Final Report

Department of Energy Grant DEFG02-02ER25543

Creating a New Generation of Software Development Environments, Compilers, and Algorithms for High Performance Computing, Networks, and Data Management

Robert A. van Engelen
Computer Science Department
Florida State University
Tallahassee, FL 32306-4530
engelen@cs.fsu.edu

January 10, 2007

Contents

1 Introduction

The PI was tenured and promoted to Associate professor in Computer Science at Florida State University in 2004 (two years into the grant project).

The funding was used to support several graduate students (MS and PhD), to prepare research publications, to deliver presentations at conferences, to purchase small equipment, and to provide summer month research support for the PI.

This report gives the expenditures and lists the details of the accomplishments, a summary of research results obtained, and the artifacts produced with the DOE grant funding.

2 Project Results and Accomplishments

2.1 Project Production

This section describes the production of the project for the duration of the award (August 15, 2002 to August 14, 2006).

2.1.1 Journal Publications Citing DOE Support (Reviewed)

The reviewed journal publications produced by the project are listed below. These publications acknowledge DOE grant support.

1. Robert van Engelen, *A Framework for Service-Oriented Computing with C and C++ Web Service Components*, accepted for publication in ACM Transactions on Internet Technologies, submitted and revised April 2006, to appear 2007.

2. Robert A. van Engelen, Kyle Gallivan, and Burt Walsh, *Parametric Timing Estimation With the Newton-Gregory Formulae*, Journal of Concurrency and Computation: Practice and Experience, Volume 18, Number 10, September 2006, pages 1434–1464.

3. Prasad Kulkarni, Wankang Zhao, David Whalley, Xin Yuan, Robert van Engelen, Kyle Gallivan, Jason Hiser, Jack Davidson, Baosheng Cai, Mark Bailey, Hwashin Moon, Kyunghwan Cho, Yunheung Paek, and Douglas Jones, *VISTA: VPO Interactive System for Tuning Applications*, in the ACM Transactions on Embedded Computing Systems, November 2006, volume 5, number 3, pages 816–863.

4. W. Kreahling, D. Whalley, M. Bailey, X. Yuan, G-R. Uh, and R. van Engelen, *Branch Elimination by Condition Merging*, in the Journal of Software Practice and Experience, 2005, Volume 35, Number 1, page 51–74.

5. Yaohang Li, Michael Mascagni, Robert van Engelen, and Qin Cai, *A Grid Workflow-Based Monte Carlo Simulation Environment*, in the Journal of Neural, Parallel, and Scientific Computations, 2004, pages 439–454.

6. Robert van Engelen, David Whalley, and Xin Yuan, *Automatic Validation of Code-Improving Transformations on Low-Level Program Representations*, in the Science of Computer Programming Journal, 2004, Volume 52, pages 257–280.

7. Robert van Engelen, Gunjan Gupta, and Saurabh Pant, *Developing Web Services for C and C++*, in IEEE Internet Computing Journal, March, 2003, pages 53–61.

2.1.2 Conference Publications Citing DOE Support (Reviewed)

The reviewed journal publications produced by the project are listed below. These publications acknowledge DOE grant support.

8. M. Head, M. Govindaraju, R. van Engelen, and W. Zhang, *Benchmarking XML Processors for Applications in Grid Web Services*, in the proceedings of Supercomputing 2006.

9. J. Birch R. van Engelen K. A. Gallivan Y. Shou, *An Empirical Evaluation of Chains of Recurrences for Array Dependence Testing*, in the proceedings of to the ACM/IEEE International Conference on Parallel Architectures and Compilation Techniques (PACT), 2006, pages 295–304, (25.6% acceptance ratio).

10. Robert van Engelen, Madhusudhan Govindaraju, Wei Zhang, *Exploring Remote Object Coherence in XML Web Services*, in the proceedings of the IEEE International Conference on Web Services (ICWS), 2006, pages 249–256.

11. Wei Zhang, Robert van Engelen, *A Table-Driven XML Streaming Methodology for High-Performance Web Services*, in the proceedings of the IEEE International Conference on Web Services (ICWS), 2006, pages 197–206, *best student paper award*.

12. Yixin Shou, Robert van Engelen, Johnnie Birch, and Kyle Gallivan, *Toward Efficient Flow-Sensitive Induction Variable Analysis and Dependence Testing for Loop Optimization*, in the proceedings of the ACM SouthEast conference, 2006, pages 1–6, (48% acceptance ratio), *best paper award* (selected out of 100 papers).

13. Wei Zhang and Robert van Engelen, *TDX: a High-Performance Table-Driven XML Parser*, in the proceedings of the ACM SouthEast conference, 2006, pages 726–731, (48% acceptance ratio).

14. Subhajit Datta and Robert van Engelen, *Effects of Changing Requirements: A Tracking Mechanism for the Analysis Workflow*, in the proceedings of the ACM Symposium on Applied Computing, 2006, 6 pages (32% acceptance ratio).

15. Michael R. Head, Madhusudhan Govindaraju, Aleksander Slominski, Pu Liu, Nayef Abu-Ghazaleh, Robert van Engelen, Kenneth Chiu, Michael J. Lewis, *Benchmark Suite for SOAP-based Communication in Grid Web Services*, Supercomputing Conference 2005 (24% acceptance ratio).

3

16. Giovanni Aloisio, Massimo Cafaro, Italo Epicoco, Daniele Lezzi, and Robert van Engelen, *The GSI plug-in for gSOAP: Enhanced Security, Performance, and Reliability*, in the proceedings of the IEEE ITCC conference 2005, IEEE Press, Volume I, pages 304–309.

17. R.A. van Engelen, J. Birch, Y. Shou, B. Walsh, and K. A. Gallivan, *A Unified Framework for Nonlinear Dependence Testing and Symbolic Analysis*, in the proceedings of the ACM International Conference on Supercomputing, 2004, pages 106-115, (22% acceptance ratio).

18. Robert van Engelen, *Constructing Finite State Automata for High Performance XML Web Services*, in the proceedings of the International Symposium on Web Services (ISWS), 2004, pages 975–981.

19. Robert van Engelen, *Code Generation Techniques for Developing Light-Weight XML Web Services for Embedded Devices*, in the proceedings of the 9th ACM Symposium on Applied Computing (SAC), Nicosia, Cyprus, 2004, pages 854–861, (41% acceptance ratio).

20. Giovanni Aloisio, Massimo Cafaro, Daniele Lezzi, and Robert van Engelen, *Secure Web Services with Globus GSI and gSOAP*, in proceedings of the International Conference on Parallel and Distributed Computing (EUROPAR) 2003, August, 2003, (6 pages, 47% acceptance ratio).

21. William Kreahling, David Whalley, Mark Bailey, Xin Yuan, Gang-Ryung Uh, Robert van Engelen, *Branch Elimination via Multi-Variable Condition Merging*, in proceedings of the International Conference on Parallel and Distributed Computing (EUROPAR) 2003, August 2003, (6 pages, 47% acceptance ratio).

22. Yaohang Li, Michael Mascagni, and Robert van Engelen, *GCIMCA: A Globus and SPRNG Implementation of a Grid-Computing Infrastructure for Monte Carlo Applications*, in proceedings of the Parallel and Distributed Processing Techniques and Applications (PDPTA) conference, Las Vegas, 2003, (CD-ROM proceedings, 5 pages).

23. Robert A. van Engelen, *Pushing the SOAP Envelope with Web Services for Scientific Computing*, in the proceedings of the International Conference on Web Services (ICWS), Las Vegas, 2003, pages 346–354.

24. Robert A. van Engelen, David Whalley, and Xin Yuan, *Validation of Code-Improving Transformations for Embedded Systems*, in proceedings of the 8th ACM Symposium on Applied Computing (SAC), 2003, March 9-12, 2003, Melbourne Florida, pages 684-691, (25% acceptance ratio).

2.1.3 Workshop Publications Citing DOE Support (Reviewed)

The reviewed journal publications produced by the project are listed below. These publications acknowledge DOE grant support.

25. Robert van Engelen, Wei Zhang, and Madhusudhan Govindaraju, *Toward Remote Object Coherence with Compiled Object Serialization for Distributed Computing with XML Web Services*, in the proceedings of Compilers for Parallel Computing (CPC), 2006, pages 441–455.

26. M. Govindaraju, A. Slominski, K. Chiu, P. Liu, R. van Engelen, and M. Lewis, *Toward Characterizing the Performance of SOAP Toolkits*, in the proceedings of the 5th IEEE/ACM International Workshop on Grid Computing, pp. 365-372, Pittsburgh, USA, 2004, (33% acceptance ratio).

27. Johnnie Birch, Robert van Engelen, and Kyle Gallivan, *Conditional Value Range Analysis for Data Dependence Testing*, in the proceedings of Compilers for Parallel Computing (CPC) 2004, pages 265–276.

28. Johnnie Birch, Robert van Engelen, and Kyle Gallivan, *Array Data Dependence Testing with the Chains of Recurrences Algebra*, in the proceedings of the IEEE International Workshop on Innovative Architecture (IWIA), January 2004, pages 71–81.

29. Burt Walsh, Robert van Engelen, Kyle Gallivan, Johnnie Birch, and Yixin Shou, *Parametric Intra-Task Dynamic Voltage Scheduling*, in proceedings of the PACT Workshop on Compilers and Operating Systems for Low Power, 2003, (CD-ROM proceedings, 6 pages).

30. Robert A. van Engelen, Kyle Gallivan, and Burt Walsh, *Tight Timing Estimation With the Newton-Gregory Formulae*, in proceedings of Compilers for Parallel Computers (CPC) 2003, January 8–10, 2003, Amsterdam, the Netherlands, pages 321–330.

2.1.4 Other Publications (Not Reviewed)

The non-reviewed publications and reports produced by the project are listed below.

31. Madhusudhan Govindaraju, Michael Lewis, Kenneth Chiu, Robert van Engelen, Sam Lang, Keith Jackson, *Web Services Performance Aspects*, (abstract) in proceedings of Globus-WORLD conference, Boston, February 2005.

32. Robert van Engelen, *gSOAP and Web Services*, C/C++ Users Journal, Vol. 23, No. 2, February 2005, pages 7–13.

33. Robert van Engelen, *gSOAP User Guide*. Accompanies the gSOAP software distribution package available from SourceForge, (approx. 240 pages), 2003–2005.

34. Robert van Engelen, *The CR# Algebra and its Application in Loop Analysis and Optimization*, Technical Report TR-041223, December 2004.

35. Robert van Engelen, Johnnie Birch, Yixin Shou, and Kyle Gallivan, *Array Data Dependence Testing with the Chains of Recurrences Algebra*, Technical Report TR-041201, December 2004.

2.1.5 Talks and Addresses

A selection of the presentations delivered during the project are listed below.

- Invited presentation *Array Dependence Analysis with the Chains of Recurrences Framework for Loop Optimization*, North Carolina State University, February 24, 2006.

- Invited presentation *Array Dependence Analysis and Vectorization with the Chains of Recurrences Framework*, Intel Corp, Santa Clara, September 12, 2005

- Invited presentation *gSOAP Web Services*, Microsoft, Redmond, September 21, 2005.

- Paper presentation *Toward Remote Object Coherence with Compiled Object Serialization for Distributed Computing with XML Web Services*, Compilers for Parallel Computing (CPC), January 2006, A Coruna, Spain.

- Paper co-presentation *Benchmark Suite for SOAP-based Communication in Grid Web Services*, Supercomputing Conference, Seattle, November 2005.

- Presentation and panel session *Web Services Performance Aspects*, GlobusWORLD conference, Boston, February 2005.

- Paper presentation *Conditional Value Range Analysis for Data Dependence Testing*, Compilers for Parallel Computing (CPC), July 2004, Seeon, Germany.

- Paper presentation *A Unified Framework for Nonlinear Dependence Testing and Symbolic Analysis*, ICS Conference, June 2004, Saint Malo, France.

- Paper presentation *An Integrated Dependence Analysis Framework based on the Chains of Recurrences Algebra*, IWIA workshop, January 11, 2004, Maui, Hawaii.

- Paper presentation, *Code Generation Techniques for Developing Light-Weight XML Web Services for Embedded Devices*, ACM Symposium on Applied Computing (SAC), Nicosia, Cyprus, March 16, 2004.

- Paper presentation *Pushing the SOAP Envelope with Web Services for Scientific Computing* at the International Conference on Web Services (ICWS) 2003, Las Vegas, USA, June 26, 2003.

- Paper presentation *Validation of Code-Improving Transformations for Embedded Systems* at ACM Symposium on Applied Computing (SAC) 2003, Melbourne FL, USA, March 10, 2003.

6

- Paper presentation *Tight Timing Estimation With the Newton-Gregory Formulae* at CPC 2003, Amsterdam, the Netherlands, January 10, 2003.

- Paper presentation *The gSOAP Toolkit for Web Services and Peer-To-Peer Computing Networks* at the IEEE International Symposium on Cluster Computing and the Grid (CCGrid2002) 2002, Berlin, Germany, May 21-24.

- Research seminar at Argonne National Labs on joint gSOAP and OGSA efforts, October 7, 2002.

2.1.6 Conferences Organized

The conferences, conference tracks, and workshops organized by the PI are listed below.

- Track chair for the ACM Symposium on Applied Computing (SAC) 2006 track on *Distributed Systems and Grid Computing*.

- Co-track chair for the GlobusWorld conference 2005 on Web Services Performance Aspects.

- Track chair for the ACM Symposium on Applied Computing (SAC) 2005 track on *Distributed Systems and Grid Computing*.

- Track chair for the ACM Symposium on Applied Computing (SAC) 2004 track on *Parallel and Distributed Systems*.

- Conference registration chair for the ACM *International Conference on Supercomputing* (ICS) 2003.

- Track chair for the ACM Symposium on Applied Computing (SAC) 2003 track on *Parallel and Distributed Systems and Networks*.

The PI served on the following program committees

- IEEE International Symposium on High Performance Distributed Computing (HPDC) program committee member, 2007.

- IEEE International Conference on Services Computing (SCC) program committee member, 2007.

- International Conference on Parallel Processing Workshop on High Performance Scientific and Engineering Computing (ICPP-HEPSEC) program committee member, 2006.

- ACM International Conference on Supercomputing (ICS) program committee member, 2003.

2.1.7 Web Sites Developed and Maintained

The Web sites and Web pages developed for the project:

- Chains of recurrence algebra package for Polaris compiler. This package provides an induction variable recognition algorithm and dependence analyzer based on the CR algebra. Demo and software is available from `http://people.scs.fsu.edu/~birch/research`.

- gSOAP toolkit web site: `http://gsoap2.sourceforge.net` (approx 30 pages). The gSOAP toolkit is an open source project and offers compilers, web servers, and tools to develop efficient SOAP/XML web services in C and C++. The software package has been downloaded over 150,000 times since 2003.

2.2 Overview of Project Expenditures and Justification

This section provides an overview of the expenditures. Justification is provided.

2.2.1 Personel

Personel appointed on the grant during the project:

- The PI was appointed on the grant for 1 month of summer salary in 2003.

- The PI was appointed on the grant for 1 month of summer salary in 2004.

- The PI was appointed on the grant for 1 month of summer salary in 2005.

- The PI was appointed on the grant for 2 months of summer salary in 2006.

- Johnnie Birch, Research Assistant, PhD candidate. Johnnie pursues a PhD in the area of restructuring compilers for scientific computing. He is expected the graduate in 2007.

- Yixin Shou, Research Assistant, PhD candidate. Yixin pursues a PhD in the area of flow-sensitive dependence analysis for compilers. Seh is expected to graduate in 2008.

- Subhajit Datta, Research Assistant, MS student (continues with PhD). Subhajit developed software engineering metrics that can benefit scientific application development by analyzing the impact of requirement changes on software. The metrics are obtained by compiler analysis methods developed for the project.

- Arthi Gokarn, Research Assistant, MS student. Arthi provides essential software development support for the gSOAP toolkit for Web and Grid services as mentioned in the grant proposal. Her accomplishments include a new toolkit for Grid computing with MATLAB. She graduated with a MS in 2004.

2.2.2 Equipment

Small equipment purchased for the project:

- Apple Macintosh Titanium Powerbook.

2.2.3 Travel

Travel sponsored by the project during the project year to deliver presentations related to the project:

1. North Carolina State University, colloquium talk, February 24, 2006.

2. CPC Workshop, January 2006, A Coruna, Spain.

3. GlobusWORLD conference, Boston, February 2005.

4. ICS Conference 2004, St. Malo, France, June 26-July 1.

5. CPC Workshop 2004, Germany, July 7-9.

6. IWIA workshop, January 11, 2004, Maui, Hawaii.

7. ACM Symposium on Applied Computing (SAC), Nicosia, Cyprus, March 16, 2004.

3 Summary

The published research results during the project year are summarized below.

1. Published 7 journal papers acknowledging grant support, and 17 conference and 6 workshop papers acknowledging grant support

2. Presented several talks at conferences and workshops.

3. Developed innovative compiler analysis methods for recurrences, induction variables, and array dependences. The work is implemented in Polaris and made available to the public. The widely-used GCC compiler uses a simplified version adapted from our work.

4. Implemented the gSOAP toolkit for efficient Grid computing with Web services. The gSOAP software is developed by the PI and is used by several DOE-sponsored projects related to Grid computing, such as Globus GT3 and Harness. The PI established an active collaboration with the Globus team Argonne National Laboratories (contact: Kate Keahey). The gSOAP software is also used by many companies such as IBM, Adobe Systems, AOL, eBay, HP, Xerox, WindRiver, OpenWave, and Siemens. gSOAP is available from http://gsoap2.sourceforge.net it has been downloaded $> 150,000$ times since the initial release in 2003). The book "Professional Open Source Web Services" devotes a chapter on gSOAP. Reviews of the PI's software can be found on developer's Web sites such as Dr.Dobbs.

5. Completed work on validating compiler transformations in collaboration with Professors D. Whalley and X. Yuan (originally funded by an NSF grant).

6. Completed work on a new type of compiler transformation to merge branches in collaboration with Professors D. Whalley and X. Yuan.

7. Investigated timing analysis methods and designed new algorithms based on Newton-Gregory interpolating polynomials.

8. Investigated and implemented improvements in gSOAP for embedded systems.

9. Developed new parametric algorithms for intra-task voltage scaling.

10. Investigated and implemented secure layer for Grid computing with gSOAP. In this work we developed a Globus secure infrastructure (GSI) plugin for gSOAP for Grid systems in collaboration with M. Cafaro (member of the European Sara project, GRB project, GridLab project).

11. Investigated and designed new fast algorithms for scientific computing and networking with Web services.

12. Worked with IBM on a port of the gSOAP software to develop the IBM Web services toolkit for mobile devices (WSTKMD) for embedded systems, cell phones, and PDAs (software available from IBM AlphaWorks).

13. Researched and co-developed Monte-Carlo Grid algorithms and application infrastructure.

4   Contact Information

Robert van Engelen
Computer Science Department
Florida State University
253 James J. Love building
Tallahassee, FL 32306-4530
phone: 850 645 0309
fax: 850 644 0058
engelen@cs.fsu.edu
http://www.cs.fsu.edu/~engelen