Tutorial Proposal for INDICON 2014

Title: An Overview of High Performance Computing, Grid & Cloud Computing

One of the objectives of tutorial is to discuss current trends on High Performance Computing (HPC), Grid and Cloud Computing technologies from application, algorithm and architecture point of view. An overview of current *multi*-to-*many* core processor programming paradigms in distributed computing infrastructure (Grids & Cloud) will be discussed. The broad areas of this tutorial are discussed as below.

High Performance Computing: Today, parallelization of several applications on accelerators and coprocessors in a hybrid message passing environment is an active area of research. General-purpose graphical processing units (GPGPUs), coprocessors (Intel Xeon Phi coprocessors), and accelerators (Altera FPGAs) are employed as a HPC accelerators with hybrid programming models. This programming model is a combination of libraries and programming languages such as MPI, OpenMP, NVIDIA CUDA & NVIDIA OpenACC, OpenCL, FPGA and Xeon Phi. The recent Intel Xeon Phi coprocessors in a message passing cluster such as PARAM YUVA-II are equipped with compiler technology and wide vector processing units to achieve the best performance for application kernels using standard, parallel programming models and languages. The GPGPUs (NVIDIA K40) based on CUDA 6.x and AMD- Fire Pro Series S9000 based on OpenCL 2.0) are used in large scale message passing clusters.

- The tutorial is aimed to explain techniques to build expertise on message passing clusters with accelerators such as Intel Xeon Phi, GPGPUs (AMD, NVIDIA), and FPGAs is to measure power consumption and performance for applications.
- The tutorial covers an overview of High Performance Computing Platforms, focusing on different accelerators and programming paradigms required for applications.
- The trends on measuring performance per watt (PARAM YUVA, IBM Blue Gene Q, Clusters with NVIDIA GPUs & AMD GPUs), will be summarized in this tutorial.
- A brief summary on multi-level parallelism based on many core threads, threading models, conservation of global memory bandwidth, mathematical libraries and efficient utilization of SIMD instructions will be highlighted with examples
- The architectural features of host multi-core processors with GPUs or coprocessors including shared virtual memory, dynamic parallelism, compliers and vectorization will be covered to know more about energy efficiency and power-aware computing.
- The tutorial discusses a summary on efforts required to measure performance and power consumption in watts for application kernels & benchmarks as a part of Green Computing and Top-500 Supercomputing projects.
- A case study of performance of benchmarks on PARAM YUVA-II using FPGA processors and other clusters systems with NVIDIA GPUs and AMD GPUs using Hadoop MapReduce and MPI in BIG DATA frame work will be discussed.

Grid Computing: Grid Computing is a key technology which is widely used in all the computational sciences. Grid uses heterogeneous and geographically spread computational resources, shared across a virtual laboratories. TeraGrid and C-DAC's GARUDA are high performance computational grids that involve heterogeneous collections of cluster of computers that are geographically located at different sites and run different system software. The commonly used grid software infrastructure is Globus,

which is a software toolkit for building grid-enabled applications and services. GARUDA is India's National grid initiative bringing together academic, scientific and research communities for developing their data and compute intensive applications. GARUDA uses multi-gigabit National Knowledge Network (NKN), thereby creating tremendous impact on collaborative research between educational institutions and Research & Development (R&D) organizations. Indian Grid Certification Authority, GARUDA Job submission portal, GARUDA Compiler Service and GARUDA Short Live Certificate (SLCS) are supported. GARDUA has peak computing power of more than 0.5 petaflop computing systems including PARAM YUVA-II. GARUDA has more than 40 collaborative partners and Virtual User Community (VoS) in scientific areas of research.

- In this tutorial, an overview of GARUDA that uses GLOBUS toolkit and grid middleware will be explained focusing on various grid services.
- An overview of GARUDA grid which is an aggregation of resources comprising of computational nodes, mass storage and scientific instruments distributed across the different geographical locations will be explained.
- The Service Oriented Architecture (SOA) based cyber infrastructure will be discussed.
- An overview of several applications which include bioinformatics-protein structure prediction on grid, OSDD chemo-informatics, initiatives on cancer grid, and collaborative class-room technologies will be discussed.
- An overview of GARUDA programming environment and benchmarks will be covered.

Cloud Computing: The concept of Cloud computing is an extension of Service-Oriented Computing, and refers to delivering hosted services over the Internet. The need of the hour is to help scientists to tackle the applications that processes huge amount of data over long time periods and resulting in equally large amount of data.

- In this tutorial, an overview of Cloud Computing infrastructure and programming paradigm will be discussed
- The tutorial is aimed to explain an overview of C-DAC's Scientific Cloud which include services such as IaaS, PaaS, SaaS and StaaS for the Scientific Community.
- An Overview of on-demand services and open storage stack efforts will be discussed.

Scientific Clouds allow to handle scientific problems, that require high compute capability, massive storage, high bandwidth and low latency networking. The researchers can access shared cluster resources or acquiring expensive hardware (clusters) at their organization, using Scientific Clouds. C-DAC's Scientific Cloud is modeled for on-demand access to the shared pool of configurable HPC resources (e.g., Servers, Storage, Networks, Applications, Software, and Services) that can be easily provisioned as and when needed, Supporting MPI / Hadoop MapReduce clusters and specific libraries on-demand proves beneficiary in scientific environments.

SuMegha Cloud is a sophisticated cloud package to build private clouds. Built using open source components -Xen, Nimbus, Openstack swift, GlusterFS and in-house developed tools -Cloud portal, CloudVault, Job Submission portal, and Problem Solving environments. It also provides a set of Golden Images of different sizes like small, medium, and large with CentOS operating systems and parallel environments with MPI libraries and Hadoop MapReduce framework.

Biography of Speakers:

Dr. VCV.Rao : VCV.Rao is an Associate Director and Head of the High-Performance Computing - Frontier Technologies & Exploration (HPC-FTE) Group, at C-DAC, Pune, India. VCV.Rao specializes in implementation of parallel algorithms on emerging parallel processing platforms (Cluster of Multi-Core Processors with accelerator devices -GPUs & CPUs). His group works on performance of application and system benchmarks and implementation of distributed computing algorithms on Distributed Computing systems with coprocessors, accelerators, power aware computing, and Out-of-Core algorithms (BIG Data Analytics) for large data processing. Mr.VCV.Rao contributed to design, develop and deployment of C-DAC's PARAM Series from the year 1994 onwards. He is actively involved in proliferation of parallel processing technology through workshops in India and contributed to PARAM series at Premier institute Projects in India. Mr.VCV.Rao is associated with C-DAC since 1993 & VCV. Rao received his Ph.D degree in Mathematics in 1993 from IIT-Kanpur. Mr.VCV.Rao was a visiting faculty at the Dept. of Computer Science, University of Minnesota (UoM), Minneapolis, and Post-Doctoral fellow at Army High Performance Computing (AHPCRC), UoM during the year 1997-98. His email address is vcvrao@cdac.in

Dr. Subrata Chattopadhyay: He is currently Associate Director at C-DAC, Bangalore. He is also the Chief Investigator of Garuda – the national grid computing initiative of India. Previously he was involved in setting up the PARAM Padma, the first Indian supercomputing facility listed from India. He was also involved in setting up of nationwide high speed communication fabric of GARUDA and deploying grid middleware across various platforms of supercomputers. He has has considerable experience on developing and managing e-Infrastructures that include High Performance Computing, Grid and Cloud technologies. From C-DAC, he was the technical manager for the EUIndia Grid project that interconnect Indian grid project – Garuda with the European grid initiatives – EGI. He is also leading another International project funded by European Commission entitled Co-ordination and Harmonization of Advanced E-Infrastructure - Research and Education Data Sharing (CHAIN-REDS). He did his Bachelors in Engineering (BE) degree from NIT, Durgapur, Masters (M.Tech) from IIT, Kanpur and Doctorate (Ph.D.) from University of British Columbia, Vancouver, Canada. He brings more than 27 years of experience both from IT industry and research organizations. His major areas of interest include HPC, grid/cloud computing and process modeling and simulations. His email address is subratac@cdac.in

Dr. B.B. Prahlada Rao: Prahlada Rao is a Joint Director –SSDH (System Software Development for HPC Grid and Cloud), C-DAC, Bangalore. He works on C-DAC HPC projects such as Cloud Computing, e-infrastructures, HPC Clusters and Grid Computing, Genetic and Evolutionary Computing, Algorithms Development in EDA/VLSI and Industrial Innovations, and Academic Research. He held different prestigious and responsible key positions in several organizations in India. He served as a Advisory R&D Engineer & Deputy General Manager (Software) at IBM Global Services India Pvt Ltd, Bangalore (2000-04), Group Leader at ST Microelectonics, Noida (1998-00), Technical Manager at Agere Systems India Pvt Ltd, Bangalore (2004-05), Team leader at Mentor Graphics, Hyderabad (1997-98), Deputy Design Engineer at HAL, Hyderabad (1982-90) and served on Deputation at ADA, Bangalore (1986-90). He also served as a Nominated member for the Technical Experts Council IBM Academy of Technology during 2003-04, Task force Member for P-CMM Core Team at IBM, and Task force Member for FPGA, SW-CMM, ISO at STMicro-elctronics on Soft Computing. He served as Faculty for MS (VLSI) Program of BITS, Pilani during 2002-2007. He did his Ph.D (1995), Computer Sc. & Automation, Indian Institute of Science, Bangalore. M.Tech (1981), Control Systems Engineering, IIT-Kharagpur, and B.Tech (1979), Electrical Engineering, JNT University, Kakinada, India. He published more than 50 Conference and journal papers with one US Patent holder and submitted other Indian Patents. His email address is prahladab@cdac.in