

USC
Information Sciences Institute
4676 Admiralty Way, Suite 1001
Marina Del Rey, CA 90292

(310) 448-8408
(310) 823-6714
deelman@isi.edu
www.isi.edu/~deelman

Ewa Deelman

Research Interests

Managing large-scale workflows and data in distributed environments. Design and exploration of collaborative scientific environments based on Grid technologies. Performance optimization in parallel and distributed environments.

Appointments

- 2003-present **Research Assistant Professor**
Computer Science Department
University of Southern California, Los Angeles, CA
- 2007-present **Project Leader, head of the Collaborative Computing Group, ISI**
- 2002-2007 **Research Team Leader**
- 2000-2002 **Computer Scientist**
Information Sciences Institute
University of Southern California, Marina Del Rey, CA

Scientific applications need to make use of a number of heterogeneous and often distributed resources that include both data and computations. My research focuses on developing tools and techniques that enable easy, efficient, and reliable execution of complex applications in distributed environments. Particular emphasis is given to managing applications specified as computational workflows. My research encompasses a number of areas including task scheduling, performance optimization, reliability techniques, resource provisioning, and workflow monitoring and troubleshooting. One of the tools developed in my group is the Pegasus Workflow Management System that is being used day-to-day in a number of scientific applications including astronomy, biology, earthquake science, gravitational-wave physics and others. The applications make use of their own computing clusters, national cyberinfrastructure, and Cloud computing resources. My research interests also include data and metadata management in scientific applications.

- 1997-2000 **Senior Software Developer**
UCLA Computer Science Dept.
Los Angeles, CA

Conducted research in parallel simulation of message-passing programs and performance modeling. *Participated in the POEMS project, which aims to design a system for Performance Oriented End-to-End Modeling of Large Heterogeneous Adaptive Parallel/Distributed Computer/Communication Systems. Major development was geared towards efficient and accurate simulation of message-passing programs (such as those using MPI) on high performance systems (such as the IBM SP). New SMP-cluster architectures were modeled. Compiler optimizations which improved the efficiency and scalability of the simulator were investigated. The POEMS project involves a collaboration between several universities: Purdue, Rice, UCLA, U.T. Austin, U.T. El Paso, U. of Wisconsin and LANL.*

Guest lectured in parallel computation courses.
Guided the research of graduate students.

Education

1997 **Doctor of Philosophy in Computer Science.**

Rensselaer Polytechnic Institute Troy, NY

"Performance Optimization of Parallel Discrete Event Simulation of Spatially Explicit Problems"

Developed an object oriented Parallel Discrete Event Simulation System that uses the optimistic approach to event processing. The system was designed to run on an MIMD machine and tested on an IBM SP. In order to achieve good performance, new synchronization techniques were formulated. Developed a new algorithm to continuously monitor the progress of all simulation processes. Designed a new approach for rolling back computation in spatially explicit problems. This approach resulted in a speedup close to linear over a single processor run. Load balancing techniques were also investigated.

Thesis Advisor: Prof. Boleslaw Szymanski

1991 **Master of Science in Computer Science**

State University of New York, New Paltz, NY

1987 **Bachelor of Arts in Mathematics**

Wells College Aurora, NY

Professional Activities

- Associate Editor responsible for Grid Computing for the Scientific Programming Journal
- Steering Committee Member, e-Science All Hands Meeting Foundation 2008-
- Chair, Global Grid Forum Workflow Management Research Group 2003-
- Chair, Workshop on Workflows in Support of Large-Scale Science, in conjunction with SC 2008
- Chair, Workshop on Workflows in Support of Large-Scale Science, in conjunction with HPDC 2006, and 2007 June 2006 & 2007
- Co-Guest Editor, Scientific Programming Journal, Special issue on Dynamic Computational Workflows: Discovery, Optimization, and Scheduling, 2007
- Co- Guest Editor Journal of Grid Computing, Special Issue dedicated to Workflow Management in Grids, 2006
- Co-Editor, Workflows in e-Science book, Springer 2007
- Workshop Chair Applications and Middleware Grid Workshop, September 2005
- Lecturer, Global Grid Forum Summer School, July 2004
- Organizing Committee Member, GGF Workshop on Workflows in Grids, 2004.
- Co-Editor of the Special Issue of the Scientific Programming Journal devoted to Grid Computing, Volume 10, Number 2, 2002.
- Workshop Chair, Applications Grid Workshop, September 2003, 2005, 2007
- Panel member at the International Conference on Dependable Systems and Network panel on "Dependability and the Grid", June 2002.
- Technical Board member of the Gridlab project, an EU funded project.
- PhD Committee member, Rensselaer Polytechnic Institute, 2002.
- External Examiner for Masters' Candidate, McGill University, 2002.
- General Co-Chair, of the 15th Workshop on Parallel and Distributed Simulation, PADS2001, Lake Arrowhead, CA, 2001.
- Organizer So-Cal Seminar on Parallel Computing Systems, January 22, 1999.
- Organizing Committee Chair, Third Workshop on Languages, Compilers, and Run-Time Systems for Scalable Computers held at Rensselaer Polytechnic Institute, May 1995.

- Pipelink Project Member, Lectured high-school students about research performed in Computer Science. The goal of this project is to interest and support young women in Computer Science, 1995.

Program Committee Member

- 4th Int'l Workshop on Workflow systems in e-Science (WSES09), Shanghai, China
- High Performance Distributed Computing (HPDC 2009), Garching, Germany, June 2009
- SC 2008, Austin Texas, November 2008
- Grid 2008, Tsukuba, Japan, September 2008
- High Performance Distributed Computing (HPDC 2008), Boston, MA, June 2008
- e-Science 2007, Bangalore, India, December 2007
- NSF/Mellon Workshop on Scientific and Scholarly Workflow. Oct 3-5, 2007, Baltimore, MD
- CoreGRID Workshop on Grid Middleware, Dresden, June 18-19, 2007
- High Performance Distributed Computing (HPDC 2007), Monterey Bay California, June 27-29
- IEEE International Conference on Grid Computing (Grid 2007) (Vice Program chair for Scheduling, Resource Management and Runtime Environments), Grid 2006, Grid 2005
- Second International Conference on Grid computing, high-performance and Distributed Applications" (GADA'07)
- Workflow Systems in e-Science 2007, International Conference on Computational Science (ICCS2007)
- The 13th International Conference on Parallel and Distributed Systems, Hsintu, Taiwan, December 2007
- 3rd International Workshop on Grid and Peer-to-Peer based Workflows (GPWW), Brisbane, Australia, 2007
- IEEE International Symposium on Cluster Computing and the Grid, CCGrid 2007, 2006, 2005
- CoreGRID Workshop on Grid Middleware 2006
- First International Multiconference on Computer Science and Information Systems
- The 006 International Conference on High Performance Computing and Communications, Grid and cluster computing topic. 2006.
- International Workshop on Workflow Systems in Grid Environments (WSGE06), October 21-23, 2006, Changsha, China.
- Challenges of Large Applications in Distributed Environments Workshop, CLADE 2009, 2008, 2007, 2006, 2005, 2004, 2003
- Fourth International Symposium on Parallel and Processing and Applications (ISPA'06), (Software and Applications Track), Sorrento, Italy, December 2006.
- 2nd International Workshop on Grid and Peer-to-peer based Workflows (GPWW 2006) ,Vienna, Austria on September, 2006, and 2005
- Scientific workflow management in e-Science, in conjunction with the International Conference on Computational Science (ICCS2006), May 2006.
- Statistical and Scientific Database Management (SSDBM) , 2006, 2005
- ICIW 2006 Web Service-based Systems and Applications track, February 2006
- International Workshop on Scientific Instruments and Sensors on the Grid, December 2005.
- Semantic Infrastructure for Grid Applications at CCGrid 2005
- European Grid Conference 2005
- Workshop on Web and Grid Services for Scientific Data Analysis (WAGSSDA), to be held in conjunction with the International Conference on Parallel Processing (ICPP-2005)
- IEEE International Conference on Services Computing (SCC 2004)
- International Conference on Semantics for a Networked World 2004
- 2nd European Across Grids Conference, 2004
- SC 2003, November 2003
- Heterogeneous Computing Workshop, 2003, 2004

Reviewer

Reviewer for journals such as the Journal of Grid Computing, Journal of Parallel and Distributed Computing, IEEE Intelligent Systems, Informatica, IEEE Transactions on Parallel and Distributed Systems, Concurrency and Computation: Practice and Experience.

Reviewer for DOE, NSF, and EU funding agencies.

Invited Talks

- Cracow Grid Workshop, Cracow, Poland, October 2008. (keynote)
- CoreGrid Workshop, Crete, Greece, April 2008.
- AAAI Spring Symposium on Semantic Scientific Knowledge Integration, Stanford University, March 2008.
- The International Workshop on Performance Analysis and Optimization of High-End Computing Systems, in conjunction with SC'07, November 2007
- International Conference on Parallel Processing and Applied Mathematics. PPAM, Gdansk, Poland, September 2007
- Computer Science Department Colloquium, Southampton University, UK April 2007
- Computer Science Department Colloquium, King's College, London, UK, April 2007
- Colloquium, University of Vienna, Austria, June 2007
- Workshop on Scientific Workflows and Business workflow standards in e-Science, in Conjunction with e-Science 2006, Amsterdam, December 2006
- Workflow Optimization in Distributed Environments, Edinburgh, October 2006
- Manchester University, School of Computer Science Seminar, October 2006
- Louisiana State University, Computer Science Department Seminar, September 2006
- Cyberinfrastructure for Ocean Observations Workshop, September 2006
- Partnerships in Innovation: Serving a Networked Nation conference, National Archives and Records Administration, November 2004
- Astronomical Data Analysis Software & Systems (ADASS), October 2004
- SC4DEVO: Service Composition for Data Exploration in the Virtual Observatory, July 2004
- SSDBM 04 panel on Data Management on the Grid, June 2004
- Ground System Architectures Workshop (GSAW 2004), April 2004
- SDSC Computational Science Seminar Series (CSSS), January 2004
- Rensselaer Polytechnic Institute, *GriPhyN: Data-Intensive Science in Grid Environments*, Troy, NY, January 2002.
- VIRGO Laboratory, GriPhyN (Grid Physics Network) and LIGO: Data-Intensive Science in Grid Environments, Cascina, Italy, September 2001.
- International Conference on Parallel Processing and Applied Mathematics. GriPhyN (Grid Physics Network): Enabling Data Intensive Science in Grid Environments. Naleczow, Poland, September 2001.
- Poznan Supercomputing and Networking Center, GriPhyN: Data-Intensive Science in Grid Environments, Poznan, Poland, September 2001.
- Mardi Gras Conference, GriPhyN (Grid Physics Network): Building a Data Grid Infrastructure for Experimental Physics, Baton Rouge, LA, February 2001.
- University of Texas at El Paso, MPI-Sim, an MPI Simulator in POEMS (Performance Oriented End-to-End Modeling System), El Paso, TX, August 2000.
- USC/ISI, POEMS: Performance Oriented End-to-End Modeling System, Marina Del Rey, CA, July 2000.
- Aerospace Corporation, Optimization in Parallel Discrete Event Simulation, Los Angeles, CA, July 1997.
- Institute of Fundamental Technological Research, Polish Academy of Sciences, *Performance Optimization in Parallel Discrete Event Simulation*, Warsaw, Poland, June 1997.

Teaching Activities

- USC, Computer Science Department, CS 599 Introduction to Grid Computing (Graduate level), Fall 2007
- UCLA, Department of Computer Science, Guest Lecturer
 - CS 133, Parallel and Distributed Programming, an undergraduate level course in parallel programming. It covers architectural and language models of parallelism. The emphasis is learning to write parallel programs. Winter Term 1998, and 1999.
 - CS 233A, Parallel Programming, graduate course in parallel programming that covers in more depth the topics of CS 133. Emphasis is on examining state of the art research and individual research projects. Winter Term 1998, and 1999.
 - CS 239, Parallel Simulation, a follow-up course to CS 233A. Spring Term 1998.

Fellowships, Honors

- Best Paper award, e-Science 2006, Amsterdam, Netherlands
- Best Paper award, 15th Workshop on Parallel and Distributed Simulation, Lake Arrowhead, CA, USA; 15-18 May 2001.
- General Electric Foundation Fellowship, Rensselaer Polytechnic Institute, 1993-1994.
- Special Distinction in the field of Mathematics, Wells College, 1987.

Media Citations

- Image of the week - Earth-quaking science in Hollywood, international Science Grid this Week, iSGTW online, January 2008, <http://www.isgtw.org/?pid=1000848>
- Feature - Montage a rising star in grid-enabled sky mosaics, international Science Grid this Week, iSGTW online, December 2007, <http://www.isgtw.org/?pid=1000731>
- Feature - Pegasus invites new communities to saddle up, international Science Grid this Week, iSGTW online, September 2007, <http://www.isgtw.org/?pid=1000664>
- ISI Leads \$13.8 Million E-Science Effort to Tame Terabyte Torrents, ISI News, <http://www.isi.edu/news/news.php?story=165>
- Rensselaer Supercomputers Battle Lyme Disease, Rensselaer Polytechnic Institute Review Vol. 17 No. 18, June 14, 1996

Grants

- Co-PI, Center for Genomic Studies of Mental Disorders, National Institutes of Health, 10/01/08 - 9/30/2013, \$2,250,000

Development of the Cyberinfrastructure and workflow technologies for the coordinating center.
- Sr. Personnel, Genomic Psychiatry Cohort, National Institutes of Health, 10/2008-9/2013, \$300,000

Development of the Cyberinfrastructure for a psychiatric study.
- Co-PI, NHGRI EpiGenVar Coordinating Center, National Institutes of Health, 7/1/2008-6/30/2012, \$500,000

Development of the Cyberinfrastructure and workflow technologies for the coordinating center.
- PI, Supporting Ocean Modeling With Workflow Technologies, Jet Propulsion Laboratory, 6/1/2008-5/31/2009, \$70,000

Development of workflows for and ocean forecasting system.

- PI, SDCI NMI Improvement: Pegasus: From Concept to Execution- -Mapping Scientific Workflows onto the National Cyberinfrastructure, National Science Foundation, 09/01/2007 - 08/31/2010, \$1,700,000

Hardening of the Pegasus workflow software that manages execution of complex scientific workflows on distributed resources.

- PI, Designing Scientific Software one Workflow at a Time, National Science Foundation, 10/1/2007-9/30/2010, \$313,556.

Exploring the use of the workflow paradigm in the context of large-scale software design.

- Sr. Personnel, Cyberinfrastructure in Support of Research: A New Imperative, National Science Foundation, 10/01/2004 - 09/30/2008, \$250,000

This project, lead by NCSA developed technologies necessary to support a wide range of NSF applications.

- Co-PI, Scalable Cross-Organization Threat and Event Discovery through Grid Workflows, Air Force Research Lab, 10/1/2006 - 9/30/2010, \$2,104,000

This project, lead by ISI develops new techniques to optimize workflows representing large-scale data mining applications and to provide seamless access to distributed data sources.

- Co-PI, CSR-AES: Collaborative Research: Intelligent Optimization of Parallel and Distributed Applications, National Science Foundation, 08/01/2006 – 07/31/2009, \$215,452.

This project, lead by ISI develops new techniques to optimize workflows in distributed Systems.

- Sr. Personnel, Enabling Earthquake System Science Through Petascale Calculations (PetaShake), National Science Foundation, 10/01/2007-9/30/2009, \$200, 000.

Providing a cyberinfrastructure and workflow technologies for earthquake science applications.

- PI, Pegasus: Supporting LIGO Workflows on the Open Science Grid, National Science Foundation, 11/1/2006 – 10/31/2007, \$102,000

This project, lead by ISI will enable efficient execution of gravitational-wave physics applications on large grid deployments in conjunction with the efficient data management.

- Co-PI, Neuroscience: Collaborative Research (CRCNS): Assembling Visible Neurons for Simulations: Merging of High-throughput 3D Microscopies with Advanced Computational Tools, National Institute of Health (NIH), 10/01/2002 – 09/30/2006, \$45,000

This project, lead by SDSC, developed workflow software necessary to support complex tomographic image reconstruction on distributed resources.

- PI, Towards Cognitive Grids: Knowledge-Rich Services for Autonomous Workflow refinement and Robust Execution, National Science Foundation, 12/15/2004 – 11/30/2006, \$200,000

This project examined the use of semantic technologies to describe resources and services on the grid.

- Co-PI, WSW-06: 2006 Workshop on Challenges of Scientific Workflows, National Science Foundation, 05/01/2006 to 10/31/2007, \$44,000

Publications

Book Editor

- Workflows for e-Science: Scientific Workflows for Grids, Ian J. Taylor, Ewa Deelman, Dennis B. Gannon, Matthew Shields (Editors), Springer, January 2007

Journals

- Kevin Lee, Norman W. Paton, Rizos Sakellariou, Ewa Deelman, Alvaro A. A. Fernandes, Gaurang Mehta, Adaptive Workflow Processing and Execution in Pegasus, *Concurrency Practice and Experience*, to appear 2009
- Gurmeet Singh, Carl Kesselman, and Ewa Deelman, An End-to-End Framework for Provisioning-Based Resource and Application Management, *IEEE Systems Journal*, to appear 2009
- Ewa Deelman, Dennis Gannon, Mathew Shields, Ian Taylor, Workflows and e-Science: An overview of workflow system features and capabilities, *Future Generations of Computer Systems*, July 2008.
- Simon Miles, Paul Groth, Ewa Deelman, Karan Vahi, Gaurang Mehta, and Luc Moreau. Provenance: The bridge between experiments and data. *Computing in Science and Engineering*, 2008.
- Yolanda Gil, Ewa Deelman, Mark Ellisman, Thomas Fahringer, Geoffrey Fox, Dennis Gannon, Carole Goble, Miron Livny, Luc Moreau, Jim Myers, Examining the Challenges of Scientific Workflows, *IEEE Computer*, *IEEE Computer*, vol. 40, pp. 24-32, 2007.
- Gurmeet Singh, Karan Vahi, Arun Ramakrishnan, Gaurang Mehta, Ewa Deelman, Henan Zhao, Rizos Sakellariou, Kent Blackburn, Duncan Brown, Stephen Fairhurst, David Meyers, G. Bruce Berriman, John Good, Daniel S. Katz, Optimizing Workflow Data Footprint, *Special issue of the Scientific Programming Journal dedicated to Dynamic Computational Workflows: Discovery, Optimisation and Scheduling*, 2007
- Jihie Kim, Ewa Deelman, Yolanda Gil, Gaurang Mehta, Varun Ratnakar. Provenance Trails in the Wings/Pegasus Workflow System, *Concurrency and Computation: Practice and Experience, Special Issue on the First Provenance Challenge*, 2007.
- Joseph C. Jacob, Daniel S. Katz, G. Bruce Berriman, John Good, Anastasia C. Laity, Ewa Deelman, Carl Kesselman, Gurmeet Singh, Mei-Hui Su, Thomas A. Prince, Roy Williams, Montage: a grid portal and software toolkit for science-grade astronomical image mosaicking, *IJCSE*, 2006
- Ewa Deelman, Tevfik Kosar, Carl Kesselman and Miron Livny, What Makes Workflows Work in an Opportunistic Environment? *Concurrency and Computation: Practice and Experience*, Volume 18, Issue 10, Pages 1187 – 1199. 2005
- Gurmeet Singh, Carl Kesselman, Ewa Deelman, Optimizing Grid-Based Workflow Execution, *Journal of Grid Computing*, Vol. 3, No. 3-4. (September 2005), pp. 201-219.
- Ewa Deelman, Gurmeet Singh, Mei-Hui Su, James Blythe, Yolanda Gil, Carl Kesselman, Gaurang Mehta, Karan Vahi, G. Bruce Berriman, John Good, Anastasia Laity, Joseph C. Jacob, Daniel S. Katz, Pegasus: a Framework for Mapping Complex Scientific Workflows onto Distributed Systems, *Scientific Programming Journal*, Volume 13, Number 3, 2005
- P. Maechling, H. Chalupsky, M. Dougherty, E. Deelman, Y. Gil, S. Gullapalli, V. Gupta, C. Kesselman, J. Kim, G. Mehta, B. Mendenhall, T. Russ, G. Singh, M. Spraragen, G. Staples, K. Vahi, Simplifying Construction of Complex Workflows for Non-Expert Users of the Southern California Earthquake Center Community Modeling Environment, *SIGMOD Record*, Volume 34 Number 3, September 2005
- Jim Blythe, Ewa Deelman, and Yolanda Gil, Automatically Composed Workflows for Grid Environments Described with Varying Levels of Detail, *IEEE Intelligent Systems*, 19(4): 16-23 (2004)

- Y. Gil, E. Deelman, J. Blythe, C. Kesselman, H. Tangmunarunkit. Artificial intelligence and grids: workflow planning and beyond, *IEEE Intelligent Systems*, Volume: 19 , Issue: 1 , Jan.-Feb. 2004, Pages:26 – 33
- A. Chervenak, E. Deelman, C. Kesselman, B. Allcock, I. Foster, V. Nefedova, J. Lee, A. Sim, A. Shoshani, B. Drach, D. Williams, D. Middleton , High-performance remote access to climate simulation data: a challenge problem for data grid technologies, *Parallel Computing* Volume 29, Issue 10 (October 2003), Pages: 1335-1356, 2003.
- E. Deelman, J. Blythe, Y. Gil, C. Kesselman, G. Mehta, K. Vahi, K. Blackburn, A. Lazzarini, A. Arbee, R. Cavanaugh, S. Koranda, Mapping Complex Workflows Onto Grid Environments, *Journal of Grid Computing*, Vol.1, No. 1, 2003., pp25-39.
- R. Williams, B. Berriman, E. Deelman, J. Good, J. Jacob, C. Kesselman, C. Lonsdale, S. Oliver, T. Prince. Multi-Wavelength Image Space: Another Grid-Enabled Science, *Journal of Concurrency and Computation: Practice and Experience*, Wiley, March 2003
- V.S. Adve, R. Bagrodia, E. Deelman, Rizos Sakellariou. Compiler-Optimized Simulation of Large-Scale Applications on High Performance Architectures. *Journal of Parallel and Distributed Computing*, Vol. 62, No. 3, Mar 2002, pp. 393-426.
- E. Deelman, B. K. Szymanski. Simulating Spatially Explicit Problems on High Performance Architectures, *Journal of Parallel and Distributed Computing*, Vol. 62, No. 3, Mar 2002, pp. 446-467.
- R. Bagrodia, E. Deelman, and T. Phan. Parallel Simulation of Large Scale Parallel Applications, *International Journal of High-Performance Computing Applications*. Volume 15, Number, 1, Spring 2001.
- V.S. Adve, R. Bagrodia, J.C. Browne, E. Deelman, A. Dube, E. Houstis, J. Rice, R. Sakellariou, D. Sundaram-Stukel, P. J. Teller, and M. K. Vernon. POEMS: End-to-end Performance Design of Large Parallel Adaptive Computational Systems, *IEEE Transactions on Software Engineering*, vol.26, no.11 p. 1027-48, November 2000.
- S. Prakash, E. Deelman, and R. Bagrodia. Asynchronous Parallel Simulation of Parallel Programs. *IEEE Transactions on Software Engineering*, 26(5), May 2000.
- T. Caraco, G. Gardner, W. Maniatty, E. Deelman, and B. K. Szymanski, Lyme disease: self-regulation and pathogen invasion. *Journal of Theoretical Biology*, 1998 Aug 21, 193(4):561-75.

Conference Papers

- Scott Callaghan, Philip Maechling, Ewa Deelman, Karan Vahi, Gaurang Mehta, Gideon Juve, Kevin Milner, Robert Graves, Edward Field, David Okaya, Dan Gunter, Keith Beattie, Thomas Jordan, “Reducing Time-to-Solution Using Distributed High-Throughput Mega-Workflows - Experiences from SCEC CyberShake”, Fourth IEEE International Conference on e-Science (e-Science 2008), 10-12 December 2008 in Indianapolis, Indiana, USA,
- Gideon Juve, Ewa Deelman, “Resource Provisioning Options for Large-Scale Scientific Workflows”, *Third International Workshop on Scientific Workflows and Business Workflow Standards in e-Science (SWBES)* in conjunction with Fourth IEEE International Conference on e-Science (e-Science 2008), 10 December 2008 in Indianapolis, Indiana, USA
- Christina Hoffa, Gaurang Mehta, Timothy Freeman, Ewa Deelman, Kate Keahey, Bruce Berriman, John Good, “On the Use of Cloud Computing for Scientific Workflows”, 3rd International Workshop on Scientific Workflows and Business Workflow Standards in e-Science (SWBES) in conjunction with Fourth IEEE International Conference on e-Science (e-Science 2008), 10 December 2008 in Indianapolis, Indiana, USA
- Ewa Deelman, Gurmeet Singh, Miron Livny, Bruce Berriman, John Good. “The Cost of Doing Science on the Cloud: The Montage Example” Proceeding of Super Computing 2008, Austin , Texas, November 2008
- Shishir Bharathi, Ann Chervenak, Ewa Deelman, Gaurang Mehta, Mei-Hui Su, Karan Vahi, “Characterization of Scientific Workflows”, 3rd Workshop on Workflows in Support of Large-Scale Science (WORKS08), Austin, TX, November, 2008

- Yang-Suk Kee, Ewa Deelman, Karan Vahi, Eunkyoo Byun, Jin-Soo Kim, “Pegasus on Virtual Grid: A Case Study of Workflow Planning over Captive Resources”, 3rd Workshop on Workflows in Support of Large-Scale Science (WORKS08), Austin, TX, November 2008
- Ewa Deelman, Ann Chervenak “Data Management Challenges of Large-Scale, Data-Intensive Scientific workflows” in WSES 08: 3rd International Workshop on Workflow Systems in e-Science, in conjunction with CCGrid 2008, May 2008, Lyon, France
- K.Lee, N. W. Paton, R. Sakellariou, E. Deelman, A. A. A. Fernandes, G. Mehta “Adaptive Workflow Processing and Execution in Pegasus”, 3rd International Workshop on Workflow Management and Applications in Grid Environments (WaGe08), in Proceedings of the Third International Conference on Grid and Pervasive Computing Symposia/Workshops, Pages 99-106, May 25-28 2008, Kunming, China
- Vijay Kumar, Mary Hall, Jihie Kim, Yolanda Gil, Tahsin Kurc, Ewa Deelman and Joel Saltz, “Designing and Parameterizing a Workflow for Optimization: A Case Study in Biomedical Imaging”, In Proc. of the Workshop on Next Generation Software, held in conjunction with IPDPS '08, April, 2008
- Gurmeet Singh, Mei-Hui Su, Karan Vahi, Ewa Deelman, Bruce Berriman, John Good, Daniel S. Katz, and Gaurang Mehta, "Workflow Task Clustering for Best Effort Systems with Pegasus", Mardi Gras Conference, Baton Rouge, LA, January 2008
- Simon Miles, Ewa Deelman, Paul Groth, Karan Vahi, Gaurang Mehta, Luc Moreau. Connecting Scientific Data to Scientific Experiments with Provenance, Third IEEE International Conference on e-Science and Grid Computing (e-Science 2007) Bangalore, India. , 2007
- Ann Chervenak, Ewa Deelman, Miron Livny, Mei-Hui Su, Rob Schuler, Shishir Bharathi, Gaurang Mehta, Karan Vahi, Data Placement for Scientific Applications in Distributed Environments, Proceedings of Grid Conference 2007, Austin, Texas, September 2007.
- Gurmeet Singh, Carl Kesselman, Ewa Deelman, Adaptive Pricing for Resource Reservations, 8th IEEE/ACM International Conference on Grid Computing (Grid 2007), Austin, Texas, September 2007.
- Yolanda Gil, Varun Ratnakar, Ewa Deelman, Gaurang Mehta, and Jihie Kim. Wings for Pegasus: Creating Large-Scale Scientific Applications Using Semantic Representations of Computational Workflows, Proceedings of the 19th Annual Conference on Innovative Applications of Artificial Intelligence (IAAI), Vancouver, British Columbia, Canada, July 22-26, 2007.
- Nandita Mandal, Ewa Deelman, Gaurang Mehta, Mei-Hui Su, and Karan Vahi, Integrating Existing Scientific Workflow Systems: The Kepler/Pegasus Example, Proceedings of the Second Workshop on Workflows in Support of Large-Scale Science (WORKS'07), in conjunction with the IEEE International Symposium on High Performance Distributed Computing Monterrey, CA, June 2007.
- Yolanda Gil, Pedro A. Gonzalez-Calero, Ewa Deelman. On the Black Art of Designing Computational Workflows, Proceedings of the Second Workshop on Workflows in Support of Large-Scale Science (WORKS'07), in conjunction with the IEEE International Symposium on High Performance Distributed Computing Monterrey, CA, June 2007.
- Gurmeet Singh, Carl Kesselman, Ewa Deelman, Adaptive Pricing for Resource Reservations The 8th IEEE/ACM International Conference on Grid Computing (Grid 2007)
- Gurmeet Singh, Carl Kesselman, Ewa Deelman, A Provisioning Model and its Comparison with Best-Effort for Performance-Cost Optimization in Grids, in IEEE International Symposium on High Performance Distributed Computing (HPDC) 2007
- Arun Ramakrishnan, Gurmeet Singh, Henan Zhao, Ewa Deelman, Rizos Sakellariou, K. Vahi, K. Blackburn, D. Meyers, and M. Samidi, Scheduling Data -Intensive Workflows onto Storage-Constrained Distributed Resources, in Seventh IEEE International Symposium on Cluster Computing and the Grid — CCGrid 2007
- Ewa Deelman and Yolanda Gil, Managing Large-Scale Scientific Workflows in Distributed Environments: Experiences and Challenges, Proceedings of Workflows in e-Science, Amsterdam, The Netherlands, 2006.
- E. Deelman, S. Callaghan, E. Field, H. Francoeur, R. Graves, N. Gupta, V. Gupta, T. H. Jordan, C. Kesselman, P. Maechling, J. Mehringer, G. Mehta, D. Okaya, K. Vahi, and L. Zhao, Managing Large-Scale Workflow Execution from Resource Provisioning to Provenance tracking: The

- CyberShake Example, Proceedings of e-Science, Amsterdam, The Netherlands, 2006. (**best paper**)
- G. Singh, C. Kesselman, and E. Deelman, Application-level Resource Provisioning on the Grid, Proceedings of e-Science, Amsterdam, The Netherlands, 2006.
 - V. Nefedova, R. Jacob, I. Foster, Z. Liu, Y. Liu, E. Deelman, G. Mehta, M.-H. Su, and K. Vahi, Automating Climate Science: Large Ensemble Simulations on the TeraGrid with the GriPhyN Virtual Data System, Proceedings of e-Science, Amsterdam, The Netherlands, 2006.
 - A. Lathers, M.-H. Su, A. Kulungowski, A. W. Lin, G. Mehta, S. T. Peltier, Ewa Deelman, and M. H. Ellisman, Enabling Parallel Scientific Applications with Workflow Tools, Proceedings of Challenges of Large Applications in Distributed Environments (CLADE), Paris, 2006.
 - Y. Gil, V. Ratnakar, and E. Deelman, Metadata Catalogs with Semantic Representation, Proceedings of International Provenance Annotation Workshop (IPAW-06), Chicago, IL, 2006.
 - J. Muench, H. Francoeur, D. Okaya, Y. Cui, P. Maechling, E. Deelman, G. Mehta, R. Moore, and T. Jordan, SCEC Earthworks Science Gateway: Widening SCEC Community Access to the TeraGrid. TeraGrid 2006 Conference, 2006.
 - Yolanda Gil, Varun Ratnakar, and Ewa Deelman. Virtual Metadata Catalogs: Augmenting Metadata Catalogs with Semantic Representations, Short paper at the Fourth International Semantic Web Conference (ISWC-05), Galway, Ireland, November 7-10, 2005.
 - E. Deelman, A. Galstyan, Y. Gil, M. Hall, K. Lerman, A. Nakano, P. Vashista, and J. Saltz, A Systematic Approach to Composing and Optimizing Application Workflows, Proceedings of Workshop on Patterns in High Performance Computing, Urbana-Champaign, 2005.
 - Daniel S. Katz, Joseph C. Jacob, G. Bruce Berriman, John Good, Anastasia C. Laity, Ewa Deelman, Carl Kesselman, Gurmeet Singh, Mei-Hui Su, Thomas A. Prince, A Comparison of Two Methods for Building Astronomical Image Mosaics on a Grid. ICPP Workshops 2005: 85-94
 - Seung-Hye Jang, Valerie Taylor, Xingfu Wu, Mieke Prajugo, Ewa Deelman, Gaurang Mehta, and Karan Vahi, Performance Prediction-based versus Load-based Site Selection: Quantifying the Difference, the 18th International Conference on Parallel and Distributed Computing Systems (PDCS-2005), Las Vegas, Nevada, 12-14 September 2005.
 - Jim Blythe, Sonal Jain, Ewa Deelman, Karan Vahi, Yolanda Gil, Anirban Mandal, Ken Kennedy, Task Scheduling Strategies for Workflow-based Applications in Grids, IEEE International Symposium on Cluster Computing and the Grid (CCGrid 2005).
 - Gurmeet Singh, Ewa Deelman, Gaurang Mehta, Karan Vahi, Mei-Hui Su, G. Bruce Berriman, John Good, Joseph C. Jacob, Daniel S. Katz, Albert Lazzarini, Kent Blackburn, and Scott Koranda, The Pegasus Portal: Web Based Grid Computing, The 20th Annual ACM Symposium on Applied Computing, SAC 2005.
 - Jim Blythe, Ewa Deelman, Yolanda Gil. Planning and Metadata on the Computational Grid, In AAAI Spring Symposium on Semantic Web Services, Palo Alto, California, USA, 2004.
 - Rattapoom Tuchinda, Snehal Thakkar, Yolanda Gil, and Ewa Deelman, Artemis: Integrating Scientific Data on the Grid, In the proceedings of the Sixteenth Innovative Applications of Artificial Intelligence, San Jose, CA, July 2004
 - Ewa Deelman, Gurmeet Singh, Malcolm P. Atkinson, Ann Chervenak, Neil P Chue Hong, Carl Kesselman, Sonal Patil, Laura Pearlman, Mei-Hui Su, Grid-Based Metadata Services, SSDBM04, Santorini, Greece, June 2004
 - I. Foster and others in the iVDGL project, The Grid2003 Production Grid: Principles and Practice, High Performance Distributed Computing, 2004. HPDC-13 2004.
 - J. C. Jacob, D. S. Katz, T. Prince, G. B. Berriman, J. C. Good, A. C. Laity, E. Deelman, G. Singh, and M.-H. Su, The Montage Architecture for Grid-Enabled Science Processing of Large, Distributed Datasets, Proceedings of the Earth Science Technology Conference (ESTC) 2004, June 2004.
 - G. B. Berriman, E. Deelman, J. Good, J. Jacob, D. S. Katz, C. Kesselman, A. Laity, T. A. Prince, G. Singh, and M. Su, Montage: a Grid Enabled Engine for Delivering Custom Science-Grade Image Mosaics on Demand, Proceedings of the SPIE Conference on Astronomical Telescopes and Instrumentation, June 2004.
 - Ewa Deelman, James Blythe, Yolanda Gil, Carl Kesselman, Gaurang Mehta, Sonal Patil, Mei-Hui Su, Karan Vahi, Miron Livny, Pegasus: Mapping Scientific Workflows onto the Grid, Across Grids Conference 2004, Nicosia, Cyprus.

- Jim Blythe, Ewa Deelman, Yolanda Gil. Planning and Metadata on the Computational Grid, In AAAI Spring Symposium on Semantic Web Services, 2004.
- Ewa Deelman, Raymond Plante, Carl Kesselman, Gurmeet Singh, Mei-Hui Su, Gretchen Greene, Robert Hanisch, Niall Gaffney, Antonio Volpicelli, James Annis, Vijay Sekhri, Tamas Budavari, Maria Nieto-Santisteban, William O'Mullane, David Bohlender, Tom McGlynn, Arnold Rots, Olga Pevunova, Grid-Based Galaxy Morphology Analysis for the National Virtual Observatory, Proceedings of SC 2003.
- G. B. Berriman , J. C. Good, A. C. Laity, A. Bergou, J. Jacob, D. S. Katz, E. Deelman, C. Kesselman, G. Singh, M.-H. Su, R. Williams, Montage A Grid Enabled Image Mosaic Service for the National Virtual Observatory, ADASS XIII, ASP Conference Series Vol XXX, F Ochsenein M Allen and D Egret eds, 2003.
- Gurmeet Singh, Shishir Bharathi, Ann Chervenak, Ewa Deelman, Carl Kesselman, Mary Manohar, Sonal Patil, and Laura Pearlman. A Metadata Catalog Service for Data Intensive Applications, Proceedings of SC 2003.
- Ewa Deelman, James Blythe, Yolanda Gil, Carl Kesselman, Scott Koranda, Albert Lazzarini, Gaurang Mehta, Maria Alessandra Papa, Karan Vahi, Pegasus and the Pulsar Search: From Metadata to Execution on the Grid., Applications Grid Workshop, PPAM 2003, Czestochowa, Poland 2003.
- J. Blythe, E. Deelman, Y. Gil, C. Kesselman, A. Agarwal, G. Mehta, K. Vahi, The Role of Planning in Grid Computing, ICAPS, 2003
- J. Blythe, E. Deelman, Y. Gil, C. Kesselman, Transparent Grid Computing: a Knowledge-Based Approach. IAAI, 2003.
- A. Chervenak, E. Deelman, I. Foster, L. Guy, W. Hoschek, A. Iamnitchi, C. Kesselman, P. Kunst, M. Ripeanu, B. Schwartzkopf, H. Stockinger, K. Stockinger, B. Tierney. Giggle: A Framework for Constructing Scalable Replica Location Services. Proceedings of Supercomputing 2002 (SC2002), November 2002.
- E. Deelman, C. Kesselman, G. Mehta, L. Meshkat, L. Pearlman, K. Blackburn, P. Ehrens, A. Lazzarini, R. Williams, S. Koranda. GriPhyN and LIGO, building a virtual data grid for gravitational wave scientists. High Performance Distributed Computing, 2002. HPDC-11 2002. Page(s): 225 -234
- E. Deelman, C. Kesselman, S. Koranda, K. Blackburn A. Lazzarini, and R. Williams. Applications of Virtual Data in the LIGO Experiment, Proceedings of the International Conference on Parallel Processing and Applied Mathematics, Naleczow, Poland, September 2001, Lecture Notes in Computer Science. Vol.2328; Berlin, Germany : Springer-Verlag, 2002, xix+915 p. (23-34)
- W. Allcock, I. Foster, V. Nefedova, A. Chervenak, E. Deelman, C. Kesselman, J. Lee, A. Sim, A. Shoshani, B. Drach, D. Williams. High-Performance Remote Access to Climate Simulation Data: A Challenge Problem for Data Grid Technologies. Proceedings of SC 2001, Denver, CO, November 2001.
- E. Deelman, R. Bagrodia, R. Sakellariou, V. Adve. Improving Lookahead in Parallel Discrete Event Simulations of Large-Scale Applications Using Compiler Analysis. Proceedings 15th Workshop on Parallel and Distributed Simulation p. 5-13, Lake Arrowhead, CA, USA; 15-18 May 2001. (**best paper**)
- E. Deelman, C. Kesselman, R.D. Williams, A. Lazzarini, T.A. Prince, J. Romano, B. Allen, A Virtual Data Grid for LIGO, Lecture Notes in Computer Science 2110 (2001) 3, also Proc. HPCN 2001, Amsterdam, June 2001.
- V. S. Adve, R. Bagrodia, E. Deelman, T. Phan, and R. Sakellariou, Compiler-Supported Simulation of Highly Scalable Parallel Applications, Proceedings of SC'99, Portland, OR, Nov 13 - 19, 1999.
- R. Bagrodia, E. Deelman, S. Docy, T. Phan; Performance Prediction of Large Parallel Applications Using Parallel Simulations. ACM SIGPLAN 1999 Symposium on Principles and Practice of Parallel Programming, Atlanta, Georgia, May 4-6, 1999.
- E. Deelman, A. Dube, A. Hoisie, Y. Luo, R. Oliver, D. Sundaram-Stukel, H. Wasserman, V. S. Adve, R. Bagrodia, J. C. Browne, E. Houstis, O. Lubeck, J. Rice, P. Teller, M. K. Vernon; POEMS: End-to-end Performance Design of Large Parallel Adaptive Computational Systems.

- Proceedings of the First International Workshop on Software and Performance '98 - WOSP '98, Santa Fe, New Mexico, October 12-16, 1998.
- E. Deelman, and B.K. Szymanski, Dynamic load balancing in parallel discrete event simulation for spatially explicit problems. Proceedings of the Twelfth Workshop on Parallel and Distributed Simulation, PADS '98, Banff, CA, May 26-29, 1998.
 - E. Deelman and B.K. Szymanski. System Knowledge Acquisition in Parallel Discrete Event Simulation. Proceedings of the 1997 IEEE International Conference on Systems Man and Cybernetics, Orlando, FL, October 12-15, 1997
 - E. Deelman and B.K. Szymanski. Continuously Monitored Global Virtual Time. Proceedings of the International Conference on Parallel and Distributed Processing Techniques and Applications, Las Vegas, NV, June 30- July 3, 1997.
 - E. Deelman and B.K. Szymanski. Breadth-First Rollback in Spatially Explicit Simulations. Proceedings of the 11th Workshop on Parallel and Distributed Simulation, PADS'97, Lockenhaus, Austria, June 10-13, 1997.
 - E. Deelman, B.K. Szymanski and T. Caraco. Simulating Lyme Disease Using Parallel Discrete Event Simulation. Proceedings of the 1996 Winter Simulation Conference, San Diego, CA, December 8-11, 1996.
 - B.K. Szymanski, E. Deelman, J. Flaherty, C. Norton, J. Teresco and L. Ziantz. Parallel Scientific Computing on the IBM SP2 at SCOREC-Rensselaer Polytechnic Institute. Proceedings of SUP'EUR 96, Krakow, Poland, September 8-11, 1996.
 - E. Deelman, T. Caraco and B.K. Szymanski. Parallel Discrete Event Simulation of Lyme Disease. Biocomputing: Proc. of the 1996 Pacific Symposium, Hawaii, January 3-6, 1996.
 - E. Deelman, W. K. Kaplow, B. K. Szymanski, P. Tannenbaum and L. Ziantz. Integrating Data and Task Parallelism in Scientific Programs. Proceeding of the Workshop on Languages, Compilers and Run-Time Systems for Parallel Programming, Kluwer Academic Publishers, Boston, 1995.

Book Chapters

- Ewa Deelman, Bruce Berriman, Ann Chervenak, Oscar Corcho, Paul Groth, Luc Moreau, Metadata and provenance management, in *Scientific Data Management, to appear 2009*
- Ewa Deelman, Miron Livny, Gaurang Mehta, Andy Pavlo, Gurmeet Singh, Mei-Hui Su, Karan Vahi, R. Kent Wenger, Pegasus and DAGMan From Concept to Execution: Mapping Scientific Workflows onto Today's Cyberinfrastructure, in "High Performance Computing and Grids in Action" (L. Grandinetti Editor), IOS Press, Amsterdam, volume 16 in the series "Advances in Parallel Computing". March 2008
- G. Bruce Berriman, Ewa Deelman, John Good, Joseph C. Jacob, Daniel S. Katz, Anastasia C. Laity, Thomas A. Prince, Gurmeet Singh, and Mei-Hui Su, Generating Complex Astronomy Workflows, in *Workflows for e-Science, 2007*
- Philip Maechling, Ewa Deelman, Li Zhao, Robert Graves, Gaurang Mehta, Nitin Gupta, John Mehringer, Carl Kesselman, Scott Callaghan, David Okaya, Hunter Francoeur, Vipin Gupta, Yifeng Cui, Karan Vahi, Thomas Jordan, and Edward Field, SCEC CyberShake Workflows—Automating Probabilistic Seismic Hazard Analysis Calculations, in *Workflows for e-Science, 2007*
- Ewa Deelman, Gaurang Mehta, Gurmeet Singh, Mei-Hui Su, and Karan Vahi, Pegasus: Mapping Large-Scale Workflows to Distributed Resources, in *Workflows for e-Science, 2007*
- Ewa Deelman, Looking into the Future of Workflows: The Challenges Ahead, in *Workflows for e-Science, 2007*
- Craig A. Lee, B. Scott Michel, Ewa Deelman and Jim Blythe, From Event-Driven Workflows Towards a Posteriori Computing, In *Future Generation Grids as part of the Springer LNCS series, Getov, Vladimir; Laforenza, Domenico; Reinefeld, Alexander (Eds.)*, 2006.
- D.S. Katz, N. Anagnostou, G.B. Berriman, E. Deelman, J. Good, J.C. Jacob, C. Kesselman, A. Laity, T.A. Prince, G.Singh, M.-H. Su, and R. Williams Astronomical Image Mosaicking on a Grid: Initial Experiences in Engineering the Grid-Status and Perspectives, B. Di. Martino, J. Dongarra, A. Hoisie, L. Yand, and H. Zima eds. Nova, 2005

- Ewa Deelman, James Blythe, Yolanda Gil, and Carl Kesselman, Workflow Management in GriPhyN, in Grid Resource Management, J. Nabrzyski, J. Schopf, and J. Weglarz editors, Kluwer, 2003.

Technical Reports

- Ann Chervenak, Ewa Deelman, Carl Kesselman, Laura Pearlman, Gurmeet Singh, A Metadata Catalog Service for Data Intensive Applications, GriPhyN technical report, 2002-11
- Leila Meshkat, William Allcock, Ewa Deelman, Carl Kesselman, Fault Location in Grids Using Fault Location in Grids Using Bayesian Belief Networks. GriPhyN technical report, 2002-8.
- K. Blackburn, P. Ehrens, A. Lazzarini, R. Williams, E. Deelman, C. Kesselman, G. Mehta, L. Meshkat, L. Pearlman, B. Allen, S. Koranda, LIGO Prototype, GriPhyN technical report 2001-18.
- E. Deelman, C. Kesselman, G. Mehta, Transformation Catalog Design for GriPhyN, Prototype of Transformation Catalog Schema. GriPhyN technical report 2001-17.
- E. Deelman, I. Foster, C. Kesselman, M. Livny, Representing Virtual Data: A Catalog Architecture for Location and Materialization Transparency, GriPhyN technical report 2001-13.
- LIGO's Virtual Data Requirements, B. Allen, E. Deelman, C. Kesselman, A. Lazzarini, T. Prince, J. Romano, R. Williams, GriPhyN technical report 2001-6 also LIGO technical report T000135-00-D.
- Continuously Monitored Global Virtual Time in Parallel Discrete Event Simulation., E. Deelman and B.K. Szymanski. Department of Computer Science Technical Report 96-18, Rensselaer Polytechnic Institute, 1996.