Workshop on Cutting Edge Computing

In conjunction with the International Conference on High Performance Computing (HiPC 2005), Goa, India [http://www.hipc.org/hipc2005/index.html]

Date: December 21, 2005
Time: 9:00AM to 1:00PM

The half-day workshop will feature invited presentations from experts in such areas as Quantum Computing, Autonomic Computing, Datamining of Scientific Data, etc. The goal is to bring to the attention of researchers and practitioners, recent R&D advances in computing technologies that are expected to dominate the future of computing.

Workshop Program

1) "X10: An Object-Oriented Approach to Non-Uniform Cluster Computing" - Vivek Sarkar, IBM T.J. Watson Research Center

ABSTRACT:
A popular emerging multiprocessor structure for future low-end and high-end systems is a Non-Uniform Cluster Computing (NUCC) configuration with nodes that are built out of multi-core SMP chips with non-uniform memory hierarchies, and interconnected in scalable cluster configurations. Unlike previous generations of hardware evolution, this shift will have a major impact on existing software. We have designed a modern object-oriented programming language, X10, for high performance, high productivity programming of NUCC systems, using the Java(TM) language as the basis for its serial subset. A member of the Partitioned Global Address Space (PGAS) family of languages, X10 highlights the explicit reification of locality in the form of places; lightweight activities embodied in async, future, foreach, and ateach constructs; a construct for termination detection construct (finish); the use of lock-free synchronization (atomic blocks); and the manipulation of global arrays and data structures. We present an overview of the X10 programming model and language, and discuss its suitability as a foundation for scalable parallelism and high-productivity programming tools.

The current members of the X10 design and implementation team are Philippe Charles, Christopher Donawa, Kemal Ebcioglu, Christian Grothoff, Allan Kielstra, Christoph von Praun, Vijay Saraswat and Vivek Sarkar.

BIO:
Dr. Vivek Sarkar is Senior Manager of the Programming Technologies Department in the IBM Research Division, and is a member of the IBM Academy of Technology. The projects under way in his department span the areas of 1) Programming Models --- PERCS/X10, XJ, Symbiotic, 2) Programming Tools --- Advanced Refactoring, MPI Tools, SAFE (Scalable And Flexible Error detection), and Legacy Transformation (Mastery), 3) Deployment, Optimization and Execution --- PDS (Progressive Deployment System), Metronome, and Jikes RVM. Dr. Sarkar is also responsible for coordinating Eclipse-related projects and activities in IBM Research, including the world-wide Eclipse Innovation Grants program.

Dr. Sarkar's personal research interest is in advancing the state of the art of programming language design, programming tools, and optimizing and parallelizing compilers to deliver improved productivity and performance benefits to users of high-performance systems. He is a co-leader of the PERCS Programming Model and Tools project, which includes work on X10. In the past, he has worked on several compiler-related projects at IBM including the PTRAN and Jikes RVM research projects at Watson, and high-order transformations and OpenMP parallelization in the IBM XL Fortran product compilers. Dr. Sarkar has published over 75 papers in the area of optimizing and parallelizing compilers. Dr. Sarkar joined IBM in 1987, after obtaining a Ph.D. from Stanford University. In 1997, he was on sabbatical as a visiting associate professor at MIT.

2) "Client Technologies: A Peek into the Future" - Apratim Purakayastha, IBM T J. Watson Research Center

ABSTRACT:
This paper will summarize some key trends in upcoming generation of clients, technological and social implications of those trends, and challenges and opportunities that lie ahead. The first major trend is the increasing diversity of devices and users. Some kind of "computing infrastructure" is being accessed by more and more diverse devices including traditional desktops and laptops, personal digital assistants, mobile phones, digital cameras, digital music players, digital televisions and phones, and smart unattended devices like radio-frequency readers and tags. The user population is also becoming increasingly diverse across geographies, age groups, and the social roles they play. In the backdrop of this diversity, the infrastructure is becoming increasingly connected. It is now reasonable to expect some form of network connection at almost all times. Perhaps equally important is the emergence of the IP protocol as a global interoperable standard. Very soon, IP will form one common bridge for applications running on diverse platforms such as computers, televisions, phones, and cameras.

The above trends have some technological and social implications. Rampant heterogeneity of devices and users will reduce the dominance of incumbents over time. For example, it is likely that desktops/laptops, televisions, cameras, and mobile phones will not all run one "operating system". They will not all evolve to be generic computing devices but specialized computing platforms that will need to work with each other to provide the overall user experience. This implies that there will be a strong focus on open interoperable standards that accommodate diverse end points. The new generation of users growing up with mobile phones will also demand "constant connect" with their social networks. Email will take a backseat to more synchronous interactions like instant messaging. Unattended clients will be first-class citizens of the computing fabric, melding with user-facing devices. Notions of presence and programmatic manipulation of these "inert" entities will emerge as well.

The above trends and implications raise some serious technical challenges as well as substantial opportunities. With diversity of devices, it will become increasingly challenging to author applications that run on multiple devices. That obstacle, however, needs to be overcome to achieve economies of scale in development. One of the big challenges will be to realize a truly content-oriented or document-oriented liquid application model such that applications can be transformed easily. Such a model will require platforms to be abstracted as a set of services and applications essentially become a declarative composition of services. The notion of deployment needs to be

revolutionized where applications are never "deployed" per se but merely "cached" as and when necessary and updated seamlessly like a content or data cache. Security becomes of paramount importance in this highly diverse and highly connected world. New techniques need to be developed that guarantee provenance of data, and establish trusted virtual domains for groups of applications. Better connectivity will be an opportunity to develop a whole new class of "live" applications and related infrastructure that will give a whole new meaning to instant messaging.

BIO:
Apratim Purakayastha is currently a senior manager in the IBM Thomas J. Watson Research Center. Over the last few years he has engaged in research in the field of mobile computing, specifically in the areas of data synchronization, messaging, notification, and context. He is one of the founding contributors to the SyncML data synchronization standard and has contributed to multiple IBM products related to mobile and pervasive computing. He has authored numerous research papers in the field and has been named an IBM Master Inventor for his patents in this area. He joined IBM in 1996 after completing his Ph.D. in Computer Science from Duke University, USA. He can be reached via email at apu@us.ibm.com.

3) "GIPSY and Genome Grid: Web based Computing Environments for Bioinformatics" - Rajendra Joshi, Bioinformatics Team, Scientific and Engineering Computing Group, CDAC.

ABSTRACT:
The last decade has witnessed extraordinary advances in the fields of biotechnology and molecular biology, the Human Genome Project being one of them. Further, more than 282 genomes have been completely sequenced and genome sequencing of additional 1259 organisms are at various levels of completion. The enormous biological sequence data thus generated necessitates the development of efficient tools for mining the information on structural and functional properties of biomolecules. Such a kind of information can prove invaluable for pharmaceutical industries, for in silico drug target identification and new drug discovery. However, in order to exploit the enormous scientific value of this information for understanding biological systems, the information must be integrated, analyzed, graphically displayed and ultimately modeled computationally. The enormity of data and complexity of algorithms make the above tasks computationally demanding, necessitating the use of computational resources beyond those available to researchers at a single location. Thus, grid computing offers an ideal platform for a biologist to analyze the data by using the computing resources and also databases available at different locations.

This paper describes the development and implementation of GIPSY, a Problem Solving Environment and Genome Grid, a Grid portal for bioinformatics applications. The GIPSY implementation design was done on basis of Model-View-Controller (MVC) architectural design recommended for interactive applications using J2EE technologies. The GIPSY was further grid enabled to give the Genome Grid, which uses Java CoG toolkit to integrate seamlessly with the Globus toolkit to bring the complete power of the compute grid onto a desktop for running an application. Some of the features include, an automated system to take user's inputs via a web page, session tracking of users across the grid, building job execution scripts for the grid environment, submission and control of the jobs and conveying the result back to the user, and aiding the user in analysis and visualization of the results. In addition, the paper also discusses about some of the

complex biological problems such as protein folding and motif finding, which can be addressed using high performance computing.

BIO:
Dr. Rajendra Joshi, a doctorate in Biochemistry from National Chemical Laboratory (University of Pune), has been associated with the area of Biotechnology & Bioinformatics for about 15 years. He has over 12 years of experience in the area of Bioinformatics, which includes experience as a researcher, faculty member at the Bioinformatics Centre, University of Pune and project manager. He is primarily responsible for building a strong bioinformatics team at C-DAC and presently directs projects taken up by the team. His major area of expertise, is in the use of high performance parallel computers for biological research. His unique strength is in the form of good knowledge of biology and parallel computing. His main research interests include, molecular dynamics simulations of nucleic acids & proteins, genome sequence analysis, metabolic pathways and software development for Problem Solving Environments and Grid based applications. He is a principal investigator for funded projects and other research projects, leading to a number of publications in international journals and conferences. He is actively involved in collaborative programmes and national scientific meets. He is a member of the International Society of Computational Biology and has numerous publications, articles and invited talks to his credit. His career objective is to build a high throughput environment where much of biology could be understood using high performance parallel computers.

4) "HPC & Challenges in Contemporary Processor Design" – Pratap Pattnaik, IBM T.J. Watson Research Center

ABSTRACT:
The prediction of Gordon Moore in 1965, only six years after the integrated circuits were invented, that the exponential growth in the number of transistors in an IC will continue, has been true for past four decades. This success in continued miniaturization in the transistor size has contributed to

- Significant increase in the processor clock speed,
- Enormous opportunity to integrated various functions on the same IC as the processor,
- Increase in the power density,
- Increase in the power consumption due to leakage currents,
- Opportunity for the Software to primarily focus on developer productivity, rather than extracting the performance from the hardware, and
- A data centric society.

With the anticipated slow down in the growth rate of processor clock speeds, for the past few years, processor designers and developers have been extensively engaged in developing a variety of architectural approaches to support the continued IT infrastructure advancements of the past decades. This talk will describe the some of the current trends and challenges faced by today's micro architects particularly in the areas of Cache Coherence protocols, micro partitioning technologies, CMP, SMT, Reliability and Serviceability and integration with the assist/offload engines and their opportunity for HPC work loads. The talk will also give some perspective from our own experience with PowerPC systems (including systems like cell processors) over the past decade.

BIO:

Dr. Pratap Pattnaik is currently senior manager of the Scalable Systems group in IBM Research. Over the past ten years, he and his team have developed a number of key technologies for IBM's high-end eServers.

His research work includes the development and design of computer systems, including both hardware and operating systems. He has also worked in the fields of parallel algorithms for molecular dynamics, solutions of linear systems, communication subsystems, fault management subsystems, quantum Monte Carlo and high-temperature superconductivity. He has over ten years of research experience in various aspects of integrated circuit design and fabrication, silicon processing, and condensed matter theory.

Dr. Pattnaik received a Ph.D. in Physics from the Massachusetts Institute of Technology in 1980.

5) "Massively Parallel Systems: Headache or Panacea?" - Manish Gupta, IBM T.J. Watson Research Center

ABSTRACT:
Conventional wisdom says that it is best to use powerful processors (in moderate numbers) to build supercomputers. Seymour Cray famously remarked, "If you were plowing a field, which would you rather use, two strong oxen or 1024 chickens?" We will describe technological trends that are pushing us in certain architectural directions for building systems. We will present our experiences from the IBM Blue Gene project on pushing the limits of scalability, with the goal of enabling breakthrough science. We will share the rationale for our key system design choices. We will describe some of the outstanding challenges in programming and managing massively parallel systems. While the speaker's bias will be quite clear, we will let members of the audience make up their own mind on the following question, do massively parallel systems represent an unnecessary headache or a powerful tool to meet our ever-growing need for more computational power.

BIO:
Manish Gupta is a Research Staff Member and Senior Manager of the Emerging System Software department at the IBM T. J. Watson Research Center. He leads research on system software for massively parallel machines and on high performance compilers and runtimes for high end servers. His team's recent activities include developing system software for the Blue Gene supercomputer, compilers and runtimes for the PERCS system, and end-to-end performance optimizations for commercial workloads on the IBM POWER5 based systems.

Manish received a Ph.D. in Computer Science from the University of Illinois at Urbana-Champaign in 1992, and has worked with IBM since then. He has co-authored over 60 papers in refereed conferences and journals in the area of high performance compilers, parallel computing, Java Virtual Machine optimizations, and optimizing commercial middleware. He has served as a co-chair of the International Workshop on Languages and Compilers for Parallel Computing (LCPC), workshops on Java for High Performance Computing, and has served on the program committee for several conferences. He has received two Outstanding Technical Achievement Awards and three Invention Achievement Plateau awards at IBM.

For further information contact the organizers.

ORGANIZERS

Harish K Grama
Vice President, IBM Software Labs
Golden Enclave, Airport Road
Bangalore 560017, India
Tel: +91-80- 2509 4547
Fax: +91-80-2527 7374
hargrama@in.ibm.com

Albee Jhoney
Technology Incubation Center
IBM Software Labs,
Golden Enclave, Airport Road,
Bangalore 560017, India
Tel: +91-80-2526 2355 extn 2501
Fax: +91-80-2527 7374
albee.jhoney@in.ibm.com