

Tak-Lon Stephen Wu

taklonwu@gmail.com

Phone: 812-606-5211

4751 12th Ave NE, APT 202, Seattle, WA, 98105

<http://html.soic.indiana.edu/~taklwu/>

Education

Indiana University , Bloomington, IN Ph.D. Candidate, Computer Science, (Part-Time after 2015 Dec.) <ul style="list-style-type: none">• Advisor: Judy Qiu• Research Committee: Filippo Menczer, Alessandro Flammini, David Leake• Major in Distributed System, minor in Database/Data Mining• Research interests in MapReduce, NoSQL (HBase), Pig / Hive, XaaS, Cloud Computing, Data Intensive Computing	Aug. 2010 – Present
Indiana University , Bloomington, IN Master's Degree in Computer Science	May, 2010
National Central University , Chung-Li, Taiwan Bachelor's Degree in Computer Science and Information Engineering	June, 2008

Work Experience

Amazon LLC , AWC EC2 Capacity, Seattle, WA <i>Software Development Engineer</i>	Dec 2015 - Present
Amazon LLC , Digital Content Experience, Seattle, WA <i>Software Development Engineer Intern</i> <ul style="list-style-type: none">• Kindle eBook sample quality research and measurement.• Build a live Java pipeline that captures, generates and identifies Kindle eBook sample quality metrics.• Fix a "Search inside" component of desktop browser reader for Kindle eBook.	May 2014 – Aug 2014
Indiana University , Community Grids Laboratory, Bloomington, IN <i>Graduate Research Assistant</i> <ul style="list-style-type: none">• Pig with Tez performance test for scientific and social network data pipeline• Pig integration with Apache Harp for iterative data analysis• Build a dynamic load balancer with using Ganglia• Build a batch-like framework salsaDPI with using Opscode Chef• Compare differences between SQL and NoSQL solutions• Build a dynamic provisioning system switches between Bare Metal and VM mode using Torque, Moab and KVM (libvirt API) for a graduate-level course with 60 students simultaneously run Hadoop, Twister MapReduce Applications	May 2010 – Dec 2015
Indiana University , Computer Science, Bloomington, IN <i>Associate Instructor</i> <ul style="list-style-type: none">• Associate Instructor for Spring 2013 C343/A594 Data Structure• Associate Instructor for Spring 2011, Fall 2012 CSCI B534 Distributed System• Associate Instructor for Fall 2010, Fall 2011 CSCI B649 Cloud Computing	Aug. 2010 – May 2012
Indiana University , Community Grids Laboratory, Bloomington, IN <i>Hourly Web developer and Graduate Research Assistant</i> <ul style="list-style-type: none">• Prepare Hadoop tutorial of VSCSE Big Data for Science 2010• Design and build an searchable web service on a Digital Academic Archive with Django and Whoosh indexing for a total of 2304 academics articles• Cloud computing research tasks with using Hadoop MapReduce• Design and Implement cloud applications (Hadoop Cap3/Blast, Amazon SQS Blast/Cap3) with using Hadoop and Amazon SQS• Generate raw data for a bioinformatics application• Coordinator of Undergraduate Research Project	Oct. 2009 – Jul.2010
Indiana University , Office of Community and School Partnerships, Bloomington, IN <i>Webmaster</i> <ul style="list-style-type: none">• Maintain website with using HTML, PHP, JavaScript, flash and SQL• Edit posters, conference banner, and brochure.	Jan. 2009 – Oct. 2009

- System Administrator of a 8 node (64 cores) Cluster running SUSE Linux
- Create internal website for sharing information

Technical Skill

- **Key skills:** Hadoop MapReduce, YARN, Pig, Hive, Tez, Spark SQL, HBase, Amazon EC2/SQS/S3
- **Distributed System:** Grids, Clouds, Microsoft Windows Azure Platform
- **Pervasive Computing:** Phidget with C#, Aduino with C++
- **Database:** SQL, Apache HBase, Apache Cassandra
- Proficient with using Java, C/C++, Python, HTML, CSS, Scheme

Publications and Technical Articles

- [1] Tak Lon Wu, Judy Qiu, Towards an Understanding of Scalable Query and Data Analysis for Social Media Data and Analysis on High-Level Dataflow Systems, in *Proceedings of the 8th Workshop on Many-Task Computing on Clouds, Grids, and Supercomputers (MTAGS) 2015*.
- [2] Tak-Lon (Stephen) Wu, Bingjing Zhang, Clayton Davis, Emilio Ferrara, Alessandro Flammini, Filippo Menczer, and Judy Qiu, Scalable Query and Analysis for Social Networks: An Integrated High-Level Dataflow System with Pig and Harp, in *Big Data in Complex and Social Networks*, My T. Thai, Hui Xiong, and W. Wu, Editors. 2015
- [3] Tak-Lon Wu, Abhilash Koppula, and Judy Qiu. 2014. Integrating Pig with Harp to support iterative applications with fast cache and customized communication. In *Proceedings of the 5th International Workshop on Data-Intensive Computing in the Clouds (DataCloud '14)*. IEEE Press, Piscataway, NJ, USA, 33-39. DOI=10.1109/DataCloud.2014.8 <http://dx.doi.org/10.1109/DataCloud.2014.8>
- [4] Tak-Lon (Stephen) Wu, [salsaDPI: a dynamic provisioning interface for IaaS cloud](#) Technical Report June 25 2013
- [5] Tak-Lon (Stephen) Wu, [MapReduce and Data Intensive Applications](#) Technical Report June 25 2013
- [6] Tak-Lon (Stephen) Wu, [An Overview of Present NoSQL Solutions and Features](#) Technical Report June 25 2013
- [7] Thilina Gunarathne, Bingjing Zhang, Tak-Lon Wu, Judy Qiu, [Scalable Parallel Computing on Clouds Using Twister4Azure Iterative MapReduce](#) in *Future Generation Computer Systems* published June 19 2012
- [8] Thilina Gunarathne, Bingjing Zhang, Tak-Lon Wu, Judy Qiu, [Portable Parallel Programming on Cloud and HPC: Scientific Applications of Twister4Azure](#), to appear in *Proceedings of Fourth IEEE International Conference on Utility and Cloud Computing (UCC 2011)*, Melbourne, Australia, December 5-8, 2011
- [9] Thilina Gunarathne, Tak-Lon Wu, Jong Youl Choi, Seung-Hee Bae, Judy Qiu, [Cloud Computing Paradigms for Pleasingly Parallel Biomedical Applications](#), accepted to the *Journal of Concurrency & Computation: Practice & Experience. Special Issues on Emerging Computational Methods for the Life Sciences*, DOI: 10.1002/cpe.1780, April 2011.
- [10] Thilina Gunarathne, Tak-Lon Wu, Judy Qiu, Geoffrey Fox, [MapReduce in the Clouds for Science](#). *Proceedings of the 2nd IEEE international Conference on Cloud Computing Technology and Science*, Indianapolis, Indiana, Nov. 30-Dec 3, 2010.
- [11] Bingjing Zhang, Yang Ruan, Tak-Lon Wu, Judy Qiu, Adam Hughes, Geoffrey Fox, [Applying Twister to Scientific Applications](#). *Proceedings of the 2nd IEEE international Conference on Cloud Computing Technology and Science*, Indianapolis, Indiana, Nov. 30-Dec 3, 2010.
- [12] Judy Qiu, Thilina Gunarathne, Jaliya Ekanayake, Jong Youl Choi, Seung-Hee Bae, Hui Li, Bingjing Zhang, Yang Ryan, Saliya Ekanayake, Tak-Lon Wu, Scott Beason, Adam Hughes, Geoffrey Fox, [Hybrid Cloud and Cluster Computing Paradigms for Life Science Applications](#) Technical Report April 17 2010 submitted to the *11th Annual Bioinformatics Open Source Conference BOSC 2010*
- [13] Thilina Gunarathne, Tak-Lon Wu, Judy Qiu, and Geoffrey Fox, [Cloud Computing Paradigms for Pleasingly Parallel Biomedical Applications](#) March 21 2010. *Proceedings of Emerging Computational Methods for the Life Sciences Workshop of ACM HPDC 2010 conference*, Chicago, Illinois, June 20-25, 2010.
- [14] Tak-Lon Wu, Shakeela Noohir Baasha and Sonali Surendra Karwa, [A Survey of Open-Source Cloud Infrastructure using FutureGrid Testbed](#). Poster at the *2nd IEEE international Conference on Cloud Computing Technology and Science*, Indianapolis, Indiana, Nov. 30-Dec 3, 2010.
- [15] Judy Qiu, Jaliya Ekanayake, Thilina Gunarathne, Jong Youl Choi, Seung-Hee Bae, Yang Ruan, Saliya

Presentation and Demo

- [Reproducible Environment for Scientific Applications \(Lab session\)](#) Presentation at Science Cloud [Summer School](#) organized by [VSCSE](#) July 31 2012
- [Hadoop Tutorial](#), [Big data for Science workshop](#), NCSA Summer School, July, 2010.

Activites and Awards

- Student Mentor, Indiana University HBCU STEM Summer Scholar Institute, Summer 2010, 2011
- Student Volunteer, the 2nd IEEE international Conference on Cloud Computing Technology and Science, Indianapolis, Dec. 2010
- Macau SAR government CTABE Postgraduate Doctoral Scholarship Award 2010-2013 (15 people awarded)