FEBRUARY 2020 VOL. 1 ISSUE 7

COMPUTATIONAL AND DATA SCIENCES

THE OFFICIAL NEWSLETTER OF IISc BANGALORE, CDS DEPARTMENT

cds.iisc.ac.in

The Department of Computational and Data Science presents the seventh issue of the first volume of the Newsletter series! Through this newsletter, we hope to give you a glimpse into our little world- what we do, where our interests lie and what we hope to achieve. The Department of Computational and Data Sciences (CDS) is an interdisciplinary engineering department covering the broad research areas of computational science and engineering, and scalable computer & data systems.

Our journey began in 1970, when the Computer Centre of the institute was established as a central computing facility. In 1990, it became Supercomputer Education and Research Centre (SERC), providing computing facility to the faculty and students of the Institute. A master's degree programme M.Tech in Computational Science, one of the first of its kind in India, was started by SERC in 1999. Therefore about 16 years later, CDS was formed in December 2015 from the academic wing of SERC.

From the Chair's Desk

It is our pleasure to share the activities of the department through this newsletter. The goal of this publication is to highlight our research activities, achievements of faculty, students and project staff and new initiatives.

The main motive of this news digest is to reach out to the world outside the campus including student , faculty, corporate and any individual who is interested in computational and data sciences. Ultimately, we hope to build a highly connected, a truly inter-disciplinary society and, of course, help us to network.

Research Highlights of CDS

Anwesha Bhowmik and Sathish Vadhiyar - HiPC 2019
Paper Title: HyDetect: A Hybrid CPU-GPU Algorithm for Community Detection

Community detection is an important problem that is widely applied for finding cluster patterns in brain, social,biological and many other kinds of networks. We propose a divide-and-conquer community detection algorithm for hybrid CPU-GPU systems. The graph representing a network is partitioned among the CPU and GPU devices of a node, and independent community detection using Louvain algorithm is carried out in both the parts. The communities are iteratively refined by a novel strategy for identifying and moving ”doubtful” vertices between the devices. The resulting accuracy is found comparable with the single device parallel Louvain algorithms.Our hybrid algorithm helped to explore large graphs that cannot be accommodated in a single device. By harnessing the power of GPUs, our hybrid algorithm is able to provide 42-73% smaller execution time compared to multicore Louvain algorithm.

Jogendra Nath Kundu, Siddharth Seth, Rahul M V, Mugalodi Rakesh, R. Venkatesh Babu, Anirban Chakraborty - AAAI 2020
Paper Title: Kinematic-Structure-Preserved Representation for Unsupervised 3D Human Pose Estimation

Estimation of 3D human pose from monocular images has gained considerable attention, as a key step to several human-centric applications. However, generalizability of human pose estimation models developed using supervision on large-scale in-studio datasets remains questionable, as these models often perform unsatisfactorily on unseen in-the-wild environments. We propose a novel kinematic-structure-preserved unsupervised 3D pose estimation framework, which is not restrained by any paired or unpaired weak supervisions. Our pose estimation framework relies on a minimal set of prior knowledge that defines the underlying kinematic 3D structure, such as skeletal joint connectivity information with bone-length ratios in a fixed canonical scale. The proposed model employs three consecutive differentiable transformations namely forward-kinematics, camera-projection and spatial-map transformation. Furthermore, devoid of unstable adversarial setup, we re-utilize the decoder to formalize an energy-based loss, which enables us to learn from in-the-wild videos, beyond laboratory settings. Comprehensive experiments demonstrate our state-of-the-art unsupervised and weakly-supervised pose estimation performance on both Human3.6M and MPI-INF-3DHP datasets.

Udit Gupta and Sathish Vadhiyar - HiPC 2019
Paper Title: Fast and Accurate Learning of Knowledge Graph Embeddings at Scale

Knowledge Graph Embedding (KGE) is used to represent the entities and relations of a KG in a low dimensional vector space. KGE can then be used in a downstream task such as entity classification, link prediction and knowledge base completion. Training on large KG datasets takes a considerable amount of time. This work proposes three strategies which lead to faster training in distributed setting. The first strategy is a reduced communication approach which decreases the AllGather size by sparsifying the Sparse Gradient Matrix (SGM). The second strategy is a variable margin approach that takes advantage of reduced communication for lower margins but retains the accuracy as obtained by the best fixed margin. The third strategy is called DistAdam which is a distributed version of the popular Adam optimization algorithm.Combining the three strategies results in reduction of training time for the FB250K dataset from twenty-seven hours on one processing node to under one hour on thirty-two nodes with each node consisting of twenty four cores.

Surbhi Aggarwal, R. Venkatesh Babu, Anirban Chakraborty - WACV 2020
Paper Title: Text-based Person Search via Attribute-aided Matching

Text-based person search aims to retrieve the pedestrian images that best match a given text query. Existing methods utilize class-id information to get discriminative and identity-preserving features. However, it is not well-explored whether it is beneficial to explicitly ensure that the semantics of the data are retained. In the proposed work, we aim to create semantics-preserving embeddings by adding an additional task of attribute prediction. Attribute prediction is also used in attribute-based person search. However, manual annotation is needed to get the ground-truth attributes, while we automatically mine attributes from the text corpus. In summary, we propose an approach for text-based person search by learning an attributes-driven space along with a class-information driven space, and utilize both for obtaining the retrieval results. We show via extensive experimentation on benchmark dataset, CUHK-PEDES, that learning the attribute-space not only helps in improving performance, giving us state-of-the-art Rank-1 accuracy of 56.68%, but also yields humanly interpretable features.

Poster Presentations at HiPC Student Research Symposium 2019

- ***Graph based Algebraic Multigrid Method:*** Manan Shah and Sashikumaar Ganesan
- ***Parallel Smoothers in Multigrid Method for heterogeneous CPU-GPU environment*** **-** Neha Iyer and Sashikumaar Ganesan (Won the Best Poster Award at HiPC 2019)
- ***Distributed Edge Storage using Erasure coding with differential reliability*** **-** Abdun Nihaal, Sheshadri K R and Yogesh Simmhan
- ***An Empirical Study on Efficient Storage of Human Genome Data*** **-** Diksha Chaudhary, Bratati Kahali and Yogesh Simmhan
- ***Drone Video Management System*** **-** Ankit Barai, Pradyumna Ym, Aakash Khochare and Yogesh Simmhan
- ***Interval centric indexing for temporal reachability queries*** **-** *Animesh Baranawal and Yogesh Simmhan*

IGCM Social Event – Mysore Trip

Indo-German Conference on Computational Mathematics (IGCM) 2019 was jointly organized by CDS, with Mahindra Ecole, Hyderabad and University of Honnehiem, Heidelberg Germany from 2nd to 4th December. CDS also organizes social events with such conferences which makes experiences of our invited speakers more memorable. This time it was a short trip to Mysore.

Tourist Guide explaining history of Keshava Temple to German Professors

The trip started with visit to Keshava Temple, Somanathapura. It was built by commander of Hoysala King Narsimha III, Somnatha in 1268 CE. Keshava Temple demonstrates rich heritage of India's Sculptural art and ancient Indian architecture. The Walls of temple have major events of Ramayana and Mahabharata inscribed on it. The main structure of temple is supported by turned pillars. Although Lathe machine was invented much latter, one cannot imagine to manufacture such masterpiece without lathe which ancient Indian Sculpturers did with their skill-sets.

The journey then proceeded towards visit to Mysuru Palace. Mysuru Palace witnesses 6 million visitors annually, 2nd highest numbers of visitors in India after Taj Mahal. Mysuru Royal Family has always supported IISc right from its establishment a century ago. Infact, Krishnaraja Wodeyar IV donated 371 acres of land and Rs. 5 lakhs in 1907 for establishment of IISc. Its official name is Amba Vilas Palace. The palace houses paintings of Royal family of Mysore, the gifts that Maharaja of Mysore received from kings and visitors across the world. The front part of the palace consists of asbestos ceiling which has paintings of different zodiac signs on it. British Architect Henry Irwin was chief designer of Palace. It took 15 (1897-1912) years to complete the construction of palace and cost of Rs. 41,47,913. The Palace is a three storey structure with Indo-Sacaracenic style. The palace is garlanded with 97,000 bulbs which are illuminated on Public Holidays and Sundays which one should not miss.

The Last place of trip was Srirangapatna island town located in Mandya district of Karnataka state on banks of river Cauvery. Located near Mysuru, it is a place of religious, cultural and historic importance. Near to it is Tipu Sultan Summer Palace, a masterpiece of Indo-Islamic architecture. It exhibits arms and ammunition used at that time and painting demonstrating Anglo-Mysore war. Srirangapatna is also nominated for place of cultural and natural heritage in UNESCO.

Visit to such historical places which are designed with elegance keeps curiosity in us alive and make us feel proud of our rich cultural heritage.

Group Photo of Social Event of IGCM 2019

Faculty Interview

- with Dr. Phani Motamarri

Q1. Can you brief your background

I did my undergraduate studies at National Institute of Technology Karnataka, Surathkal (NITK), followed by a master's degree at IISc, both in Mechanical Engineering. I had a brief stint as a researcher at General Motors R&D -- India Science Lab in Bangalore for two years before heading to the US for my PhD. I got my Ph.D. from University of Michigan, Ann Arbor in the broad area of Computational Materials Physics. After a brief postdoc there, I worked as a research faculty at University of Michigan for 4.5 years before joining IISc-CDS as an Assistant Professor in Dec 2019.

Q2. Can you please tell us about your research interests and how you see your research field developing over the next few years.

My research focus is broadly in the area of computational material science, in particular, ab-initio modelling of materials (using quantum mechanical theories). Other research interests include high performance computing (HPC), finite element methods, computational solid mechanics and open-source ab-initio code development (DFT-FE).

Today, numerical simulation codes spanning areas like combustion, climate modelling, earthquake modelling etc., can efficiently leverage parallel computing architectures to scale very well on existing supercomputers. However, state of the art ab-initio material modelling codes are restricted to very small material systems and mostly remain in the high throughput calculation mode to fill up parallel computers. Traditionally, these never have been HPC centric codes. One of the significant objectives of my research is to remove this restriction by developing appropriate mathematical techniques and HPC oriented computational methodologies to conduct fast, scalable and accurate large-scale ab-initio calculations, that can take advantage of the disruptive advancements happening in computing architectures today. Furthermore, my research has implications in ICME (Integrated Computational Materials Engineering) an approach to design materials that involve linking material models at multiple length scales. To this end, my research will be able to provide useful inputs to higher scale models which can in turn be used to predict macroscopic material properties. Accelerated materials discovery is another area where my research can have a significant impact. In this case, one needs accurate numerical simulation methodologies which must be fast and accurate in order to search for materials with desired functionality by screening many materials.

Q3. What is your philosophy of education

I believe that students should place utmost emphasis on building strong foundations in their core research subjects and academic institutions play a very important role in imparting this education. Furthermore, getting them exposed to solving real life problems is very crucial. The first and foremost step in any engineering research is to build a proof of concept of your research idea and demonstrate this on small benchmark problems. However, I feel your research should also prioritize translating your proof of concept idea to a realistic implementation so that other people could make use of your research in their real-life application. This would create a significant impact of your research. Sometimes, this effort can lead you to new research problems as well.

Q4. How is doing a Ph.D. in US different from doing it from India?

I can tell you about PhD experience in the US as I am still new to Indian system for commenting about PhD in India. In top US universities, most research groups focus on the impact of their research in addition to focusing on scientific contribution and hence in the process become famous. Moreover, top universities in the US enter trendy research areas much ahead than rest of the world. They provide you a world-class experience of both the research infrastructure and the research ecosystem. In the US, most students are pushed very hard to be productive because it is more like an employee-employer relationship. This can be both good and bad. It is good in a way as you have more research productivity but sometimes it can deter you from attacking deep research problems. Furthermore, you meet people who are experts in their fields when you travel for conferences in the US and the connections you make in this process help broaden your research perspective across scientific domains. Conventional engineering branches have substantial interdisciplinary flavour in their research in the US which does not exist in India to a large extent. However, things are changing in India as well. I think, departments like CDS in IISc are doing an excellent job in filling this gap.

Dr. Phani Motamarri
Assistant Professor

Q5. How did University of Michigan help to shape your research interests?

University of Michigan, Ann Arbor (UoM) is an excellent place for research. It is home to 25 Nobel prize winners, 6 Turing award winners and ranked consistently top 5-10 in the US in almost all branches of engineering. It has played a significant role in shaping my research interests that are interdisciplinary in nature. Although I was working in the department of mechanical engineering, I was able to conduct research that combined ideas from quantum mechanics, material science, solid mechanics, numerical methods and high-performance computing.

UoM helped me to realize the proof of concept methods developed in my PhD into a massively parallel open-source computational framework (DFT-FE) for conducting fast and accurate material modelling using density functional theory with adaptive finite-element based methods. These methods were nominated as the ACM Gordon Bell Prize finalist at Supercomputing conference at Denver, Colorado in 2019. This computational framework has given me a big runway for my future research in terms of pushing the current limits in our ability to study complex material systems and provide deeper insights into various aspects of material properties at nanoscale.

Q6. What are the main research areas your lab would focus on? What would you expect from students who wish to join your lab?

My newly formed research lab " Computational Materials Physics Lab (CMPL)" will focus on advancing the current predictive capabilities of computational based design of materials. In particular, the focus will be on developing mathematical techniques and HPC driven computational algorithms that can leverage the latest heterogenous parallel computing architectures and future exa-scale machines for ab-initio material modelling. Furthermore, my lab would take advantage of the developed techniques to address interesting material science problems with applications geared towards mechanics of materials, battery materials, catalytic materials and next generation bio-molecular electronic devices.The most important skill that I would look in prospective students is their motivation to learn and willingness to think creatively. They should be passionate about learning new computational methods, condensed matter physics, parallel computing on heterogeneous architectures, finite-element methods.

Q7. What are your professional goals in the next 5 years.

I would like to see myself as a leading researcher in the development of HPC driven real-space computational methods for accelerating large-scale ab-initio calculations and thereby contribute to high fidelity material modelling leveraging the developed methods. I will make sure these methods will be made available to community in the form of an open-source code so that everyone can benefit from these methods. I would want to see a submission sent to ACM Gordon Bell prize from IISc, this time.

Q8. In Computational Science Research, there are three aspects -- one is understanding the governing equations for the problem at hand, second aspect involves the development of robust numerical methods to solve the problem and third one is the efficient and scalablen implementation taking advantage of the system architectures. How should a researcher having a background only in one of them handle other aspects of this problem?

A computer systems scientist would not be interested in material modelling or modelling a fluid flow problem. A mechanical/aerospace/materials engineer may be a domain expert but may not have enough background in parallel programming aspects or knowledge in scalable computer architectures. This gap must be bridged to have a practically useful implementations which can work on realistic problems. To this end, I feel a domain scientist must invest substantial amount of time to understand high-performance computing aspects required for their problem. As a person who knows the physics of the problem and a good understanding of the numerical method employed, it is the domain scientist who can figure out the best scalable implementation strategy if he is well versed with HPC aspects. For instance, it is important for the domain expert to be well versed with parallel programming paradigms with CPUs/GPUs, asynchronous programming techniques and sometimes, he may need to examine the scope of using mixed precision to improve the throughput performance of their method or reformulating their computational technique to minimize communication costs and memory access costs etc., This sometimes requires abstracting out the physics of the problem and the underlying computational method to seek out collaboration from computer systems scientists. Infact, a domain scientist developing these skill sets can set himself apart from others because he can speak both languages (of a domain scientist and a computational scientist as well).

Q9. How was your transition from Mechanical Engineering towards interdisciplinary HPC centric research which combines ideas from various scientific domains?

My transition was basically in two steps. The first step was about identifying and taking the right courses. This was exciting because you are learning something new and different from what you have seen in traditional mechanical engineering courses However, I felt there had to be a balance in terms of how much coursework was needed and how much stuff I need to pick up myself during my research. This was where I used to read, have informal discussions with people who are experts in that area, travelled to conferences and connect with people. I also learnt a lot by participating in scientific hackathons where you get to interact with excellent computational scientists and these interactions in fact helped me with many of the implementation innovations, I carried out in my research. The transition was not that difficult indeed!

Q10. How would you go about motivating researchers who are going through a low point in research?

As a mentor, I would not mind sitting with them and help them debug their code for instance. I would try to make them ask the right questions that would help them think through the problem they face. It appears simple but believe me if you know the right questions to ask, you are halfway through the solution for the problem you are seeking.

Q11. Can you please point out the similarities and differences you see in the course structure in IISc and in University of Michigan?

I have asked this question to myself many times. I feel course structure in IISc is on par with best universities in the US. Infact, some courses are more rigorous in IISc, as in the US the graduate courses must cater to students from a wide variety of backgrounds. I had an overlap of one course which I took in IISc and again at University of Michigan because of the core course requirement. In IISc, I got a rigorous mathematical perspective, and, in the US, the same course was handled giving more emphasis on physical intuition. Both perspectives are unique in their respects and I feel good having experienced both.

12. Can you tell about the books that have inspired you as a researcher?

In my graduate school days at IISc, a book that provided me a lot of inspiration and built my foundations is " Continuum Mechanics" by Prof C. S. Jog who happened to be my M.Tech advisor as well. As a person who was very fond of tensor analysis in that book, I found very nice parallels when I first got introduced to Dirac notation from the book ``Principles of Quantum Mechanics" by R. Shankar. My interest in quantum mechanics got intrigued with this book which formed the backdrop for my journey at University of Michigan later.

Alumni Experiences from CDS Alumni Reunion,2019

NANDITA NAYAK,M.Tech in Computational Science 2006-2008, Sr. Computer Vision Engineer at NVIDIA

It was a very enriching experience overall here at IISc. It is a lot more interactive here, actively doing projects, a lot of assignments, working at nights, going for tea at tea-board (Now, it is called Prakruthi) in the middle of the night and having discussions with classmates around. The whole experience was to define what you are going to do in future. A lot of my peers went for research at different universities. Many of my peers defined what they want to be, being at IISc. I do think that the industry and academics are a lot closer now than they used to be before as almost one drives the other. In United States, I have noticed that the academics is really driven by the industry, a lot of funding comes from their engaging actively with the industries, many professors take sabbatical and lead startups for couple of years and come back. Now industries are also highly encouraging students to pursue PhD. I encourage you to think of where your research is heading, if it can actually be used in the real world, what products you can make out of it so that you can have your own company in the future and be a job creator.

KALAPRIYA KANNAN

PhD in Computer Engineering 2002-2007,Senior Research Engineer at IBM Research

The first thing that I learned from IISc is that many people think that once they get into their PhD they think it is the final stage of their learning. But I think PhD is just the beginning of learning and there is a lot to learn beyond that also, once you enter into jobs or research. In IBM research what I experienced is that in every two years the topics of agenda changes. It's completely different. So if you don't keep updating yourself, you are out of the team and you are out of the market. The second thing is, it gives you a holistic experience of learning, like most of us come to IISc with good academic backgrounds and when we present our fantastic ideas submitting a paper and then often we get a rejection. It could also be a cycle of rejections and it needs a lot of effort before it finally gets accepted. So it gives you a holistic view that its not just your book and paper but you also get to know about rejection and downfall and its all okay to have such experiences in the long run. The third thing is that it changes your perspective completely. During MTech, we are very confident of our own work and we tell "this is the *only* way....", but once you do your PhD, you actually start telling, "I *think* it is like this ..". The good part is that now you are open to accept information, views and opinions from others which is actually very important. Without listening, I think you will not go anywhere in your professional life. You should enjoy your journey more than just your thesis and you should give back to the institute because it has produced you. My main job is research and I do university relationships as a part of my passion.

ANSHU SHUKLA

M.S. in Computational Science 2014-2016, Data Platform Engineer at Ericsson

I passed out from DREAM Lab in 2017. I have not opted for PhD till now. So, I might end up with a whole new perspective as I am not that senior. I think at IISc we are more inclined towards the theory side and pretty less on the actual core systems point of view. I think we need a really broader view and a lot of effort is required to be good in systems research. As you can see outside India, the universities have very strong relationships with the industry and systems community as well. After my masters completion, I am working in Ericsson Research cloud team. The main problem is that there nobody is talking about what IISc and IITs are doing in cloud research community but they check what the universities abroad are doing in their domain. If you think of the theory side, every company will look upon us as our theoretical side is strong. I think strong relationship with the industry is required. Professors and senior level people can take a decision of what can be done to overcome these inefficiencies. In terms of research, we are always ahead of others but we are lagging in systems research. In Database, IISc is good. It may be that we are doing good but are not getting adequate exposure. This is the main perspective that I can think of. Other thing is, IISc is obviously the best place to learn. The main bottom-line thing that I learned from IISc is that you have no limits and you can always push yourself. In Industry atleast, nobody is there to push you. In research based industry, they expect you to do your best. But in IISc, you always have an example that this guy is much ahead than me. I really miss that in industry.

Alumni Entrepreneur Interview

- with Abhishek Upperwal

Q1. You are the founder and CEO of Soket Labs. What is Soket about? Where do you envision Soket in a few years?

Soket Labs is a deep-tech startup with a vision to build better cities with Data. All this started back in IISc when I was exploring different datasets originating from cities around the world. It was evident that Indian cities lacked the data culture. We are hardly collecting any data and even if we are, governments have no idea how to use it. I believe it is up to people like us who understand data and have the right expertise to come forward and innovate.This inspired me to build "Pravah". It's is a data exchange platform for high velocity and real-time data. The amazing part is, it's decentralised. This means data does not flow through a centralised server, instead, our technology helps a data consumer find the right set of data producers in a peer-to-peer network and initiate a direct data stream. You can compare it with the BitTorrent or ADC protocol used in DC++ (I am sure all of you use it to download new movies and lecture videos). This architecture reduces data transmission latency, improves scalability with no single point of failure. Through our platform consumers can get access to real-time datasets, be it locations of public buses or metros, air quality data, solar power plant data, traffic congestion, and various other datasets.All this hard work is to build an ecosystem where city data is available for all to innovate and build amazing applications. Being a big believer in open source, all our code is licensed under MIT open license. Feel free to checkout the tech stack at *https://github.com/pravahio/go-mesh* or visit *https://pravah.io*.

Q2. How did you get involved with Blockchain?

This was a time when I had just joined CDS (back in 2016). I was fascinated by the complexity of the Blockchains. It is an amalgamation of so many different technologies and domains; computer networking (peer-to-peer), cryptography, economics, and my favorites maths and distributed systems (consensus/byzantine fault tolerance protocols). It's an engineer's dream to study and contribute to something so new and complex. After failing to even understand Ethereum's whitepaper, I started reading it's source code (yup, programming languages are sometimes better than the English language). I was able to learn the nitty-gritty details of almost everything that made up a Blockchain system. Instead of being an application developer on Blockchains, I was more interested to understand it's internals and theoretical working. I think this knowledge played an important role in how I architectured and designed Pravah.

Q3. Can you tell us about your other ventures: Merkle Labs and Qubit Technologies?

Merkle Labs is more of an initiative for open Blockchain technology for our society. I saw the potential of Blockchains for the public good and how it can bring transparency and accountability to our system. I have been working on a few projects like Public Grievances redressal on Blockchains and "Open Complaint Network" which can help anyone file an FIR using smart contracts.
Qubit was my earlier venture which I started with one of my professors during my undergraduate. It was a service firm that use to help businesses develop technological solutions.

Q4. Can you tell us about your experience at CDS?

CDS and IISc have been my home and I have enjoyed every bit of it from an amazing campus to late-night or even full night stay at the MARS lab. My motivation to pursue masters was an optimisation problem I encountered during my undergrad project. While working on a computer vision problem and going through numerous research papers, I could not wrap my head around the concept of Jacobians and Hessians. It was damn confusing for my little mathematical brain. I knew that was the limit of my knowledge.

CDS has not just taught me about Jacobians and Hessians, it has expanded my brain to think about complex problems and solutions to them. I got an opportunity to work and learn from some of the best researchers in the world and find new friends who are brilliant at what they do. The most amazing thing that I learned and still apply in my life is how to work efficiently under some "astronomical" stress but take my words, it's worth it.

Q5. What has been your experience as a young entrepreneur in India and what advice would you give to aspiring entrepreneurs?

I would describe it in three words; Passion, Excitement, and Stress. Startups are like non-convex functions unfolding in time on the x-axis. You have local minima (this is where you go into stress mode) and local maxima (this keeps you excited and motivated). Being unbounded in time, no global maxima exist and as an entrepreneur, my priority is to jump from one local maxima to another by taking the most optimal route. My passion for the problem keeps me excited and helps push me to the upward slope and setbacks (they are very common) to help me learn what not to do.
To be frank, this is the best time to start your venture especially if you are young. At the core of any startup is innovation and all of you are amazing at it. The only component that is needed, is leadership. I would just say, lead your innovation and build a business around it. If the most brilliant people at IISc don't solve some of the biggest challenges our world faces, no one will.

Sample Use of SahasraT Supercomputer by CDS in 2018

BIOMOLECULAR COMPUTATION LAB

~1.31 million core hours used on SahasraT

To understand the role of kinase activator on AMPK protein kinase and its mutant in breast cancer at structural and functional level, molecular dynamics silmulations using the GROMACS package is employed. 10x speed-up was achieved by using SahasraT where 30 nanoseconds' worth of molecular dynamics simulation was completed in a day.

COMPUTATIONAL MATHEMATICS GROUP

~0.3 million core hours used on SahasraT

ParMoon, an open source finite element package for the scalable (parallel) solution of partial differential equations, was modified to exploit GPUs along with hybrid MPI-OpenMP parallelism already present. SahasraT's GPU nodes were used to push the degree of parallelism and use upto 3 GPUs in parallel.

MIDDLEWARE AND RUNTIME SYSTEMS (MARS) LAB

~0.2 million core hours used on SahasraT

A multi-node multi-device algorithm for finding minimum spanning tree (MST) that uses a divide-and-conquer approach by partitioning the input graph across multiple nodes and devices and performing independent Boruvka's MST computations on the devices followed by a novel hybrid merging algorithm that ensures that the combined results on a node never exceeds the memory capacity of any single node was proposed. It was deployed on SahasraT to use up to 16 GPU nodes of SahasraT, completely utilizing all the CPU and GPU cores in those nodes. 24-88% performance improvements were seen over existing state-of-art and almost linear scalability was seen for large graphs.In another research, performance models were built to predict weak scaling of scientific applications using small-scale runs. Upto 16K cores of SahasraT were used to predict sacalibility in LAMMPS and SMG applications.

Faculty Interview

- with Dr. Sathish S. Vadhiyar

Dr. Sathish Vadhiyar
Associate Professor

Q1. Can you share your academic journey with us?

I will start from how I got interested in doing Masters when I was an undergrad. When I was in the third year of my undergraduate program, one of my cousins planted the idea of doing masters abroad in my mind. One thing that I have found out about myself over the years is that competition has always suffocated me, and I have always tried to find escape routes, which eventually turn out to be at least as good as the others. So, when almost all my fellow undergrads were looking for good placements, I found this escape route of writing GRE and applying for masters which eventually led me to Clemson University.

Clemson University was largely a coursework-oriented institute with a good environment for campus placement. Clemson had tie ups with companies like Qualcomm and Sun Microsystems at that time. The journey was kind of new to me, but at the same time looking back it was not perhaps that eventful. Once again, my tendency of getting suffocated from competition made me apply for PhD when most of the others were applying for jobs. I eventually joined the University of Tennessee, Knoxville for a PhD.

I was always clear about getting back to India after doing my PhD. When I was in the last semester of my PhD program in the University of Tennessee, I started applying for jobs in India. I got a response from Prof S. M. Rao, the then SERC chair and the opportunity was exciting to me. I was attracted by the "Supercomputing" buzz word, just like many of you are currently attracted by "data science" and "machine learning" buzz words. I was hoping that it would be more of a technical role where I can lead an HPC application optimization group like the project lead role in national supercomputer labs in the US. It was well in the middle of the application process that I came to know that they're considering me for the assistant professor role. After some initial skepticism and a few emails with the chair, I was convinced and eventually joined here in 2003. Prof S. M. Rao made great efforts in getting me here and I must thank him for that. I remember thinking at the time that if I were the chair, these are exactly the steps that I should take to get potential candidate to the department.

Q2. What was it like working with Prof. Jack Dongarra?

Three things stand out about my experience doing PhD with Prof. Jack Dongarra.

Firstly, his group was set up hierarchically with a total of around fifty people comprising students and project staff. And I'm guessing he was at the peak of his career at around the time when I joined. The presence of this large dynamic group with varying capabilities resulted in a conducive environment for acquiring knowledge. There were always people around to have enriching discussions with and to answer questions I had.

Secondly, Prof. Jack Dongarra provided many opportunities, to me and others who worked with him, including travel to conferences and establishing connections with the well-known players in the field. I had the chance to be part of large multi-institutional project groups. Some stalwarts used to visit us and give talks and have informal get together with us.

The system in the USA is that faculty are funded by different projects and it is through the projects that their students and labs are sustained. So, it is important for the students to contribute towards these funded projects even if these projects don't necessarily relate to their PhD topics. I happened to work in two to three funded projects. In fact, some of the most cited papers of mine belong to these projects, even though they were not entirely related to my doctoral dissertation. This is the third aspect which stands out about working in Prof. Dongarra's group.

It was a highly motivating environment due to these and many other aspects. It was one of the best phases of my career as well as my life. I hardly had any complaints at all during my PhD and that's a big statement to make.

As to how it was like working with the great Prof. Dongarra himself, I have to disappoint you unfortunately since, because of the large size of the group, I didn't get much chance to observe him to comment about the working of his mind. He was generally fun loving; he was never strict. He used to travel a lot and that's one of the reasons we didn't have much time to spend with him. Jack has the knack of picking up projects and topics that are very timely, that will be well appreciated and useful to a large community across the world. These are some of the traits I think that we all should try to inculcate.

Q3. How has HPC evolved over the past few decades?

I can speak about how HPC has evolved in the past three decades. There are three distinct milestones here. First is the emergence of MPI-based distributed programming. Then came multicore systems, and then the era of GPUs. I think we are now in a phase where machine learning is very much tied to HPC.

Among these milestones, personally for me, the two significant ones are MPI programming and emergence of GPUs. MPI really interested me and got me involved in HPC. It was quite fascinating to write MPI programs. We are really indebted to the GPU era primarily because the availability of low-cost GPUs spread HPC in India. About a couple of decades back, you could hardly find anyone in India doing significant HPC. And now you can buy GPUs and place them in your desktop for real cheap and then many thousands of cores are available to you. So that made people to think about harnessing those cores and that really opened a lot of HPC-based research in India. That was quite satisfying to witness.

Q4. What are some exciting prospects in HPC in the near future?

Anyone in the field of HPC will tell you that it is the dawn of the exa-scale era in about three to four years' time. This poses exciting questions, both algorithmically and in terms of building middleware and runtime systems. Fault tolerance at exa-scale is another thing that will be emphasized in the coming years. Asynchrony is going to play a big role in dealing with large-scale irregular applications. More efficient methods to use heterogeneous computing resources will emerge. All of that will come into play and we have made humble starts on all these fronts, so we look forward to the future.

Q5. Can you share your experiences as the SERC chairman with us?

Well, it has been an interesting learning process so far. We have a very good disciplined team here in SERC. It has been a good teamwork both within the building and with the HPC users of the institute. I should also express my sincere thanks to my CDS students who have adjusted to my change in role to SERC Chair in 2018. The change brought in changes in times and venues of the meetings, and change in my character as well involving much more expression of dissatisfaction and more yelling! Of course, they didn't have a choice!

Q6. How do you envision future research in MARS lab?

I want my group to work on large-scale projects as a team, building and contributing towards a comprehensive framework that encompasses different modules of work that are managed by individuals, but all of them coming together and contributing towards the framework. I'd like the lab to also work on developing novel programming models for parallelism. I would like the lab to work on end-end solution of societal problems, even though many of the modules will not be relevant to our areas of interest. And I think emergence of exa-scale computing is presenting us those opportunities.

Q7. What are your interests/hobbies outside research and work?

Well, I hardly have a life outside research, but I can say the primary activity outside my research work is watching my kid grow and being part of his growing process.

Q8. Can you share your research approach and philosophy? What is your advice to students at the dawn of their research career?

I think the most important thing in research is to define a "good" problem statement. The goodness can be either how intellectually stimulating the problem is or how relevant it is, and its ability to serve the community at large. Coming up with a well-defined "good" problem statement to me is crossing 60% of the bridge, whatever that is. So, it's better to invest a decent amount of time in that. Another important aspect is to make an objective assessment of the relevance of our work. Sometimes, we tend to get carried away by what we do and start believing that it is important work. Making a true assessment throughout the research process is crucial. I think we've always endeavored for these in our works. So, you can categorize these as my research approaches and philosophies.

Regarding the advice that I can give, honestly, I think most of you guys are more mature than me in many ways. So, you really don't need my advice. And I sincerely believe in that.

But there are certain reflections which I would probably share with myself. Firstly, it is imperative to take ownership of your research as early as possible. There is a lot of difference between you identifying a leakage in your house and finding a leakage when you are visiting your friend's house. So, in the first case, the problem is yours. You develop an ownership to the problem and good solutions automatically appear. In the second case, you throw your hands in the air and speak about vague solutions that may or may not work. Secondly, the underlying theme or story of your research thesis is important. Rather than working on disparate things, cohesion in thesis should be strived for. It is important to be aware of the big picture and analyze your current activity from the perspective of the big picture.

Sometimes we also get caught up in doing something just because it appears cool and we don't put enough thought on whether it has already been done or whether it will be useful. Being honest with yourself, taking a hard look at your research and evaluating your work objectively is very important.

CDS Seminar Series

- ***" Super-fast Graph Queries even Without Indexes" by Prof. Srikanta Bedathur from IIT Delhi***
- ***"Erasure Coded Computations" by Prof. Ananth Grama, Samuel Conte Professor of Computer Science at Purdue University***
- ***" Cross-Accelerator Language Support" by Dr. Henry Gabb, senior principal engineer in the Intel Architecture, Graphics, and Software Group***
- ***"Exploiting Low Precision Floating-Point Formats in the Solution of Linear Systems of Equations" by Dr.Srikara Pranesh, research associate in the department of mathematics at the University of Manchester***
- ***Data Flow Execution Models -- a Third Opinion by Prof. Vivek Sarkar, Georgia Institute of Technology***
- ***"Higher Order Methods For Machine Learning Appilcations" by Prof. Ananth Grama, Samuel Conte Professor of Computer Science at Purdue University***
- ***"A fine-grained perspective onto object interactions " by Prof. Dima Damen, Associate Professor (Reader), University of Bristol, UK***

Placement Experiences

PREYASHI AGARWAL

I am an M.Tech Research student in the Department of CDS. I associate with the Cloud Systems Lab and work in the domain of Microservice architecture. I got placed in Wells Fargo and Samsung R&D during the last placement session. As I was focusing for system profiles, I started my preparations by refreshing domain knowledge in the areas of Data structures, Algorithms, Databases, Virtualization and Operating Systems. I had also practiced on different online coding platforms. In addition to that, the companies I have been interviewed for were interested to know in depth about my research area as well. My advice with regard to placement preparations would be to start early, though the preparation time would vary from person to person. It is also important to identify the domain where one wants to work and prepare for the type of work and technologies that companies focus on in that domain.

CDS Placement Highlights

Placement Session at IISc starts at in the month of October. CDS Department is one of leading departments in IISc for placements. The roles offered to students has diversity due to it interdisciplinary flavour. The following is the list of organisations in which students of our department got placed this year along with the roles offered.

1. Wells Fargo : Software Developer and Analyst Roles
2. Goldman Sachs: Analyst Role
3. Microsoft : Software Development and Data Scientist
4. Flipkart : Data Scientist
5. Samsung R&D Bangalore: Reseacher and Developer roles
6. Qualcomm : Machine Learning Engineer, Multimedia Engineer role
7. Mastercard : Data Scientist
8. Zendrive : Data Scientist
9. Exxon Mobil: Computational Engineer Role
10. Robert Bosch Engineering and Business Solutions India Ltd: Computational Engineer Role
11. Walmart : Software Development Role
12. Nutanix : Software Development Role
13. ZS Associates: Data Scientist Role
14. Archeron Group: Data Scientist Role
15. Ernst and Young: Data Scientist Role
16. Cisco : Software Development Role
17. Citrix : Data Scientist and Software Development role

Alumni Experiences

UDIT GUPTA

When I joined IISc in 2017 it was the only premier institute in the country offering a full-fledged M.Tech program in Data Science. It had the vision to foresee the requirements of industry and academia which had limited expertise in this domain. Following course, many other institutes including IITs now have similar programs. But M.Tech CDS is much different from the rest as it is offered by a department which specifically caters to this domain. It would not be an overstatement to say that CDS in IISc was my first choice when I was looking to enroll in an M.Tech program. I was from electronics background and was working in a PSU in the same domain before joining IISc. It was an exciting time to come back to academia after 2.5 years in industry and learn about the latest and most talked about topics. The course structure is very broad where we not only learn about the implementations but also the theory behind its working. IISc is known for its cutting-edge research and when you see people around doing great and impactful things it encourages you to perform at that level yourself and hence we collectively evolve. I was able to convert my M.Tech dissertation into an IEEE research publication with the help of my advisor Prof. Sathish Vadhiyar.

I am currently working as a data scientist at Zendrive. The company encourages safe driving behaviour by analysing smartphone sensor data. It's a fast-growing startup whose engineering team is based out of Bangalore and customers are primarily in US. Work flexibility and good structure makes it an exciting organisation to work for.

UTKARSH SHREEMALI

My journey at CDS, IISc has been a very memorable one. Studying in an institute of such repute, attending the seminars and lectures of highly knowledgeable professors, participating in various informative discussions with my peers and friends, all contributed towards making my journey at IISc a wonderful dream come true. The grind of courses and research at CDS make a person capable of gaining in-depth understanding of a field. The resources provided to us were at par with the best universities in the world and helped students do state-of-the-art research. I would like to give a special thanks to my advisor, Prof. Anirban Chakraborty, who guided me in my research and encouraged me to come up with my own research problems and solutions to these problems.

I am currently employed with optimization team in the Machine Learning Group at Qualcomm. We work towards developing and optimizing the performance of Qualcomm's Snapdragon Neural Processing Engine.

Credits - Dhanya Mathews, Koushik Sen, Manan Shah,Renga Bashyam, Shreya Roy, Siddharth Seth
Edited by Dr. Anirban Chakraborty and Dr. Sathish S Vadhiyar