

The Next Generation Software Workshop – IPDPS'07

Frederica Darema
Computer and Information Science and Engineering Directorate
National Science Foundation
Arlington, VA 22230
(darema@nsf.gov)

Abstract

This workshop provides a forum for an overview, project presentations, and discussion of the research fostered and funded initially by the NSF Next Generation Software (NGS) Program, and more recently through the Advanced Execution Systems (AES) and the Systems Modeling and Analysis (SMA) components of the follow-up Computer Systems Research (CSR) Program¹. The present workshop is part of the series of the Next Generation Software Workshops that have been conducted since 2001, on a yearly basis and in conjunction with IPDPS. The topics addressed in the workshop are on research in systems' software technology in the scope of NGS and of the AES and the SMA components, namely: methods for systems modeling and analysis, performance engineering, dependability, fault tolerance, and power-management, programming environments, enhanced compiler capabilities, tools for the development, dynamic runtime support and dynamic composition of complex applications executing on heterogeneous, parallel and distributed computing platform assemblies, such as computational grids, encompassing, advanced multiprocessing nodes, mid-range and high-end platforms, clusters, embedded and sensor systems, and special purpose processing systems.

1 Introduction

Research on novel directions in systems software in the areas of computing systems performance engineering methods and dynamic runtime

¹ The NSF Next Generation Software (NGS) Program was developed and managed (1998-2004) by F. Darema. This Program Announcement and the list of current awards can be found at: <http://www.nsf.gov/cgi-bin/getpub?nsf00134>, and under the Next Generation Software (NGS) Program in http://www.nsf.gov/home/grants_awards.htm. The two components (TPES and CADSS) of the NGS program have been incorporated in the subsequently announced Computer Systems Research (CSR) Program (NSF04-609, NSF05-629, and NSF07-504) (http://www.nsf.gov/funding/pgm_summ.jsp?pims_id=13385&org=CNS) respectively as the SMA and AES components.

environments has been fostered through the Next Generation Software (NGS) Program (1998-2004), and subsequently through the Advanced Execution Systems (AES) and the Systems Modeling and Analysis (SMA) components of the successor program, the Computer Systems Research (CSR) Program (2004-todate). The majority of the projects supported within the scope of NGS, AES and SMA, involve multidisciplinary research, spanning several computer sciences sub-areas. In addition there is emphasis for system-level, multilevel and multimodal approaches, encompassing modeling and measurements, and integration there-of, as well as “phenomenological” approaches, and emphasis on driving and validating the systems software advances with end-user applications the advances in computer sciences technologies.

The SMA component is aimed to enable systematic modeling and analysis methods to assist the design and management of computing systems hardware and software, and support of dynamic runtime applications environments. To enable that, the program supports development of systematic methodologies and tools for analysis and prediction of the performance of applications, of hardware platforms, and of system software. The emphasis placed in the sought methodologies is a system-level approach, so that analysis and prediction of system behavior and performance is not based on isolated components or layers but in a coherent combination of modeling, simulation, emulation of all layers and components. The approach fostered by the SMA component is to consider the computing system in terms of an architectural framework consisting of: the applications, the system software and the underlying hardware layers. Key elements of the methodologies emphasized by the SMA component, include development of multi-level and multi-modal methods and tools for describing the application software, the system software, and the system hardware. Such models and tools encompass modeling and simulation of components at multiple levels of detail and abstraction, as well as incorporation of performance measurements. Additional key capability is the ability to combine these multi-

level/multi-modal methods and tools into “performance frameworks” (in a “plug-and-play” fashion) as needed for understanding, analysis, and prediction of behavior and performance of individual components or layers as well as the system as a whole. Thus the methodology fostered by the SMA component enables component-level and system-level performance analysis and prediction.

The AES component of the program fosters novel research in two key technology areas: new compiling technology and new application composition technology, which will enable applications to effectively execute under dynamic runtime resource availability as is manifested in complex and heterogeneous grid computing environments. In the new compiling system technology advocated under AES, part of the compiler gets embedded in the runtime, and interacts with the underlying system resource managers as well as measurements and performance descriptors of the applications and the underlying hardware and software systems. This kind of new compiling system (the Runtime Compiling System – RCS) will have the ability to adaptively optimize the mapping of applications on the underlying dynamic platform assembly. In conjunction of imparting to the RCS technology the desired capabilities, the AES component also fosters the development of novel programming models, application interface, and debugging and check-pointing technologies.

In addition the program advocates the need for new technology for knowledge-based, RCS-invoked dynamic assembly of application components capabilities, to optimize composition of an application depending on the underlying resources, and supports research along these directions. The AES component also supports integration of these technologies into application support environments, and demonstration of the technologies on important production-class applications.

Through the most recent CSR Program solicitation an additional thematic area has been included (Cross-Systems Integration – CSI, in NSF07-504) placing emphasis in advances in systems software supporting dynamic integration of highly heterogeneous and environments, ranging from real-time data acquisition systems (such as sensors and other instruments) to the high-end computational grids, and driven by dynamic

applications, such as those encountered in Dynamic Data Driven Applications Systems^{2,3} environments.

2 Workshop Agenda

The presentations in the 2-day NGS Workshop at IPDPS2007, on March 25-26, 2007, provide a sample of the kind of research projects that are advancing technology along the directions envisioned in NGS Program, and the AES and SMA components of the CSR Program. The titles and authors of each presentation are listed below; the name underlined is that of the Principal Investigator(s) of the sponsored project. The Workshop includes in addition Panel Sessions to discuss overarching issues, broader impact, education issues, technology transfer, and potential new and additional future research directions.

Agenda for the NGS Workshop at IPDPS2007

Sunday, March 25, 2007

Session 1: 8:30am-10:30am

Introduction to the Workshop – Frederica Darema

ParalleX: A Study of A New Parallel Computation Model

Guang R. Gao, Thomas Sterling, Rick Stevens, Mark Hereld, and Weirong Zhu

Improving MPI Independent Write Performance Using A Two-Stage Write-Behind Buffering Method
Wei-keng Liao, Avery Ching, Kenin Coloma, and Alok Choudhary

Automatic MPI application transformation with ASPHALT

Anthony Danalis, Lori Pollock, and Martin Swamy

Formal Analysis for Debugging and Performance Optimization of MPI

Ganesh L. Gopalakrishnan and Robert M. Kirby

Automatic Parallelization of Scripting Languages: Toward Transparent Desktop Parallel Computing

² F. Darema: DDDAS - Dynamic Data Driven Applications Systems, Proceedings of the ICCS'03, June 2003; Grid Computing and Beyond: The Context of Dynamic Data Driven Applications Systems, Proceedings of the IEEE, V.93, 3, March 2003.

³ www.cise.nsf.gov/dddas and www.dddas.org

Xiaosong Ma, Jiangtian Li, and Nagiza F. Samatova

Virtual Execution Environments: Support and Tools
Apala Guha, Jason D. Hiser, Naveen Kumar, Jing Yang,
Min Zhao, Shukang Zhou, Bruce R. Childers, Jack W.
Davidson, Kim Hazelwood, and Mary Lou Soffa

Session 2: 11:00am-12:30pm

Intelligent Optimization of Parallel and Distributed Applications

Bhupesh Bansal, Umit Catalyurek, Jacqueline Chame,
Chun Chen, Ewa Deelman, Yolanda Gil, Mary Hall,
Vijay Kumar, Tahsin Kurc, Kristina Lerman, Aiichiro
Nakano, Yoon-ju Lee Nelson, Joel Saltz, Ashish
Sharma, Priya Vashishta

Scheduling issues in optimistic parallelization
Milind Kulkarni and Keshav Pingali

New Results on the Performance Effects of Autocorrelated Flows in Systems

Evgenia Smirni, Qi Zhang, Ningfang Mi, Alma Riska,
and Giuliano Casale

The Adaptive Code Kitchen: Flexible Tools for Dynamic Application Composition

Pilsung Kang, Mike Heffner, Joy Mukherjee,
Naren Ramakrishnan, Srinidhi Varadarajan, Cal
Ribbens, and Danesh K. Tafti

DOSA: Design Optimizer for Scientific Applications
David A. Bader and Viktor K. Prasanna

The TMO Scheme for Wide-Area Distributed Real-Time Computing and Distributed Time-Triggered Simulation

K. H. (Kane) Kim and Stephen W. Jenks

Session 3: 1:30am-3:30pm

Adaptive Scheduling with Parallelism Feedback

Kunal Agrawal, Yuxiong He, Wen-Jing Hsu, and
Charles E. Leiserson

Weaving Atomicity Through Dynamic Dependence Tracking

Suresh Jagananthan

A Key-based Adaptive Transactional Memory Executor

Tongxin Bai, Xipeng Shen, Chengliang Zhang, William
N. Scherer, Chen Ding, and Michael L. Scott

Optimizing Sorting with Machine Learning Algorithms

Xiaoming Li, Mar'ia Jes 'us Garzar'an, and David
Padua

Knowledge and Cache Conscious Algorithm Design and Systems Support for Data Mining Algorithms

Amol Ghoting, Gregory Buehrer, Matthew Goyder,
Shirish Tatikonda, Xi Zhang, Srinivasan Parthasarathy,
Tahsin Kurc, and Joel Saltz

Memory Optimizations For Fast Power-Aware Sparse Computations

Konrad Malkowski, Padma Raghavan and Mary Jane
Irwin

A global address space framework for locality aware scheduling of block-sparse computations

Sriram Krishnamoorthy, Umit Catalyurek, Jarek
Nieplocha, Atanas Rountev, and P. Sadayappan

Speedup using Flowpaths for a Finite Difference Solution of a 3D Parabolic PDE

Darrin M. Hanna, Anna M. Spagnuolo, and Michael
DuChene

Session 4: 4:00am-6:00pm

NGS: Service Adaptation in Open Grid Platforms

Krishnaveni Budati, Jinh Kim, Abhishek Chandra, and
Jon Weissman

Creating a Robust Desktop Grid using Peer-to-Peer Services

Jik-Soo Kim, Beomseok Nam, Michael Marsh, Peter
Keleher, Bobby Bhattacharjee, Derek Richardson,
Dennis Wellnitz, and Alan Sussman

Locality-aware Buffer Management: Algorithms Design and Systems Implementation for Data Intensive Applications - A Brief Progress Report

Xiaodong Zhang

Designing Efficient Systems Services and Primitives for Next-Generation Data-Centers

K. Vaidyanathan, S. Naravula, P. Balaji, and D. K.
Panda

Supporting Quality of Service in High-Performance Servers

Yan Solihin, Fei Guo, Seongbeom Kim, and Fang Liu

Enhancing Energy Efficiency in Multi-tier Web Server Clusters via Prioritization

Tibor Horvath, Kevin Skadron, and Tarek Abdelzaher

Autonomic Power and Performance Management for Large Scale Data Centers

Bithika Khargharia, Salim Hariri, Ferenc Szidarovszky,
Hesham El-Rewini, Manal Hourri, Samee Khan, Ishfaq
Ahmad, and Mazin S. Yousif

Improving Data Access Performance with Server Push Architecture

Xian-He Sun, Surendra Byna, and Yong Chen

Session 5: 7:30pm-9:00pm

Panel Discussion: Programmatic Directions

Monday, March 26, 2007

Session 6: 8:30am-10:00pm

SimX meets SCIRun: A Component-based Implementation of a Computational Study System
Siu-Man Yau, Eitan Grinspun, Vijay Karamcheti, and Denis Zorin

VIPProf: Vertically Integrated Full-System Performance Profiler
Hussam Mousa, Chandra Krintz, Lamia Youseff, and Rich Wolski

Model Predictive Control for Memory Profiling
Sean Callanan, Radu Grosu, Justin Seyster, Scott A. Smolka, and Erez Zadok

Understanding Measurement Perturbation in Trace-based Data
Todd Mytkowicz, Amer Diwan, Matthias Hauswirth, and Peter F. Sweeney

PROTOFLEX: FPGA-accelerated Hybrid Functional Simulator
Eric S. Chung, Eriko Nurvitadhi, James C. Hoe, Babak Falsafi, and Ken Mai

Models and Heuristics for Robust Resource Allocation in Parallel and Distributed Computing Systems
David L. Janovy, Jay Smith, Howard Jay Siegel, and Anthony A. Maciejewski

Session 7: 10:30am-12:30pm

Model-Driven Performance Analysis Methodology for Distributed Software Systems
Swapna S. Gokhale, Aniruddha Gokhale, Jeff Gray, Paul Vandal, Dimple Kaul, Arundhati Kogekar, and Yuehua Lin

J-Sim: An Integrated Environment for Simulation and Model Checking of Network Protocols
Ahmed Sobeih, Mahesh Viswanathan, Darko Marinov, and Jennifer C. Hou

Early Results with Precision Abstraction: Using Data-flow Analysis to Improve the Scalability of Model Checking
Adam Brown, James C. Browne, and Calvin Lin

Static Verification of Design Constraints and Software Correctness Properties in the Hob System

Patrick Lam, and Martin Rinard

ExPert: Dynamic Analysis Based Fault Location via Execution Perturbations
Neelam Gupta and Rajiv Gupta

An Analysis of Availability Distributions in Condor
Rich Wolski, Daniel Nurmi Kim, John Brevik

Identifying and Addressing Uncertainty in Architecture-Level Software Reliability Modeling
Leslie Cheung, Leana Golubchik, Nenad Medvidovic, Gaurav Sukhatme

A Markov Reward Model for Software Reliability
YoungMin Kwon and Gul Agha

Session 8: 1:30pm-3:00pm

Modeling Modern Micro-architectures using CASL
Edward K. Walters II, J. Eliot B. Moss, Trek Palmer, Timothy Richards, and Charles C. Weems

Rethinking Automated Synthesis of MPSoC Architectures
Brett H. Meyer and Donald E. Thomas

A Reconfigurable Chip Multiprocessor Architecture to Accommodate Software Diversity
Engin Ipek, Meyrem Kirman, Nevin Kirman, and Jos'e F. Mart'inez

Scalable, Dynamic Analysis and Visualization for Genomic Datasets
Kai Li, Matthew Hibbs, Grant Wallace Maitreya Dunham, Rachel Sealfon, and Olga Troyanskaya

A Distributed Execution Environment for Large Data Visualization
Jian Huang, Huadong Liu, Micah Beck, Jinzhu Gao, and Terry Moore

Annotation Integration and Trade-off Analysis for Multimedia Applications
Radu Cornea, Alex Nicolau, and Nikil Dutt

Session 10: 3:30pm-5:00pm

Panel Discussion: Research, Technology and Education Drivers and Opportunities (Industry Representatives)

Session 9: 5:00am-5:30pm

Closing Discussion

Adjournment: 5:30pm