Final Report

US-India Network Enabled Research Collaboration Workshop

5th- 7th December 2010

Lalit Hotel, Delhi

Table of Contents

Final Report:
US-India Network Enabled Research Collaboration Workshop

held in Delhi, India from 5th to 7th December 2010

James G. Williams - Indiana University
N Mohan Ram – ERNET India
George M McLaughlin – Indiana University
S Ramakrishnan – Former Director General, C-DAC, India

Introduction

The motivation for the workshop resulted from the coming together of a number of factors including:

- the positioning of India as a significant player in the global economy;
- the desire by the US and India to have a stronger collaborative engagement (culminating in the visit by President Obama to India in November 2010 and the joint announcement by President Obama and Prime Minister Singh regarding increased US-India science collaborations, particularly n the areas of climate change and civilian space research;
- the implementation of significant high bandwidth connectivity between India and the US via the shorter trans-Pacific route;
- the implementation of the ambitious National Knowledge Network (NKN) in India; and
- existing research collaboration programs and potential new ones involving US and Indian researchers that are able to exploit the new connectivity to significantly enhance these collaborations.

However, the fact that the opportunities for enhanced network-enabled collaboration are now in place does not mean that these will be exploited automatically. User communities need to be made aware of the possibilities and what they need to do to take advantage of them. Similarly, the networking professionals need to understand what the real needs of the user communities are and ensure that the optimal engineering, protocols and performance attributes are in place in a way that ensures an easy-to-use experience for the users.

Identifying champions and using exemplars of effective network-enabled collaborations that are already happening is a powerful motivator. Running tutorials and having discipline-based presentations of what has been done and planning for what could be done with the new infrastructure helps to catalyze activities. Bringing all of these elements together in a carefully designed workshop involving researchers, network engineers, government officials and funding bodies was an inevitable outcome.

Acknowledgements

Indiana University gratefully acknowledges the support of the US National Science Foundation via award OISE 0960487.

Workshop Objectives and Motivation

The workshop aims were to drive network-enabled collaboration between research groups in India and the United States through innovative applications using the advanced cyberinfrastructure involving the national research and education networks in India (ERNET, NKN) and the US (Internet2, ESNet, NLR), the international circuits connecting the two countries and the other nation-wide cyberinfrastructure facilities in India and the US.

The workshop provided the opportunity for the cyberinfrastructure service providers and the research users from a number of disciplines to come together to jointly develop a way forward that will lead to significantly enhanced opportunities for collaboration between India and the US.

The key outcomes will be the formative stages of the development of action plans for the infrastructure/service providers (ERNET/ NKN/ Internet2/ NLR/ TransPAC3) to enhance/refine their service offerings to support the emerging India-US e-Research Collaboration requirements articulated during the workshop.

The plans will include a list of deliverables that the participants would commit to implement. These will then be trialled by the collaborating research teams over the next year. One possibility is to have these milestones be part of significant demonstrations at international conferences, noting that the 32nd APAN meeting will be held in India in August of 2011 and would provide a good opportunity to demonstrate progress.

Workshop Organization

The workshop was jointly organized by Indiana University in the US and ERNET in India. The workshop was supported by the US National Science Foundation and the Government of India.

The website at http://usindiaworkshop.indiana.edu/ was established early in the process and was used as the main vehicle for all aspects of the workshop. Its content was progressively expanded as the plans developed. Organizing, Program, Tutorial, Fellowship and Sponsorship Committees were formed. In practice, a subset of the Program Committee met regularly (typically fortnightly) using Skype (and in the latter stages the multi-party video version) to progress most aspects of the workshop.

As the workshop drew closer a second website was deployed by ERNET focusing on the local arrangements and the India participant aspects. http://www.events.ernet.in/

A framework for the workshop was developed and progressively refined, as were guidelines for presenters, session chairs and participants, see
http://usindiaworkshop.indiana.edu/framework-guidelines

The workshop attracted 132 registrants from universities, national laboratories and government organizations. There were 103 registrants from India and 25 from the US, and 4 others. Most of the US participants were funded by the NSF workshop award and by supplementary funding from the Vice President for Information Technology at Indiana University.

The tutorials, workshop program and participant list are available both on the web site and as Appendices to this report.

Workshop and Tutorial Day Structure and Content

The workshop was preceded by a Tutorial day. Six tutorials were offered - three concurrent tutorials in the morning and a further three in the afternoon. The topics were intentionally aligned to the topics of the workshop and were well received, well attended and of high quality. The tutorial; leaders and their teams were drawn from subject-matter experts from both India and the US.

The tutorials covered:

(i) Technology, Cyberinfrastructure, Networks and Advanced Communication Systems;
(ii) Open Source Drug Discovery;
(iii) Cloud Computing for Scientific Research and Education;
(iv) Grid-based Cancer Research;
(v) The HUBzero™ Platform for Scientific Collaboration; and
(vi) What makes for a successful Indo-US network-enabled collaboration?

The Tutorial program, abstracts, and what participants should gain from participation are listed in Appendix1 of this report. Details of the tutorial leaders and presenters are shown at:
http://usindiaworkshop.indiana.edu/speaker-bios

Presentations are available at:
http://internationalnetworking.iu.edu/us-india-workshop/workshop-program

The workshop itself started with a well received inaugural session addressed by Minister of State for Communications and Information Technology (MCIT) of the Government of India Mr. Sachin Pilot. Afterwards, Prof. Subra Suresh, Director, National Science Foundation (NSF), addressed the audience via a video link from Washington. Both Leaders spoke of the importance of ICT and, in particular, of research networks like ERNET and the National Knowledge Network in India and the international links between India and the US in enabling opportunity and enhancing cooperation. Both underlined the significance of the workshop in promoting Indo-US collaboration.

The rest of the two days of the workshop covered the following topics, all involving presentations from both the Indian and US sides:

- Technology, Cyberinfrastructure, Networks and Advanced Communication Systems;
- Geosciences, Climate Change, Weather Prediction and Observing Systems:
- Astronomy, Astroinformatics and Astrophysics;
- Cyberinfrastructure for Medical Research;
- Bioinformatics; and
- Nanotechnology and High Energy and Computational Physics

In addition, two demonstrations were held during the working sessions - Accessing the resources of the National Library of Medicine remotely across the networks, and illustrating the power of Earth Observation data from NOAA by analyzing satellite data on night time lights around India using data hosted in the US. In addition, a string quartet from the Conservatorium of the Cleveland Institute of Music performed remotely across the network, while Indian dancers performed locally at the workshop dinner. Further details on these demonstrations are available at:
http://internationalnetworking.iu.edu/us-india-workshop/demos

The Workshop program is listed in Appendix1 of this report. Abstracts are available in the document at:
http://internationalnetworking.iu.edu/sites/usindiaworkshop.indiana.edu/files/US-India%20Workshop%

Details of the session leaders and presenters are shown at:
http://usindiaworkshop.indiana.edu/speaker-bios

Presentations are available at:
http://internationalnetworking.iu.edu/us-india-workshop/workshop-program

Video streams of the sessions are available at: http://www.events.ernet.in/video-archive-sessions

Outcomes and next steps

The feedback at the end of the workshop was very positive (number, diversity and quality of participants and their level of participation in the deliberations; a healthy mix of speakers from both sides in each session; a mix of on-going or previously known and motivated teams for furthering joint collaboration projects, especially with the assistance of advanced network infrastructure(s); bringing together cyber-infrastructure presenters and discipline people). It emerged from the 3 days of deliberation that there exists immense scope for collaborative research between USA and India. It was felt that for ongoing collaborative efforts in the field of Astronomy, Earth Science, High Energy Physics, Cancer grid and Medical research, the participating institutes in Indian side should be provided high bandwidth connectivity to significantly enhance their international collaborative opportunities.

Other positive feedback was obtained from the survey, following the workshop (see **Survey methodology and results** next). Several science collaborations were catalyzed as a direct outcome of the workshop (see section headed **Specific activities and collaborations catalyzed by the workshop** later in the report).

Participants expressed views that issues relating to support of science communities, including funding support, must be articulated in specific terms and taken up at the appropriate level, whenever necessary; similarly, issues with cyber-infrastructure at both sides must be coordinated and resolved in a way that meets real needs.

Participants expressed a strong desire to continue the workshop series with varying locations in India (Bangalore and Mumbai were suggested) and the US.

The organizers plan to convene at the APAN meeting in August of 2011 (to be held in Delhi) to assess progress and, together with various research groups, demonstrate some of the activities catalysed by the workshop through live network interactions between the US and India.

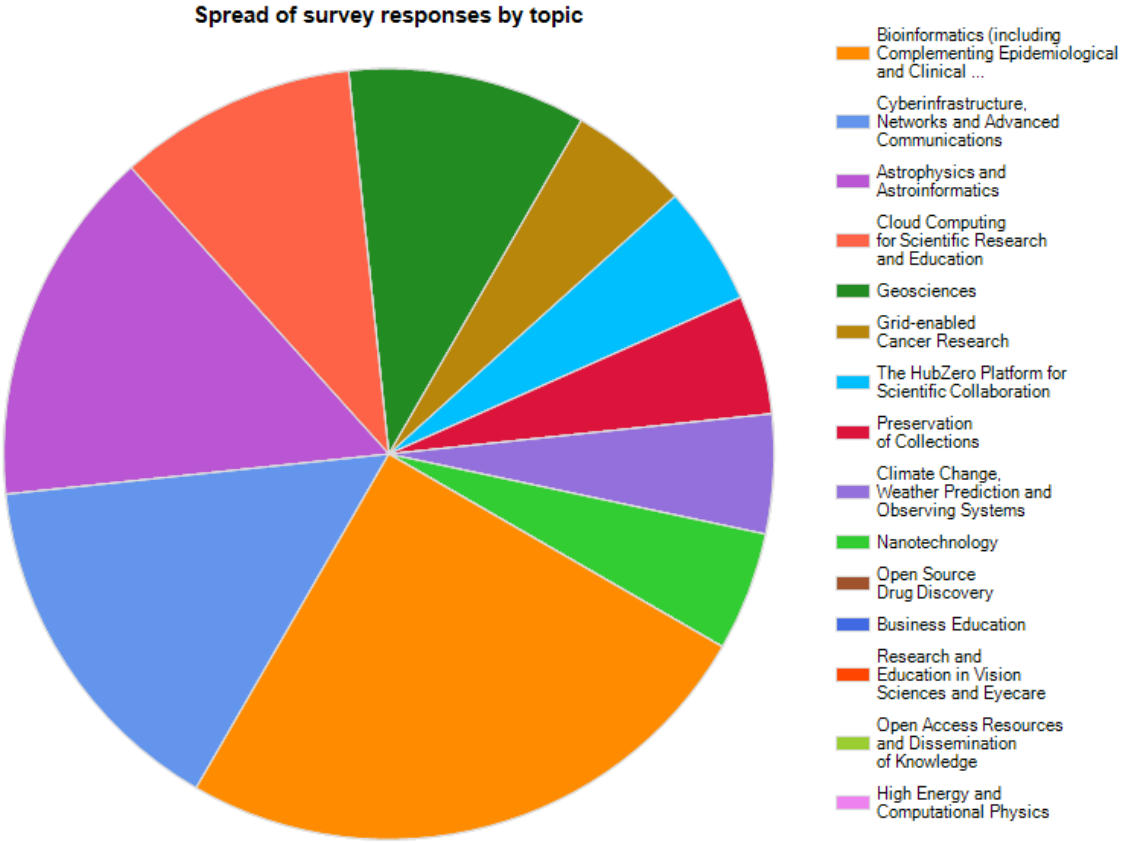
Survey methodology and results

A structured web-based survey was prepared to aid in development of the next steps and was sent to all session leaders, presenters and tutorial leaders (40 in all) of the 15 topic areas covered by the workshop with the aim of identifying:

- potential exemplars that can be quickly put, in place with possible demos at the APAN meeting in Delhi in August 2011);
- obstacles inhibiting uptake of network-enabled collaboration
- the nature of assistance required to enable uptake of network based collaboration
- other feedback from the researchers side

The survey is available at: http://www.surveymonkey.com/s/indo-us_network-enabled_collaboration

By 2nd April 2011, 20 of the 40 leaders and presenters surveyed had responded with similar numbers from both India and the US.



Summary results from the survey:

- 70% of the respondents are already engaged in some form of collaboration with researchers in the other country;
- of those not yet involved in network-enabled research collaboration, all stated that collaborators on the other side were yet to be identified. In addition a number were not sure how to exploit the new connectivity;
- of those currently involved in joint collaborations, almost all rely on high capacity networking;
- responses indicated a readiness for around half a dozen exemplars of network-enabled collaboration that could be pursed, particularly in areas of astronomical transient event detection; preservation of collections; genomics; hazardous weather prediction and climate modeling; developing an Indian Nano-Hub; and collaboration on grid-based cancer research;

Issues raised by the respondent group:

The most often raised concern and recommendation from the feedback groups was the issue of better communication about India-US funding opportunities for collaborative work. This involved a number of different comments and recommendations.

- Wider publicity of collaborative projects and better dissemination of information from both the US and Indian sides and directed toward scientists is needed;
- An exchange of list of experts in USA & India in the field of interest will engender more collaboration;
- A list of funding agencies able to support such collaborations is important;
- Funding to facilitate the research, including funds for students, post-docs, travel, etc.
- Single point of contact or help to reach out to potential collaborators, more formal open calls inviting potential partners from both India & US. Formalization of partnership at the Government level (we can expect much easier steps to collaborate)
- Collaborative projects need to be formalized and key personnel identified and empowered;
- More opportunities for collaborations to get started and to get funded;
- Pilot funding for collaborative projects that have promise will be helpful. Some sort of annual event (alternating between the countries) that has success stories and tracks progress would keep momentum going.
- A suitable committee should function over the net to encourage preparation of a work-plan, seek and consolidate report cards, help bug fixing and help meet well set performance indices. Currently, the picture on who wants what to achieved is not clear
- There is little information conveyed since the Delhi Workshop regarding the status and availability of the network facility for researchers to try and plan their work.
- Increase the number of satellite observing data streams exchanged between USA and India. Some of the data is only available now for a cost, which vastly restricts the volume of data that can be affordably transferred. To have any substantial volume of meaningful data exchanged it has to be free. The access could be restricted to non-profit research activities;
- High speed data transfer and dedicated networks are necessary to enable faster processing and dissemination of derived products from satellite missions.
- Greater visibility for role of Virtual Organization environments (like HUBzero) in collaborative research;

- Link these VO environments to ongoing collaborations to see if there is a fit and if they can provide value to the collaboration;
- Computational Infrastructure and bandwidth. We need to set up a high end cluster and provide access through the web to the interested organizations to run these tools. This will be on similar lines as the Purdue nano-hub, but in India much more hand holding is needed for effective utilization of the resources
- Until recently IUSO lacked high speed connection to the data center. This will be addressed soon. On the other end, there is concern for storage and preservation. We will need to determine the capability on the Indian side as soon as we start to develop projects where we will collect high resolution images in India and are ready to send to US. We understand that they have capability but we will not be sure until we test. In addition, we will be offering courses (that will include high resolution images and video) that can be accessed through breeze in real time. This capability will be tested soon too. We would like to capture and store these presentations in a digital library that can be accessed by students and practitioners as they wish.
- Understanding the relationship between NKN and ERnet and who will be supporting high-performance network based science interactions between the US and India

This feedback and other sources were used to form the basis for the recommendations at the end of this report

Specific activities and collaborations catalyzed by the workshop

While the organizers are aware that a number of collaborative interactions have been catalysed by the workshop, listed here are two specific cases that are good exemplars:

(i) Indo-US rapid detection and analysis of astronomical transient events

The Inter-University Centre for Astronomy and Astrophysics *(*IUCAA*)*, Pune, India and the Caltech Astronomy Department are working together on developing tools and technology for detection of astronomical transients i.e. objects that vary in brightness over short time periods (eg supernova, new asteroid, binary flares).

Transient astronomy requires fast data processing as some of the events evolve within a few seconds and immediate follow-up observations are required to understand the nature of those events.

Since observations are carried out at different locations depending upon the available facilities in the observatories and the geographical location of the observatory, fast communication and data transfer for processing becomes the major bottleneck affecting the effectiveness of the entire process.

It is critical that analysis of these objects happens close to real-time as follow-up observations based on results from the analysis are requires. Some of the methods are being developed by in India and some follow-up facilities are also in India (the origin of the transient detection currently happens in the US). In the near future the number of transients will increase necessitating larger bandwidth and robust networks.

The new high capacity research networking now available between these to institutions opens up new opportunities for advancing collaboration leading to a better understanding of transient event.

Since some of the surveys have already started delivering data, the institutions are ready to make use of the new infrastructure and the APAN Delhi meeting will provide a focus to demonstrate what can be done.

A planning group involving astronomers from Caltech and IUCAA has been established and will work with the APAN meeting networking taskforce to prepare both the science activity and importantly the end-to-end network throughput to demonstrate this activity at the APAN meeting in Delhi in August 2011.

(ii) Bioinformatics Analysis of Cyanobacteria for Carbon Sequestering and Biofuel Production

A proposal from IUPUI (Matthew Palakal), involving the establishment of a pilot collaborative research program on carbon sequestration with the Indian Institute of Technology (IIT), Bombay (Pramod Wangikar) has been submitted to Catalyzing New International Collaborations program as a result of the workshop.

The study involves two specific tasks: Prof. Wangikar, an expert in cyanobacteria, will provide operon data from his laboratory and Prof. Palakal's (PI) lab will conduct bioinformatics analysis and genome-wide studies for species selection. Robist high performing network connectivity between the two sites will be critical elements of the data and analysis sharing.

This project has two long-term objectives: (i) fundamental understanding of the regulatory mechanisms of cyanobacteria—so that these mechanisms can be reengineered to produce biofuel directly rather than as a byproduct of the biomass and, more broadly, (ii) the ability to reengineer the bacteria for production of useful products. As part of meeting the above stated objectives, this work looks at the development of a suite of algorithms and software tools customized for the large-scale bioinformatics analysis of the cyanobacteria genome.

Cyanobacteria can be developed as an excellent microbial cell factory that can harvest solar energy and convert atmospheric CO_2 to useful products such as biofuels. If the outcome and the long-term objectives of this work prove to be viable, they will have enormous, global-level impact on the production of clean energy, especially in burning clean coal. The project will also have a broad impact on education and training: this project will educate young professionals—professionals who need to understand how to use cutting-edge technology to produce clean energy

Conclusions

In the opinions of the organizers, the workshop was a success and met its objectives. There were active discussions and questions throughout the sessions. There was widespread discussion in hallways and during/after lunches and dinners. US and India participants had ample opportunity to interact with each other socially and professionally. The fact that many participants expressed interest in the planned follow-up meetings (APAN in August 2011 and Washington in March 2012) indicates their level of involvement in the workshop and their enthusiasm for continuing that involvement.

RECOMMENDATIONS

Based on the feedback provided by workshop and tutorial participants both during and after the workshop, the organizers make the following recommendations for consideration:

1. Develop a mechanism to better inform researchers of funding opportunities to facilitate Indo-US collaboration.
   {Action: Indian and US funding agencies and Indo-US S&T Forum]

2. Compile and promulgate to all interested a list of experts in US and India in particular disciplines to engender matchmaking and facilitate collaboration
   [Action: TBD]

3. Explore the option of a single point of contact or help to reach out to potential collaborators to advise on opportunities, in particular pilot funding for collaborative projects that have promise
   [Action: TBD]

4. Promote greater visibility for the role of Virtual Organization environments (like HUBzero) in collaborative research;
   [Action: TBD]

5. Consider setting up a high end cluster in India and provide access through the NKN, ERNET and TransPac to the interested organizations to run these tools. This could be based on similar lines as the Purdue Nano-hub, with initial support being provided across the network infrastructure from the US
   [Action: Funding Agencies]

6. Build on India's involvement in international organizations that focus on exploiting cyberinfrastructure for innovative research, such as the Pacific Rim Access and Grid Middleware Assembly (http://www.pragma-grid.net/)
   [Action – TBD]

7. Existing Indian institutes that are, or will be, partners in collaborative research with institutes in United States should be provided high speed access bandwidth (minimum of 1 Gbps) to exploit the 2.5 Gbps bandwidth going out from India towards USA via Singapore. These institutes include: the Inter University Centre for Astronomy and Astro-Physics(IUCAA) Pune; Tata Institutes of Fundamental Research (TIFR) Mumbai; National Remote Sensing Agency (NRSA) Hyderabad; University of Hyderabad; CDAC Pune; AIIMS New Delhi; NARI Pune; TMH Mumbai.
   [Action: NKN, DIT]

8. In order to use the high speed national and international bandwidth available at these institutes, it is also recommended that internal network at these institutes to be upgraded accordingly.
   [Action: Indian side]

9. Take steps to put in place follow-on workshops to build on the momentum and help cement the exploitation of network-enabled Indo-US collaboration.

[Action: Williams to follow-up on the US side to determine the level of interest within the NSF in a continuing series of workshops aiming for the next one to be held in Washington DC in March of 2012].

10. Set-up a mechanism for regular review of progress of workshop outcomes and preparation for next workshop.
    [Action: Workshop organizers. Next planned face-to-face meeting at APAN 32, New Delhi, week commencing 22 August 2011]

11. Set up at least one demonstration (preferably more) of collaborative activity catalyzed by the workshop at the APAN 32 meeting in Delhi in August 2011
    [Action: McLaughlin and Singh in consultation with CalTech Astronomy Department and IUACC to demonstrate use of high speed networks for rapid response to transient astronomical events]

12. Further develop awareness raising activities to help user communities to understand how they go about gaining access to and exploiting the new network infrastructure.
    [Action; NKN, ERNET, Indiana University]

Appendix 1: Tutorial program and abstracts

Sunday 5 December 2010

9.00 - 12.30 **Concurrent Morning Tutorial M1 - Technology, Cyberinfrastructure, Networks and Advanced Communication Systems**

Tutorial leaders: Dr. Radha Nandkumar and Dr. Subrata Chattopadhyay, CDAC

Session 1: Interoperability of the Grid, Mr. B. Asvija, CDAC

This tutorial gives a brief overview of Grid Computing, its need, evolution and the current state of affairs on various popular implementations worldwide. It describes the role and importance of the Grid Middleware in sustaining the massive, distributed, collaborative infrastructure of the grid. An overview of different grid middle-ware implementations will also be presented. Highlighting the differences among the various grid architectures and technologies that build the current day grids, it tries to emphasize on the need for interoperability among these infrastructures to address global collaborative research challenges. Interoperability areas, at various layers of technology, management and operation, will be enlisted. Detailed analysis of achieving interoperability at the various technology layers shall be carried out. An understanding of various attempts from global grid community in this regard will be highlighted. Finally, a case study of the joint collaboration between the European EGEE and the Indian Garuda grid, will be described.

Session 2: GridChem/ParamChem (Dr. Sudhakar Pamidighantam, NCSA)

In this tutorial we will introduce and give detailed description of GridChem and ParamChem virtual organizations. GridChem represents a Computational Chemistry Grid, a production cyber infrastructure serving computational chemistry community in areas such as quantum chemistry, molecular dynamics and computational biochemistry. This NSF(USA) funded project deploys multiple popular molecular modeling applications across a Grid consisting of several HPC sites and provides intuitively familiar application specific user interfaces for research and education in molecular/material modeling. ParamChem is a new virtual organization that is being built on the core GridChem services for molecular forcefield parameterization purposes that will be used in molecular mechanics and molecular dynamics. The tutorial consists of a lecture on the various components of GridChem and ParamChem user interfaces and will provide a detailed view of how web services are integrated into application specific information provisioning frameworks for the end user benefit. Examples will be provided for specific implementation of services and user interfaces.

Some implementation details as to how the user data is managed will be demonstrated. A brief discussion on sustainability of such virtual organizations will be presented. Specific examples for usage for a user and portal administrator will be provided. A set of training accounts will be used for interested participants to use the GridChem Client to create and submit a molecular model for a brief simulation to provide a firsthand experience.

9.00 - 12.30 **Concurrent Morning Tutorial M2 - Open Source Drug Discovery - "Ligand-based Virtual Screening Using Ensemble Classifiers"**

Tutorial leaders:
Dr. Max Kuhn - Director, Nonclinical Statistics at Pfizer, United States of America
Professor Sanghamitra Bandyopadhyay - Machine Intelligence Unit, Indian Statistical Institute, Kolkata
Professor Ujjwal Maulik - Department of Computer Science and Technology, Jadavpur University, Kolkata
Dr. Sundarajan - Group coordinator of the Scientific and Engineering Computing Group, CDAC-Pune
Dr. Jaleel - Assistant Professor, Malabar Christian College, Calicut
Dr. Andrew Lynn - Associate Professor, School of Communication and Information Systems Jawaharlal Nehru University, New Delhi
Dr. S. Ramachandran - Scientist EII, Institute of Genomics and Integrative Biology
The tutorial would be assisted by members from their groups, and coordinated by Nisha Chandran, Jawaharlal Nehru University

The Open Source Drug Discovery (OSDD) project envisages the sourcing of geographically distributed scientific manpower and resources towards collectively solving the problem of drug discovery against M. tuberculosis. The project is presently building an infrastructure of federated resources consisting of analytical instruments, computational resources and personnel distributed across the country. The project seeks to bridge the unequal distribution of resources required for science and education by connecting resource-rich CSIR laboratories and central universities with participants students and research personnel from smaller less equipped institutions. The community of users for the OSDD project is now over 4000 persons, and has generated interest in a wider global community.

Existing tutorials deal with portal development and grid middleware involved with the cyberinfrastructure for scientific communities. This tutorial is designed as the development of applications, documentation and use in a real world scientific project, using work-flows, wikis linked to grid technology. The main focus of implementing these methods would be through the use of R :- An open source application for

statistical analysis, using the package CARET (Classification and Regression Training), applied to the problem of drug discovery.

9.00 - 12.30 **Concurrent Morning Tutorial M3 - Cloud Computing for Scientific Research and Education**

Tutorial Sessions:

Dr. Sorav Bansal - Department of Computer Science and Engg., IIT, Delhi
The talk will introduce the area of virtualization and cloud computing, followed by a discussion on my research group's recent work at IIT Delhi. In introduction topics, I will provide an overview of the virtualization technology followed by an introduction on cloud computing. In research topics, I will discuss our ongoing efforts to apply virtualization to three goals: security, reliability, and performance. I will discuss our approach and present preliminary results.

Prasad Dharmavaram - VMware R&D, Bangalore
In this session, the speaker will talk about the evolution of the virtual data center - A Journey from Desktop to Cloud. The presentation will provide an overview of virtualization and the various cloud computing models. VMWare solutions in the IAAS, PAAS, and SAAS space will also be discussed. The speaker will also explain how educational and research institutes can use the power of 'on-demand' computing provided by internal, external, and hybrid clouds for their compute needs.

Jothi Padmanabhan - Yahoo R&D, Bangalore
This talk will start with introducing the problem of large scale data processing and the challenges associated with it. It will then go on to introduce the Map/Reduce programming paradigm and how M/R can be used to solve the large scale data processing problems at Yahoo! and elsewhere. This will be followed by an overview of its most popular implementation – Apache Hadoop (http://hadoop.apache.org). Hadoop is in wide use at leading Internet companies including Yahoo!, Amazon & Facebook. The presentation will walk through the high level design and architecture of the Hadoop Map-Reduce Framework and the Hadoop Distributed File System (DFS). The objective of the talk is to introduce Hadoop to an audience that is interested in technical and technological aspects of Hadoop.

Dr. A. Paventhan - ERNET India R & D, Bangalore
This talk will present how economic benefits of cloud as seen by business enterprises can be extended to educational domain. Also, ERNET India's research and development plan for cloud educational services and its application in secondary

school education will be covered.

12.30 - 14.00 Lunch

14.00 - 17.30 **Concurrent Afternoon Tutorial A1 - caBIG® (cancer Biomedical Informatics Grid)**

1. caBIG Overview and Indo-US Cancer Research Grid (20 mins)
Deputy Director, Center for Biomedical Informatics and Information Technology (CBIIT) and Interim CIO, US National Cancer Institute

2. caBIG Technical infrastructure (30 mins)
covering SOA; SAIF/ECCF; Semantics; and caGRID
Deputy Director, Center for Biomedical Informatics and Information Technology (CBIIT) and Interim CIO, US National Cancer Institute, and Ravi Madduri, Argonne National Laboratory, University of Chicago

3. caBIG Capabilities (40 mins)
(i). Clinical Services: John Speakman, Chief Program Officer, Associate Director, Clinical Products and Programs, Center for Biomedical Informatics and Information Technology (CBIIT)
and Interim CIO, US National Cancer Institute

(ii). Life Sciences: Ian Fore, Associate Director, Tissue Banking and Pathology Tools, US National Cancer Institute

4. caBIG Deployment (50 mins)
John Speakman
Ganesh Shankar, caBIG Deployment Lead and Manager, Advanced IT Core, Indiana University Simon Cancer Center
CS Ramesh, Associate Professor, Thoracic Surgery and Officer in Charge, Clinical Research Secretariat, Tata Memorial Hospital
Gaur Sunder Team Coordinator. Medical Informatics, Centre for Development of Advanced Computing (C-DAC)
Madhulika Tripathi, BioMantra

5. Q&A (40 mins)

14.00 - 17.30 **Concurrent Afternoon Tutorial A2 - The HUBzero™ Platform for Scientific**

Collaboration

Tutorial leaders:
Michael McLennan, Senior Research Scientist, the Rosen Center for Advanced Computing, Purdue University
William K. Barnett, Senior Manager for Life Sciences in Research Technologies, and Associate Director of the Pervasive Technology Institute, Indiana University.

"HUBzero™ is an open-source software platform used to create web sites or "hubs" for scientific collaboration, research, and education. It has a unique combination of capabilities that appeals to many people engaged in research and educational activities. A little like YouTube.com, HUBzero allows people to upload content and "publish" to a wide audience, but instead of being restricted to short video clips, it handles many different kinds of scientific content. A little like Google Groups, HUBzero lets people work together in a private space where they can share documents and send messages to one another. A little like Askville on Amazon.com, HUBzero lets people ask questions and post responses, but about scientific concepts instead of products.

Perhaps the most interesting feature of HUBzero is the way it handles simulation and modeling programs, or "tools." A little like SourceForge.net, HUBzero allows researchers to work collaboratively on the source code of their simulation programs and share those programs with the community. But instead of sharing only by offering source code bundles to download, HUBzero also offers live published programs available for use instantly and entirely within an ordinary web browser. Computationally demanding runs can be dispatched to remote computing resources in a way that is completely transparent to users. In effect, each hub is an "app store" for a scientific community connected to a cloud of resources for execution, complete with a library of training materials and other collaboration features.

HUBzero brings this functionality together in a package that campus IT organizations can use to create their own online communities or "virtual organizations." HUBzero is not meant to be a closed solution, but rather, an open platform, supported by a consortium of universities, that other institutions can build upon. For more details, see http://hubzero.org

In this tutorial, participants will learn about the features of HUBzero and how existing hubs are being used to support research and educational activities. They will learn how to publish simulation tools and other resources within a hub and track their use by a worldwide community. Specific sessions are:
• What are virtual organizations and how does HUBzero help them work?
• How do the features of HUBzero support various types of projects?
• Tutorial: Uploading tools and other resources to "publish" on a hub

14.00 - 17.30 **Concurrent Afternoon Tutorial A3 - What makes for a successful Indo-US network-enabled collaboration? Development and Enhancement of International Collaborations, especially between Indiana University and Institutions of Higher Education in India**

Tutorial Leaders:
Bradley Wheeler, Vice President for Information Technology, Professor of Indiana University
P. Sarita Soni, Associate Vice President for Research, Professor of Optometry and Vision Science, Indiana University

Drawing on the Indiana University experience on collaboration with partners in India, this tutorial will focus on the arts, business and vision science experiences of developing collaborations and enhancing these collaboration through high performance computing networks. This tutorial will also discuss the use of open access resources to enhance dissemination of knowledge.

In this tutorial, participants will learn how collaborations develop, responsibility of partnerships and how these will be enhanced through high performance computing. There will be an hour long breakout session into five small groups so that the participants can discuss the five presentations in detail with the presenters and bring up new possibilities for collaboration between Indiana University and Indian Institutions of higher education.

Specific topics that will be covered:

• **Business Education** (30 minutes) - Presenters: Professors **M. A. Venkataramanan**, Ashok Soni and Ishwar Murthy
• **Preservation of Collections** (30 minutes) – Presenters: Ruth Stone and Shubha Chaudhuri
• **Education and Research in Vision Science and Eye Care** (50 minutes) – Presenters: Sarita Soni, Akila Ganesan, L. Srinivasa Varadharajan, and Shrikant Bharadwaj
• **Open Access Resources and Dissemination of knowledge** (30 minutes) – Presenter: Brad Wheeler

Following these presentations, there will be 3 breakout sessions for presenters and attendees to further discuss these and other topics.

Appendix 2: Workshop Program

Monday, 6 December

Workshop - day 1

9.00 - 9.10 **Welcome from Indian Hosts**
N. Mohan Ram, Director General, ERNET

9.10 - 9.20 **Lamp Lighting Ceremony**

9.20 - 9.30 **Introduction to the Workshop - background, objectives, and what we want to achieve from the workshop**
Jim Williams, Indiana University & ACE and TP3, Principal Investigator

9.30 - 9.45 Marjorie Lueck, Office of International Science and Engineering, National Science Foundation and introduction of *Video address* and announcement by Subra Suresh, Director, National Science Foundation

9.45 - 9.55 **Welcome from Indiana University**
Brad Wheeler, Vice President for Information Technology and CIO, Indiana University

9.55 - 10.10 **Theme Address**
***Shri. Shashi Kant Sharma**, Secretary, DIT, GOI*

10.10 - 10.20 **Address by Guest of Honour**
***Shri. Kapil Sibal**, Hon'ble Minister for Communications & IT, GOI*

10.20 - 10.30 **Vote of Thanks**
Dipak Singh, Director, ERNET

10.30 - 11.00 **Morning Tea**

Session: "Technology, Cyberinfrastructure, Networks and Advanced Communication Systems

Session Chair: S. Ramakrishnan, former Director General, Centre for Development of Advanced Computing (CDAC)

The aim of this session is to present the network infrastructure that is available to researchers today or will be available to researchers in the relatively near term (1-2 years). This session should prepare the researchers for their sessions regarding the required Cyberinfrastructure for their network-enabled collaborations. These are short talks <10mins each). Reference material (urls, printed material) should be available to supplement the overviews.

11.00 - 11.10 **Current and evolving R&E network infrastructures in India**
***Prof. S.V. Raghvan**, Scientific Secretary to the Principal Scientific Advisor, Govt. of India*

11.10 - 11.20 **India's National Knowledge Network (NKN)**
Dr. B.K. Gairola, Director General, National Informatics Centre (NIC), Govt of India

11.20 - 11.30 **ERNET**
N. Mohan Ram, Director General, ERNET India

11.30 - 11.40 **TBD**
Gulshan Rai, CERT-In, India

11.40 - 11.50 **Current and evolving R&E network infrastructures in the US**
Ana Preston, Internet2

11.50 - 12.00 **Taj - High-Performance Networking Facilitating US-India Science Collaboration**
Greg Cole, GLORIAD/TAJ

12.00 - 12.10 **TransPAC3**
Jim Williams, Indiana University & ACE and TP3, Principal Investigator

12.10 - 12.20 **US National Science Foundation International Network Programs relevant to India**
Kevin Thompson, NSF

12.20 - 12.30 **The World Bank's Role in Catalysing Advanced Networks for Learning and Knowledge Dissemination**
***Michael Foley**, World Bank, South Asia Region*

12.30 - 12.40 **TEIN3**
Dai Davies, DANTE, UK (remote presentation)

12.40 - 12.55 **Identifying key issues, challenges, obstacles, opportunities associated with the new network infrastructures**

Panel Session and participant Interactive process led by session chair - S. Ramakrishnan

12.45 - 14.00 **Lunch**

14.00 - 14.05 **Session: Geosciences, Climate Change, Weather Prediction and Observing Systems**
***Session Chair:** Chaitan Baru, San Diego Supercomputer Center (SDSC), University of California, San Diego*

Overview by session chair including collaboration and funding opportunities

14.05 - 14.25 **Fostering India and US Geoscience Collaboration through iGEON: A cyber infrastructure for Geo Scientists**
Dr. Rajeev Wankar, Department of Computer and Information Sciences, University of Hyderabad

14.25 - 14.45 **Environmental Networks and Cloud Computing: Opportunities for Collaboration**
Chaitan Baru, San Diego Supercomputer Center (SDSC), University of California, San Diego

14.45 - 15.05 **Potential for expanded exchange of earth observation satellite data for climate, weather and environmental studies**
Chris Elvidge, National Geophysical Data Centre, NOAA

15.05 - 15.25 **Satellite and ground observing systems for climate change studies –requirement for network enabled research collaboration**
K.V.S. Badarinath, Head of the Atmospheric Sciences Division, National Remote Sensing Agency, Hyderabad

15.25 - 15.50 **Afternoon Tea**

15.50 - 15.55 **Session: Astronomy, Astroinformatics and Astrophysics**
***Session Chair**: Dr. Dipankar Bhattacharya, Inter University Centre for Astronomy & Astrophysics(IUCAA)*

Overview by session chair including collaboration and funding opportunities.

15.55 - 16.15 **Virtual Observatory - India**
Prof. Ajit Kembhavi, Inter University Centre for Astronomy & Astrophysics(IUCAA)

16.15 - 16.35 **Astronomy With Massive Data Streams**
Dr. Ashish Mahabal, CalTech (remote)

16.35 - 16.50 **Transient Sources**
Sajeeth Philip, Inter University Centre for Astronomy & Astrophysics(IUCAA)

16.50 - 17.00 **Identifying key issues, challenges, obstacles, opportunities from Day 1 for action plan**

17.00 - 17.30 **A conversation with Marjorie Lueck, Office of International Science and Engineering, about National Science Foundation collaborative opportunities**

17.30 - 18.30 **Drinks, nibbles, and Demonstrations (Nighttime Lights, Climate Change, National Library of Medicine Resources)**

19.00 - 21.00 **Workshop Dinner - with String Quartet playing from Cleveland, Ohio, US - and performances by local Indian Dance Troupe**

During the dinner, and to celebrate our coming together for this workshop, there will be performances locally by Indian dancers and a remote performance by a string quartet from the Cleveland Institute of Music - a premiere U.S. conservatory. The Vera Quartet is thrilled to be performing one of their favorite pieces for the audience in India. Anne-Sophie LaCharite-Roberge, first violin, is a Senior from Quebec, Canada. Michelle Abraham, second violin, is an Artist Diploma student from Connecticut. Caitlin Lynch, viola, is an Artist Diploma student from Oregon. Katie Tertell, cello, is a 2nd year Masters student from Virginia. The quartet has been together for two years and participates in the Intensive String Quartet program at CIM

Performing: *String Quartet Number Three in A Major, Op.41; 4th Movement – Finale: Allegro Molto Vivace – Quasi Trio*; by Robert Schumann (1810-1856)

Tuesday 7th

December

Workshop - day 2

Morning — **Session: Cyberinfrastructure for Medical Research**
Session Chair: Anil Srivastava

9.00 - 9.10 — **Welcome to day 2**
Overview by session chair including collaboration and funding opportunities

9.10 - 9.50 — **caBIG® Cooperation for Cancer Care**
Dr. Hemant Darbari, Executive Director, Centre for Development of Advanced Computing (C-DAC) and
Dr. George Komatsoulis, Deputy Director, NCI Center for Biomedical Informatics and Information Technology (NCI CBIIT)

9.50 - 10.10 — **Complementing Epidemiological and Clinical Research with Bioinformatics Applications**
***Dr. Sheela Godbole**, National AIDS Research Institute (NARI), Indian Council of Medical Research*

10.10 - 10.30 — **National Library of Medicine Resources for Researchers**
Hemant Shah, MD, Senior Research Medical Informatician at the Henry Ford Health System

10.30 - 11.00 — **Morning Tea**

11.00 - 11.05 — **Session: Bioinformatics**
Session Chair: Prof. Indira Ghosh, Jawaharlal Nehru University (JNU), Delhi
Overview by session chair including collaboration and funding opportunities

11.05 - 11.25 — **Operonic structure of key metabolic pathways in cyanobacteria**
Prof. Pramod P. Wangikar, Department of Chemical Engineering, Indian Institute of Technology Bombay

11.25 - 11.45 — **Cyber-infrastructures for Genetic Design Automation**
Dr. Jean Peccoud, School of Biomedical Engineering and Sciences, Virginia Tech

11.45 - 12.05 — **Selection of data features for machine learning in the prediction of genome-wide protein:protein interactions**
Dr. Shekhar C. Mande, Centre for DNA Fingerprinting and Diagnostics, Hyderabad

12.05 - 12.25 — **Systems Biology, Global Analysis, and Biomarker Target Identification**
Prof. Mathew J. Palakal, School of Informatics/Department of Computer & Information

Science, Indiana University-Purdue University Indianapolis, Indiana, USA

12.25 - 13.30 **Lunch**

13.30 - 13.55 **Session: Nanotechnology and High Energy and Computational Physics**
Session Chair: Dr. P.S. Dhekne, Bhabha Atomic Research Centre(BARC) Mumbai
Overview by session chair including collaboration and funding opportunities

13.35 - 13.55 **Computational Nanoelectronics: Current Status & Trends**
Prof. V. Ramgopal Rao, Electrical Engineering Department, IIT Bombay

13.55 - 14.15 **Atomistic Modeling and Simulation Tools for Nanoelectronics and their Deployment on nanoHUB.org**
Michael McLennan, Senior Research Scientist, the Rosen Center for Advanced Computing, Purdue University

14.15 - 14.35 **LHC-CMS Tier-2 Facility at the Tata Institute for Fundamental Research (TIFR)**
Prof. Atul Gurtu, TIFR LHC Tier 2 Facility, Mumbai (India)

14.35 - 14.55 **Computing and Networking Needs for Lattice Field Theory**
Prof. Steven Gottlieb, Indiana University

15.00 - 15.30 **Afternoon Tea**

15:30 - 16.30 **Discussion - where to from here?**

A panel led by four scientists, two from India and two from the US, and taking into account the contributions made during the workshop, and drawing on audience participation will debate and deliberate on ways to significantly enhance Indo-US network enabled collaboration. In doing so the panel and audience will try to identify key issues, challenges, obstacles, and opportunities needed for the development of action plans, and identify next steps and future deliverables. The output of this session will form the basis for the second part of the US-India Workshop, tentatively planned for October 2011 in Washington, DC

16.30 - 16.40 **Summing up of workshop**
Jim Williams, Indiana University & ACE and TP3, Principal Investigator & N. Mohan Ram DG, ERNET India

16.40 - 16.45 *Closing Remarks - Dipak Singh, Director, ERNET*

16:45 - 16:50 *Vote of Thanks (**Praveen Misra**, Additional Director, ERNET)*

Appendix 3: List of Participants

From India:

Serial	Name	Surname	Institution	
1.	Santosh	Shukla	National Botanical Research Institute (NBRI)	Participant
2.	Balakrishnan	Athiyaman	NCMRWF, Delhi	Participant
3.	John	George	NCMRWF, Delhi	Participant
4.	Rajesh	Ghorpade	Dr. Panjabrao Deshmukh Agricultural University, Akola	Participant
5.	Vinod	Kumar	Directorate of Rapeseed Mustard Research	Participant
6.	Nirmal	Bisai	Institute for Plasma Research	Participant
7.	Sarah	Ponrathnam	IUCAA, Pune	Participant
8.	Bharat	Somkuwar	Institute of Bioresources & Sustainable Development	Participant
9.	Thanudas	B	Indian Space Research Organization	Participant
10.	Jayaraman	Somasundaram	Indian Institute of Soil Science (IISS), ICAR,	Participant
11.	Amit	Saxena	CDAC, Mumbai	Participant
12.	Sreekantha	B A	CDAC, Mumbai	Participant
13.	Jayanta	Parial	CDAC Kolkata	Participant
14.	Saubhik	Datta	CDAC, Kolkata	Participant
15.	A.S.	Kamble	DIT	Participant
16.	Kishor Gupta	Potta	C-DAC	Participant
17.	Bharat	Tiwari	ERNET India	Participant
18.	Meherban	Singh	ERNET India	Participant
19.	Amit	Chaudhuri	CDAC, Kolkata	Participant
20.	Gigi	Joseph	Bhabha Atomic Research Centre	Participant
21.	Biswajit	Sarkar	VECC, Kolkata	Participant
22.	Surojit	Saha	VECC, Kolkata	Participant
23.	Shyam Sunder	Singh Laishram	DOEACC Centre Imphal	Participant
24.	Buboo	Singh Angom	DOEACC Centre Imphal	Participant
25.	Krishna Kumar	Ramani	Medical Research Foundation	Participant
26.	P.K.	De	PGCIL	Participants
27.	Milind	Jagtap	CDAC, Mumbai	Participant
28.	Mangala	N	C-DAC, Bangalore	Participant
29.	Dr. Prahlada	Rao	C-DAC, Bangalore	Participant

30.	Radhakrishna	Pillai. C	VSSC, Trivandrum	Participant
31.	Abhijit	Sarma	Guwahati University	Participant
32.	Rajendra	Joshi	C-DAC, Pune	Participant
33.	Asha	Bhusan	AIIMS, Delhi	Participant
34.	Vasant	Avaghade	C-DAC, Pune	Participant
35.	Dr. Indira	Ghosh	JNU	Speaker
36.	Dr. Ajay	Gupta	University of Delhi	Participant
37.	Nitin	Dawar	University of Delhi	Participant
38.	Subramanian N	Neelakantan	C-DAC, Bangalore	Participant
39.	Rajendran	Balaji	C-DAC, Bangalore	Participant
40.	Muraleedharan	N	C-DAC, Bangalore	Participant
41.	Soumitri	Mishra	ERNET India	Participant
42.	Vipin	Bhatnagar	Panjab University	Participant
43.	O P	Sharma	DIT	Participant
44.	Pradeep	Chopra	DIT	Participant
45.	S. A.	Kumar	DIT	Participant
46.	B. K.	Murthy	DIT	Participant
47.	Tushar Kanti	Roy	MTNL, Delhi	Participant
48.	Srinivas	S	CSIR - Chennai	Participants
49.	P.S.	Joshi	Bhabha Atomic Research Centre	Participants
50.	Praveen	Misra	ERNET Delhi	Participants
51.	Dipak	Singh	ERNET Delhi	Participants
52.	Rama	Pandey	ERNET Delhi	Participants
53.	Bharti	Saluja	ERNET Delhi	Participants
54.	Uma	Shankar		Participants
55.	Majinder	Singh	Presto System	Participants
56.	Shivam	Verma	Presto System	Participants
57.	Ashwin	Mittal	Wipro	Participants
58.	Prashant		Wipro	Participants
59.	Padam		AGC Networks	Participants
60.	Rajesh		AGC Networks	Participants
61.	Geeta	Kathpalia	DIT	Participants
62.	Atul	Gurtu	TIFR	Speaker
63.	V	Ramgopal Rao	IIT Bombay	Speaker
64.	Sanjeeth	Philip	IUCAA	Speaker
65.	Anil	Sagar	DIT	Speaker
66.	Ishwar	Murti	IIM, Bangalore	Speaker
67.	Anil	Kumar	Extreme Network	Participants
68.	P.S.	Dekne	BARC	Speaker
69.	Andrew	Lynn	JNU	Speaker

70.	Prasad	Dharmavaram	VMware India	Speaker
71.	Jyothi	Padmanabhan	Yahoo India	Speaker
72.	Prof. Ajit	Kembhavi	IUCAA	Speaker
73.	Dr. Subrata	Chattopadhyay	C-DAC, Bangalore	Speaker
74.	B.	Asvija	C-DAC, Bangalore	Speaker
75.	Dr. Sheela	Godbole	NARI, ICMR	Speaker
76.	Dr. S.	Ramachandran	Institute of Genomics & Integrative Biology, India	Speaker
77.	Dr. Hemant	Darbari	C-DAC, Pune	Speaker
78.	Madhulika	Tripathi,	Biomantra	Speaker
79.	N. Mohan	Ram	ERNET Delhi	Speaker
80.	Prof. Rajeev	Wankar	University of Hyderabad	Speaker
81.	KVS	Badarinath	NRSA	Speaker
82.	Dr. Dipankar	Bhattacharya	IUCAA	Speaker
83.	Sorav	Bansal	IIT Delhi	Speaker
84.	Akila	Ganesan	Medical Research Foundation	Speaker
85.	Paventhan	Arumugam	ERNET India	Speaker
86.	L. Srinivasa	Varadharajan	Medical Research Foundation	Speaker

Fellowship Registration:

Serial	Name	Surname	Institution
87.	Sunil	Kumar	School of Information Technology, PAU, Ludhiana
88.	Sivakumar	Belliraj	Sri Sathya Sai Institute of Higher Learning
89.	Abhilash	Gangadharan	Jawaharlal Nehru University
90.	Anmol	Hemrom	Jawaharlal Nehru University
91.	Kunal	Sharma	Jawaharlal Nehru University
92.	Rashmi	Kumari	Jawaharlal Nehru University
93.	Shawez	Khan	Jawaharlal Nehru university
94.	Swati	sinha	Jawaharlal Nehru University
95.	Roopa	Singh	Jawaharlal Nehru University
96.	Pankaj	Narang	JawaharLal Nehru University
97.	Nisha	Chandran	Jawaharlal Nehru University
98.	Reema	Singh	Jawaharlal Nehru University
99.	Swapna	Menon	Jawaharlal Nehru University
100.	Srikant	Verma	Institute of Genomics and Integrative Biology
101.	Shreya	Chakraborty	Institute of Genomics and Integrative Biology
102.	Rahul	Mandal	Institute of Genomics and Integrative Biology
103.	Gawri	Sriniwasa	PES School of Engineering

US and other Participants

Serial	First Name	Family or Surname	Institute	Country
104.	William	Barnett	Indiana University	USA
105.	Andrew	Kuhn	Pfizer Global Research	USA
106.	Ganesh	Shankar	IUSCC	USA
107.	Michael	McLennan	Purdue University	USA
108.	Steven	Gottlieb	Indiana University	USA
109.	Sarita	Soni	Indiana University	USA
110.	Ashok K	Soni	Indiana University	USA
111.	Radha	Nandkumar	NCSA - UIUC	USA
112.	Sudhakar	Pamidighantam	NCSA	USA
113.	William	Barnett	Indiana University	USA
114.	George	McLaughlin	Indiana University	Australia
115.	George	Luis	US National Solar Observatory	USA
116.	Frank	Hill	US National Solar Observatory	USA
117.	Chris	Elvidge	NOAA	USA
118.	Ruth	Stone	Indiana University	USA
119.	Dr. Jean	Peccoud	Virginia Tech	USA
120.	James	Williams	Indiana University	USA
121.	Yasuichi	Kitamura	NICT Japan	Japan
122.	George	Komatsouli	National Cancer Institute, NIH	USA
123.	Kevin	Thomspson	OCI NSF	USA
124.	Brad	Wheeler	Indiana University	USA
125.	Ian	Fore	NCI CBIIT	USA
126.	Anil	Srivastava	NCI CBIIT	USA
127.	Michael	Joseph Foley	World Bank	USA
128.	Mathew	Joseph Palakal	IUPUI	USA
129.	Takatoshi	IKEDA	NICT	Japan
130.	Jin	TANAKA	KDDI	Japan
131.	Gregory	Scott Cole	Taj Program	USA
132.	Venkataramanan	M	Indiana University	USA