I was the technical lead on the project in charge of infrastructure management and workflow platform setup.

CloudBioLinux – is a set of automation scripts that enable a complete bioinformatics analysis benchmark to be installed and configured on cloud or local resources. I am one of the top contributors to this open source project. Specifically, I have contributed code that deals with configuring all the components required to install and configure Galaxy application, installation procedures for a number of domain-specific tools, core framework generalization changes, as well as documentation.

BioCloudCentral - a web application used to provision, manage, and monitor compute clusters in the cloud. I developed the application using a range of current web technologies (e.g., Django, AngularJS, Celery, nginx, gunicorn) and ensure the application is accessible as a public web service. Throughout year 2013, a public instance of this application was used to launch approximately 200 cloud clusters per month.

Galaxy ObjectStore – a Galaxy instance needs to manage terabytes of data, which has grown to be a substantial data management challenge. 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 Galaxy 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 – I worked on the 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, data scaling, infrastructure and application management, and cloud usability. Specifically, I worked with Amazon Web Services (AWS) and Eucalyptus middleware as well as accompanying tools (e.g., boto, euca2ools, RabbitMQ, SGE) to devise this project.

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. I 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) – I devised and developed a set of core grid services that enable collection and retrieval of relevant application- and resource-specific information. Together, these services enable 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 –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 (Debian/Ubuntu), OS X
<i>Programming Languages</i>	Python, JavaScript, Java
<i>Cloud/Grid Technologies</i>	boto, AWS, OpenStack, Globus, GridWay
<i>Domain-specific technologies</i>	Ansible, Galaxy, bioinformatics genomics tools
<i>Web Technologies</i>	Django, Backbone.js, jQuery, AngularJS

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)

SERVICE RECORD

Special Issue editor for Scalable Computing: Practice and Experience (SCPE): Special Issue on Distributed Computing with Applications in Bioengineering (Vol 17, No 2), 2016

Conference program committee member: ccGrid 2014, Galaxy Australasia Workshop (GAW) 2014, MIPRO 2014, MIPRO 2015, MIPRO 2016, Gateways 2016, Gateways 2017, Gateways 2018

Conference scientific program committee: Galaxy Australasia Meeting 2017 (GAMe2017)

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 Fall 2004 – Summer 2007

Ambassador for the student chapter of ACM at UAB Spring 2005 – Spring 2007

Student volunteer at Supercomputing 2005 (SC05) November 2005

President for the student chapter of ACM at UAB Spring 2004 – Spring 2005

Webmaster for the student chapter of ACM at UAB Fall 2002 – Fall 2003

Participated at the regional ACM Programming Contest in Daytona, FL

November 2003

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 – 2014</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, kiteboarding, sailing, tennis, windsurfing, hiking, snow skiing, water skiing, mountain biking