

Ketan C Maheshwari

ketancmaheshwari@gmail.com
ketan@mcs.anl.gov
+1 630 252 0908

480 74th street,
Downers Grove,
IL 60516 USA

OBJECTIVE To work on challenging research pertaining to large-scale computer systems.

WORK EXPERIENCE ◇ **Postdoctoral Researcher** Distributed Systems Lab, Argonne National Laboratory, February 2011 - March 2012 and March 2013 - current.

In this position, I work in a dual role of supporting large-scale applications on HPC infrastructures as a user-engagement catalyst and a researcher in the area of distributed HPC infrastructures.

Main activities are writing and running parallel applications on clouds, XSEDE, Cray XE6 and Blues supercomputers at University of Chicago. The applications include massive protein docking, weather and soil modeling, earthquake simulations, and power grid modelling. The research is funded by NSF, DOE and NIH.

◇ **Postdoctoral Research Associate** CS Department, Cornell University, April 2012-March 2013.

The research conducted in this position was in the area of Smart Power Grid Design supported by ARPA-e GENI program. The application involved modeling the computations required in order to perform a state estimation in the smart grid and a provision of time critical reactive control thereof. I lead the effort to build a first software infrastructure to collect and process remote data over wide-area network and store into AWS cloud.

◇ **ICT Developer(L5)** Informatics Institute(IvI), University of Amsterdam(UvA), June 2006 to May 2007.

Worked on Reliable File Transfer Service (RFTS) of Globus Toolkit 4.0.x to achieve transfer of large number of big files between Academisch Medical Centrum(AMC) and Storage Resource Broker(SRB) at SARA, Amsterdam. Integrated the RFTS client into the Virtual Resource Browser developed within the vl-e project.

◇ **MCA Intern/Programmer**, Atharwa Management and Strategic Systems (now, Arth Infosoft Private ltd.(AIPL)), Gandhinagar, Gujarat, India(Spring 2001)

SKILLS ◇ Competent in python, bash, awk, L^AT_EX, vim

◇ Working knowledge of sed, gnuplot, MPI, openMP, OpenStack

◇ gnu/Linux, ranked top 200 of Unix/Linux StackExchange community

◇ Fluent spoken/written English, basic French

EDUCATION ◇ **University of Nice at Sophia Antipolis**, France.

PhD with **Highest Honors**, January 2011.

The research included workflow expression and enactment on large-scale distributed computing infrastructures for data intensive applications under the guidance of Dr. Johan Montagnat. The projects were related to the field of biomedical imaging. First was the 'Cardiac Image Processing Application' involving processing 4D (3D + time) images of cardiac motion cycle and extracting quantitative information about myocardial movement. Second was the 'Drug Discovery Application' that dealt with in-silico docking of protein molecules by filtering a million potential candidates based on their 'docking energy'.

◇ **University of Amsterdam**, Amsterdam, The Netherlands.

Research M.Sc.(NVAO-accredited) in Grid Computing, August 2007.

Masters Thesis Work

To achieve integration and interoperability among several Workflow Management Systems(WfMS) and Problem Solving Environments(PSE) in order to realize a Medical Image Analysis Support System at the University of Amsterdam Academisch Medical Centrum, Amsterdam.

CourseWork

Distributed Stochastic Simulation, Introduction to Grid Computing, Scientific Visualization and Virtual Reality, Grid Hardware Infrastructures, Distributed Programming Methods, Profile Project – Grid Computing (Task Farming fMRI jobs on DAS-2 Cluster using Nimrod-G and Globus), Scientific Computing on the Grid, Computational Finance, Theory and Application of Multi-Threading, Concurrent Systems.

- ◇ **Gujarat University**, Ahmedabad, India.

Masters in Computer Applications(MCA), September 2001.

Major Courses: Fundamentals of Programming, Database Systems, Operating Systems, Data and File Structures, Object Oriented Analysis and Design, Discrete Mathematical Structures, Numerical Methods, Software Engineering, Client Server Technologies, Networking Technologies, Computer Architecture, Financial Accounting, Managerial Economics.

- ◇ **Gujarat University**, Ahmedabad, India.

Bachelors in Science, June 1998.

SCHOLARSHIP DELTA (Dutch Education: Learning at Top Level Abroad) Scholarship for the academic year 2006/07.

GRANTS &
SERVICE

- ◇ PI on \$ 2500 Amazon AWS grant
- ◇ CoPI on 200K SU XSEDE compute grant
- ◇ Designated organizer, SRMPDS'14
- ◇ Invited talk at FERC at Washington, DC and LANS at Argonne
- ◇ Session chair on Industrial Forum at Super Computing 13 conference, ICPP 14 conference
- ◇ Reviewer on SC'11 and '14 (3), IEEE TCC (1), ACM Computing Surveys (1), ICPADS'12 (1), Cluster'13 (3), CLOSER'14 (1), ICPP'14 (2), ICCCN'13 (3), ACM TOPC (1), JORS (1), SRMPDS'14 (3), Cluster journal (1)
- ◇ Professional member of IEEE and ACM

Publications & Theses

- [1] Thoshitha Gamage, David Anderson, David Bakken, Kenneth Birman, Anjan Bose, Carl Hauser, *Maheshwari, Ketan*, and Robbert van Renesse. Mission-critical cloud computing for critical infrastructures. In David Bakken, editor, *Smart Grids: Clouds, Communications, Open Source, and Automation*, pages 1–19. CRC Press, 2014.
- [2] L.B. Costa, H. Yang, E. Vairavanathan, A. Barros, *Maheshwari, K.*, G. Fedak, D. Katz, M. Wilde, M. Ripeanu, and S. Al-Kiswany. The case for workflow-aware storage:an opportunity study. *Journal of Grid Computing*, pages 1–19, 2014.
- [3] Daniel S. Katz, Sou-Cheng T. Choi, Hilmar Lapp, *Ketan Maheshwari*, Frank Löffler, Matthew Turk, Marcus D. Hanwell, Nancy Wilkins-Diehr, James Hetherington, James Howison, Shel Swenson, Gabrielle Allen, Anne C. Elster, G. Bruce Berriman, and Colin C. Venters. Summary of the first workshop on sustainable software for science: Practice and experiences (wssspe1). *Journal of Open Research Software*, 2(1):e6, 2014.
- [4] *Ketan Maheshwari*, Justin Wozniak, Hao yang, Daniel Katz, Matei Ripeanu, Victor Zavala, and Michael Wilde. Evaluating Storage Systems for Scientific Data in the Cloud. In *ScienceCloud at HPDC*, Vancouver, Ca, May 2014. IEEE/ACM. *Best paper award*.

- [5] *Ketan Maheshwari*, Eun-Sung Jung, Jiayuan Meng, Venkatram Vishwanath, and Raj Kettimuthu. Improving Multisite Workflow Performance using Model-based Scheduling. In *International Conference on Parallel Programming*, Minnesota, MN, USA, September 2014. IEEE XPLORE.
- [6] *Ketan Maheshwari*, Alex Rodriguez, David Kelly, Ravi Madduri, Justin M. Wozniak, Michael Wilde, and Ian Foster. Extending the Galaxy portal with parallel and distributed execution capability. In *The 4th International Workshop on Data-Intensive Computing in the Clouds (DataCloud at SC13)*, Denver, Co, USA, November 2013. IEEE/ACM.
- [7] Eun-Sung Jung, *Ketan Maheshwari*, and Rajkumar Kettimuthu. Pipelining/overlapping data transfer for distributed data-intensive job execution. In *ICPP*, pages 791–797, 2013.
- [8] *Ketan Maheshwari*, David Kelly, Scott J. Krieder, Justin M. Wozniak, Daniel S. Katz, Mei Zhi-Gang, and Mainak Mookherjee. Reusability in science: From initial user engagement to dissemination of results. volume abs/1309.1813, 2013.
- [9] *Ketan Maheshwari*, Alex Rodriguez, David Kelly, Ravi Madduri, Justin Wozniak, Michael Wilde, and Ian Foster. Enabling Multi-task computation on Galaxy-based Gateways using Swift. In *Science Gateway Institute Workshop (SGIW at Cluster'13)*, Indianapolis, IN, USA, September 2013. IEEE Computer Society.
- [10] Justin M. Wozniak, Timothy G. Armstrong, *Ketan Maheshwari*, Ewing L. Lusk, Daniel S. Katz, Michael Wilde, and Ian T. Foster. Turbine: A distributed-memory dataflow engine for high performance many-task applications. *Fundamenta Informaticae*, 128(3):337–366, 2013.
- [11] Marc Parisien, Xiaoyun Wang, George Perdrizet, Corissa Lamphear, Carol A Fierke, *Ketan Maheshwari*, Michael J Wilde, Tobin R Sosnick, and Tao Pan. Discovering rna-protein interactome by using chemical context profiling of the rna-protein interface. *Cell Reports*, May 2013.
- [12] *Ketan Maheshwari*, Ken Birman, Justin Wozniak, and Devin Van Zandt. Evaluating Cloud Computing Techniques for Smart Power Grid Design Using Parallel Scripting. In *Cluster Cloud and Grid Computing (CCGrid)*, TUDelft, Delft, Netherlands, May 2013. IEEE/ACM.
- [13] *Ketan Maheshwari*, Marcus Lim, Lydia Wang, Ken Birman, and Robbert van Renesse. Toward a reliable, secure and fault tolerant smart grid state estimation in the cloud. In *Innovative Smart Grid Technologies*, Washington DC, USA, February 2013. IEEE-PES.
- [14] *Ketan Maheshwari*, Allan Espinosa, Daniel S. Katz, M. Wilde, Z. Zhao, Ian Foster, Scott Callaghan, and Philip Maechling. Job and Data Clustering for Aggregate Use of Multiple Production Cyberinfrastructures. In *Data Intensive Distributed Computing*, Delft, the Netherlands, June 2012. ACM.
- [15] *Ketan Maheshwari*, Justin Wozniak, Mihael Hategan, Allan Espinosa, Daniel S. Katz, and M. Wilde. Flexible Cloud Computing through Swift Coasters. In *Proceedings of Cloud Computing and its Applications, short paper*, Chicago, April 2011. Open Cloud Consortium.
- [16] Justin Wozniak, *Ketan Maheshwari*, Mihael Hategan, David Kelly, Jon Monette, Daniel S. Katz, M. Wilde, Duncan Roweth, and David Strenski. Swift - a parallel scripting language for petascale many-task applications. In *Proceedings of the 54th Cray User Group Conference(CUG 2012)*, Stuttgart, Germany, May 2012. Cray inc., Cray User Group inc.
- [17] Justin Wozniak, Timothy Armstrong, *Ketan Maheshwari*, Ewing Lusk, Daniel S. Katz, M. Wilde, and Ian Foster. Turbine: A distributed-memory dataflow engine for extreme-scale many-task applications. In *Scalable Workflow Enactment Engines and Technologies(SWEET'12)*, Scottsdale, AZ, USA, May 2012. ACM.

- [18] Mihael Hategan, Justin Wozniak, and *Ketan Maheshwari*. Coasters: Uniform Resource Provisioning and Access for Clouds and Grids. In *Proceedings of the Utility and Cloud Computing(UCC'11)*, pages 114–121, Melbourne, Australia, December 2011. IEEE, IEEE Computer Society.
- [19] Timothy Armstrong, Justin Wozniak, M. Wilde, *Ketan Maheshwari*, Daniel S. Katz, Matei Ripeanu, Ewing Lusk, and Ian Foster. ExM: High level dataflow programming for extreme scale systems. In *4th USENIX Workshop on Hot Topics in Parallelism(HoTPar)*, poster, Berkeley, CA, June 2012. ACM, USENIX Association.
- [20] *Ketan Maheshwari* and Johan Montagnat. Scientific workflows development using both visual-programming and scripted representations. In *International Workshop on Scientific Workflows(SWF'10)*. IEEE, July 2010.
- [21] Johan Montagnat, Tristan Glatard, Damien Reimert, *Ketan Maheshwari*, Eddy Caron, and Frédéric Desprez. Workflow-based comparison of two Distributed Computing Infrastructures. In *5th Workshop on Workflows in Support of Large-Scale Science(WORKS'10)*, pages 1–10, New Orleans, LA, USA, November 2010.
- [22] Allan Espinosa, Daniel S. Katz, M. Wilde, *Ketan Maheshwari*, Ian Foster, Scott Callaghan, and Philip Maechling. Data-intensive CyberShake computations on an opportunistic cyberinfrastructure. In *Proceedings of the 2011 TeraGrid Conference: Extreme Digital Discovery(TG'11)*, TG '11, pages 14:1–14:2, Salt Lake City, Utah, USA, 2011. ACM, ACM.
- [23] *Ketan Maheshwari*, Tristan Glatard, Joël Schaerer, Bertrand Delhay, Sorina Camarasu, Patrick Clarysse, and Johan Montagnat. Towards Production-level Cardiac Image Analysis with Grids. In *HealthGrid'09*, June 2009.
- [24] *Ketan Maheshwari*, Paolo Missier, Carole Goble, and Johan Montagnat. Medical Image Processing Workflow Support on the EGEE Grid with Taverna. In *Intl Symposium on Computer Based Medical Systems(CBMS'09)*. IEEE, August 2009.
- [25] Johan Montagnat, Benjamin Isnard, Tristan Glatard, *Ketan Maheshwari*, and Mireille Blay-Fornarino. A data-driven workflow language for grids based on array programming principles. In *Workshop on Workflows in Support of Large-Scale Science(WORKS'09)*, pages 1–10, November 2009.
- [26] *Ketan Maheshwari*, Silvia Olabarriaga, Charl Botha, Jeroen Snel, Johan Alkemade, and Adam Belloum. Problem Solving Environment for Medical Image Analysis. In *Computer Based Medical Systems, 2007 (CBMS-07)*, Maribor, June 2007. IEEE.
- [27] Silvia D. Olabarriaga, Piter T. de Boer, *Ketan Maheshwari*, Adam Belloum, Jeroen G. Snel, Aart J. Nederveen, and Maurice Bouwhuis. Virtual lab for fmri: Bridging the usability gap. In *Proceedings of the Second IEEE International Conference on e-Science and Grid Computing, E-SCIENCE '06*, pages 53–, Washington, DC, USA, 2006. IEEE Computer Society.
- [28] *Ketan Maheshwari*. *Data-intensive Scientific Workflows: Representation of Parallelism and Enactments on Distributed Systems*. PhD thesis, Ecole Polytech Universitaire, Sophia Antipolis, Nice, January 2011.
- [29] *Ketan Maheshwari*. Problem Solving Environment for Medical Image Analysis Application Development. Master thesis, University of Amsterdam, Amsterdam, August 2007.