Collaborative provenance for workflow-driven science and engineering

Altıntaş, İ.

Publication date
2011

Publications

[1] Lin, Abel W., Ilkay Altintas, Chris Churas, Madhusudan Gujral, Jeffrey Grethe and Mark Ellisman (2011 (In print.)). *REST: From Research to Practice (C. Pautasso and E. Wilde, Eds.)*. Chap. 16: Case Study on the Use of REST Architectural Principles for Scientific Analysis: CAMERA - Community Cyberinfrastructure for Advanced Microbial Ecology Research and Analysis. Springer.

[2] Altintas, Ilkay, Abel W. Lin, Jing Chen, Chris Churas, Madhusudan Gujral, Shulei Sun, Weizhong Li, Ramil Manansala, Mayya Sedova, Jeffrey S. Grethe and Mark Ellisman (2010*a*). Camera 2.0: A data-centric metagenomics community infrastructure driven by scientific workflows. In *SWF 2010 in conjunction with 6th World Congress on Services (SERVICES 2010)*. IEEE Computer Society. pp. 352–359.

[3] Altintas, Ilkay, Daniel Crawl, C.J. Crosby and Peter Cornillon (2010*b*). *Geoinformatics: Cyberinfrastructure for the Solid Earth Sciences (Keller, G. Randall and Baru, Chaitanya Eds.)*. Chap. Scientific Workflows for the Geosciences: An Emerging Approach to Building Integrated Data Analysis Systems. Cambridge University Press.

[4] Altintas, Ilkay, Jing Jing Chen, Mayya Sedova, Amarnath Gupta, Shulei Sun, Abel W. Lin, Madhusudan Gujral, Manish K. Anand, Weizhong Li, Jeffrey S. Grethe and Mark Ellisman (2010*c*). Extending the data model for data-centric metagenomics analysis using scientific workflows in camera. In *Proceedings of HPC for Life Sciences Workshop at eScience 2010*.

[5] Altintas, Ilkay, Manish Anand, Daniel Crawl, Shawn Bowers, Adam Belloum, Paolo Missier, Bertram Ludäscher, Carole Goble and Peter Sloot (2010*d*). Understanding collaborative studies through interoperable workflow provenance. In *Provenance and Annotation of Data and Processes* (Deborah McGuinness, James Michaelis and Luc Moreau, Eds.). Vol. 6378. pp. 42–58. Springer Berlin / Heidelberg.

[6] Altintas, Ilkay, Manish Kumar Anand, Adam Belloum, Gokhan Ertaylan, Bartosz Balis, Marian Bubak and Peter M.A. Sloot (2010*e*). Collaborative provenance for workflow-driven science - a position paper. Submitted to FGCS.

[7] Altintas, Ilkay, Manish Kumar Anand, Trung Vuong, Shawn Bowers, Bertram Ludäscher and Peter M.A. Sloot (2010*f*). A data model for analyzing user collaborations in workflow-driven escience. Submitted to the International Journal of Computers and Their Applications (IJCA), Special Issue on Scientific Workflows, Provenance and Their Applications.

[8] Anand, Manish, Shawn Bowers, Ilkay Altintas and Bertram Ludäscher (2010). Approaches for exploring and querying scientific workflow provenance graphs. In *Provenance and Annotation of Data and Processes* (Deborah McGuinness, James Michaelis and Luc Moreau, Eds.). Vol. 6378 of *Lecture Notes in Computer Science*. Springer Berlin / Heidelberg. pp. 17–26.

[9] Barseghian, Derik, Ilkay Altintas, Matthew B. Jones, Daniel Crawl, Nathan Potter, James Gallagher, Peter Cornillon, Mark Schildhauer, Elizabeth T. Borer, Eric W. Seabloom and Parviez R. Hosseini (2010). Workflows and extensions to the kepler scientific workflow system to support environmental sensor data access and analysis. *Ecological Informatics*. **5**(1), 42–50.

[10] Missier, Paolo, Carole Goble, Saumen Dey, Anandarup Sarkar, Biva Shresta, Bertram Ludäscher, Shawn Bowers, Ilkay Altintas and Manish Kumar Anand (2010). Linking multiple workflow provenance traces for interoperable collaborative science. In *Proceedings of the 5th Workshop on Workflows in Support of Large-Scale Science*. WORKS'10. ACM. New York, NY, USA. pp. 1–8.

[11] Mouallem, Pierre, Daniel Crawl, Ilkay Altintas, Mladen A. Vouk and Ustun Yildiz (2010). A fault-tolerance architecture for kepler-based distributed scientific workflows. In *Proceedings of Scientific and Statistical Database Management, 22nd International Conference (SSDBM 2010)*. Vol. 6187 of *Lecture Notes in Computer Science*. Springer. Berlin, Heidelberg. pp. 452–460.

[12] Sun, Shulei, Jing Chen, Li Weizhong, Jeffrey Grethe, Ilkay Altintas, Abel W. Lin, Steve Peltier, Karen Stocks, Eric E. Allen, Mark Ellisman and John Wooley (2010). Community cyberinfrastructure for advanced microbial ecology research and analysis - the camera resource. *Nucleic Acids Research, Database Issue*.

[13] Wang, Jianwu, Prakashan Korambath, Seonah Kim, Scott Johnson, Kejian Jin, Daniel Crawl, Ilkay Altintas, Shava Smallen, Bill Labate and Kendall N. Houk (2010). Theoretical enzyme design using the kepler scientific workflows on the grid. In *Procedia Computer Science, ICCS 2010*. Vol. 1. pp. 1169 – 1178.

[14] Goderis, Antoon, Christopher Brooks, Ilkay Altintas, Edward A. Lee and Carole A. Goble (2009). Heterogeneous composition of models of computation. *Future Generation Computer Systems*. **25**(5), 552–560.

[15] Ludäscher, Bertram, Ilkay Altintas, Shawn Bowers, Julian Cummings, Terence Critchlow, Ewa Deelman, David De Roure, Juliana Freire, Carole Goble, Matthew Jones, Scott Klasky, Timothy Mcphillips, Norbert Podhorszki, Cláudio T. Silva, Ian Taylor and Mladen A. Vouk (2009). *Scientific Data Management: Challenges, Technology, and Deployment (Shoshani, Arie and Rotem, Doron, Eds.)*. Chap. Chapter 13: Scientific Process Automation and Workflow Management. Computational Science Series. Chapman and Hall/CRC.

[16] Wang, Jianwu, Daniel Crawl and Ilkay Altintas (2009*a*). Kepler + Hadoop: A general architecture facilitating data-intensive applications in scientific workflow systems. In *WORKS '09: Proceedings of the 4th Workshop on Workflows in Support of Large-Scale Science*. ACM New York, NY, USA. Portland, Oregon. pp. 1–8.

[17] Wang, Jianwu, Ilkay Altintas, Parviez R. Hosseini, Derik Barseghian, Daniel Crawl, Chad Berkley and Matthew B. Jones (2009*b*). Accelerating parameter sweep workflows by utilizing ad-hoc network computing resources: An ecological example. In *Services, IEEE Congress on*. IEEE Computer Society. pp. 267–274.

[18] Altintas, Ilkay (2008). Lifecycle of scientific workflows and their provenance: A usage perspective. In *SERVICES '08: Proceedings of the 2008 IEEE Congress on Services - Part I*. IEEE Computer Society. Washington, DC, USA. pp. 474–475.

[19] Abramson, David, Colin Enticott and Ilkay Altintas (2008). Nimrod/k: towards massively parallel dynamic grid workflows. In *SC '08: Proceedings of the 2008 ACM/IEEE conference on Supercomputing*. IEEE/ACM. Piscataway, NJ, USA. pp. 1–11.

[20] Crawl, Daniel and Ilkay Altintas (2008). A provenance-based fault tolerance mechanism for scientific workflows. In *Provenance and Annotation of Data and Processes (IPAW 2008, Revised Selected Papers)* (Juliana Freire, David Koop and Luc Moreau, Eds.). Vol. 5272 of *Lecture Notes in Computer Science*. Springer. pp. 152–159.

[21] Ludäscher, Bertram, Norbert Podhorszki, Ilkay Altintas, Shawn Bowers and Timothy M. McPhillips (2008). From computation models to models of provenance: the rws approach. *Concurrency and Computation: Practice and Experience*. **20**(5), 507–518.

[22] Moreau, Luc, Bertram Ludäscher, Ilkay Altintas, Roger S. Barga, Shawn Bowers, Steven P. Callahan, George Chin Jr., Ben Clifford, Shirley Cohen, Sarah Cohen Boulakia, Susan B. Davidson, Ewa Deelman, Luciano A. Digiampietri, Ian T. Foster, Juliana Freire, James Frew, Joe Futrelle, Tara Gibson, Yolanda Gil, Carole A. Goble, Jennifer Golbeck, Paul T. Groth, David A. Holland, Sheng Jiang, Jihie Kim, David Koop, Ales Krenek, Timothy M. McPhillips, Gaurang Mehta, Simon Miles, Dominic Metzger, Steve Munroe, Jim Myers, Beth Plale, Norbert Podhorszki, Varun Ratnakar, Emanuele Santos, Carlos Eduardo Scheidegger, Karen Schuchardt, Margo I.

Seltzer, Yogesh L. Simmhan, Cláudio T. Silva, Peter Slaughter, Eric G. Stephan, Robert Stevens, Daniele Turi, Huy T. Vo, Michael Wilde, Jun Zhao and Yong Zhao (2008). Special issue: The first provenance challenge. *Concurrency and Computation: Practice and Experience*. **20**(5), 409–418.

[23] Wang, Jianwu, Ilkay Altintas, Chad Berkley, Lucas Gilbert and Matthew B. Jones (2008). A high-level distributed execution framework for scientific workflows. *IEEE International Conference on eScience*. pp. 634–639.

[24] Goderis, Antoon, Christopher Brooks, Ilkay Altintas, Edward Lee and Carole Goble (2007). Composing different models of computation in kepler and ptolemy ii. In *Computational Science – ICCS 2007* (Yong Shi, Geert van Albada, Jack Dongarra and Peter Sloot, Eds.). Vol. 4489. Springer Berlin / Heidelberg. pp. 182–190.

[25] Vouk, M., I. Altintas, R. Barreto, J. Blondin, Z. Cheng, T. Critchlow, A. Khan, S. Klasky, J. Ligon, B. Ludaescher, P. Mouallem, S. Parker, N. Podhorszki, A. Shoshani and C. Silva (2007). Automation of network-based scientific workflows. In *Grid-Based Problem Solving Environments* (Patrick Gaffney and James Pool, Eds.). Vol. 239 of *IFIP International Federation for Information Processing*. Springer Boston. pp. 35–61.

[26] Altintas, Ilkay, Oscar Barney and Efrat Jaeger-Frank (2006*a*). Provenance collection support in the kepler scientific workflow system. In *Provenance and Annotation of Data (IPAW 2006, Revised Selected Papers)* (Luc Moreau and Ian Foster, Eds.). Vol. 4145 of *Lecture Notes in Computer Science*. Springer Berlin / Heidelberg. pp. 118–132.

[27] Altintas, Ilkay, Oscar Barney, Zhengang Cheng, Terence Critchlow, Bertram Ludaescher, Steve Parker, Arie Shoshani and Mladen Vouk (2006*b*). Accelerating the scientific exploration process with scientific workflows. In *SciDAC 2006, SCIENTIFIC DISCOVERY THROUGH ADVANCED COMPUTING* (Dr. William M Tang, Ed.). Vol. Vol. 46. pp. 468–478.

[28] Hou, Chien-Yi, Ilkay Altintas, E. Jaeger-Frank, L. Gilbert, R. Moore, A. Rajasekar and R. Marciano (2006). A scientific workflow solution to the archiving of digital media. In *Workshop on Workflows in Support of Large-Scale Science, 2006*. WORKS '06. IEEE. pp. 1–10.

[29] Jaeger-Frank, E., C.J. Crosby, A. Memon, V. Nandigam, J. Conner, J.R. Arrowsmith, I. Altintas and C. Baru (2006*a*). A three tier architecture applied to LiDAR processing and monitoring. *Scientific Programming*. **14**(3), 185–194.

[30] Jaeger-Frank, E., C.J. Crosby, A. Memon, V. Nandigam, J.R. Arrowsmith, J. Conner, I. Altintas and C. Baru (2006*b*). Three tier architecture for LiDAR interpolation and

analysis. In *1st International Workshop on Workflow Systems in e-Science in Conjunction with ICCS 2006* (Vassil N. Alexandrov, G. Dick van Albada, Peter M. A. Sloot and Jack Dongarra, Eds.). Vol. 3 of *Lecture Notes in Computer Science*. Springer. Berlin, Heidelberg. pp. 920–927. 1st International Workshop on Workflow Systems in e-Science (WSES06).

[31] Ludäscher, Bertram, Ilkay Altintas, Chad Berkley, Dan Higgins, Efrat Jaeger-Frank, Matthew Jones, Edward Lee, Jing Tao and Yang Zhao (2006). Scientific workflow management and the kepler system. *Concurrency and Computation: Practice and Experience, Special Issue on Scientific Workflows*. **18**(10), 1039–1065.

[32] Sloot, Peter M.A., Alfredo Tirado-Ramos, Ilkay Altintas, Marian T. Bubak and Charles A. Boucher (November 2006). From molecule to man: Decision support in individualized e-health. *IEEE Computer*. **39**(11), 40–46.

[33] Sudholt, Wibke, Ilkay Altintas and Kim Baldridge (2006). Scientific workflow infrastructure for computational chemistry on the grid. In *Computational Science – ICCS 2006* (Vassil Alexandrov, Geert van Albada, Peter Sloot and Jack Dongarra, Eds.). Vol. 3993 of *Lecture Notes in Computer Science*. Springer Berlin / Heidelberg. pp. 69–76.

Abramson, David, Jagan Kommineni and Ilkay Altintas (2005). Flexible io services in the kepler grid workflow system. In *E-SCIENCE '05: Proceedings of the First International Conference on e-Science and Grid Computing*. IEEE Computer Society. Washington, DC, USA. pp. 255–262.

[34] Baldridge, Kim K., Jerry P. Greenberg, Wibke Sudholt, Steve Mock, Ilkay Altintas, Celine Amoreira, Yohann Potier, Adam Birnbaum, Karan Bhatia and Michela Taufer (2005*a*). The computational chemistry prototyping environment. In *Proceedings of the IEEE*. Vol. 93. IEEE. pp. 510–521.

[35] Baldridge, Kim K., Wibke Sudholt, Jerry P. Greenberg, Celine Amoreira, Yohann Potier, Ilkay Altintas, Adam Birnbaum, David Abramson, Colin Enticott and Slavisa Garic (2005*b*). *Parallel Computing for Bioinformatics (ed. A. Y. Zomaya)*. Chap. Cluster and Grid Infrastructure for Computational Chemistry and Biochemistry. John Wiley & Sons.

[36] Jaeger, Efrat, Ilkay Altintas, Jianting Zhang, Bertram Ludäscher, Deana Pennington and William Michener (2005). A scientific workflow approach to distributed geospatial data processing using web services. In *Proceedings of the 17th international conference on Scientific and statistical database management*. Lawrence Berkeley Laboratory. Berkeley, CA, US. pp. 87–90.

[37] Altintas, Ilkay, Adam Birnbaum, Kim K. Baldridge, Wibke Sudholt, Mark Miller, Celine Amoreira, Yohann Potier and Bertram Ludaescher (2004*a*). A framework for the

design and reuse of grid workflows. In *Scientific Applications of Grid Computing: First International Workshop; Lecture Notes in Computer Science* (P. Herrero, M.S. Perez and V. Robles, Eds.). number 3 In *Lecture Notes in Computer Science*. Springer-Verlag GmbH. pp. 119–132.

[38] Altintas, Ilkay, Chad Berkley, Efrat Jaeger, Matthew B. Jones, Bertram Ludäscher and Steve Mock (2004*b*). Kepler: An extensible system for design and execution of scientific workflows. In *Proceedings of the 16th International Conference on Scientific and Statistical Database Management (SSDBM 2004)*. IEEE Computer Society. Washington, DC, USA. pp. 423–424.

[39] Altintas, Ilkay, Efrat Jaeger, Kai Lin, Bertram Ludaescher and Ashraf Memon (2004*c*). A web service composition and deployment framework for scientific workflows. In *ICWS '04: Proceedings of the IEEE International Conference on Web Services*. IEEE Computer Society. Washington, DC, USA. p. 814.

[40] Altintas, Ilkay, Sangeeta Bhagwanani, David Buttler, Sandeep Chandra, Zhengang Cheng, Matthew Coleman, Terence Critchlow, Amarnath Gupta, Wei Han, Ling Liu, Bertram Ludaescher, Calton Pu, Reagan Moore, Arie Shoshani and Mladen Vouk (2003). A modeling and execution environment for distributed scientific workflows. In *15th International Conference on Scientific and Statistical Database Management (SSDBM)*. Boston, Massachusetts. pp. 247–250.

[41] Ludascher, Bertram, Ilkay Altintas and Amarnath Gupta (2003). Compiling abstract scientific workflows into web service workflows. In *Proceedings of the 15th International Conference on Scientific and Statistical Database Management*. pp. 251–254. IEEE Computer Society. Washington, DC, USA.

[42] Ludäscher, Bertram, Ilkay Altintas and Amarnath Gupta (2002). Time to leave the trees: From syntactic to conceptual querying of xml. In *XML-Based Data Management and Multimedia Engineering — EDBT 2002 Workshops* (Akmal Chaudhri, Rainer Unland, Chabane Djeraba and Wolfgang Lindner, Eds.). Vol. 2490 of *Lecture Notes in Computer Science*. pp. 774–778. Springer Berlin / Heidelberg.

[43] Altintas, Ilkay (2001). A comparative study for component oriented design modeling. Master's thesis. Middle East Technical University. Ankara, Turkey.

[44] Altintas, Ilkay and Ali H. Dogru (2001). Case studies using course-projects for component oriented modeling. In *Workshop on Transdisciplinary Education, Research & Training*.

[45] Dogru, Ali H. and Ilkay Altintas (2000). Modeling language for component oriented software engineering: Coseml. In *The 5th Biennial World Conference in Integrated Design & Process Technology*.