Internship Program 2010
IBM Research - India

IBM Research - India invites applications for the 2010 Internship Program from May for students interested in all areas of Computer Science and related fields at our locations in New Delhi, Bangalore and Hyderabad. We are seeking highly motivated undergraduate and post graduate students, who are interested in experiencing an exciting period of research. The selected students will have the opportunity to work closely with an outstanding research team on challenging problems that range from leading-edge exploratory work to prototyping real-world systems. During the internship, the students will also have the opportunity to participate in the dynamic technical environment of the largest Industrial Research organization in the world and network with other top students from many different fields and universities. We offer internship positions in following research groups:

- High Performance Computing
- Information Management
- IT Service Management
- Next Generation Systems & Smarter Planet Solutions
- Optimization, Data Mining, Text and Speech Analytics
- Services Software Engineering
- Services Science
- Telecom Research

Interns bring fresh ideas and perspective to the lab and help us conduct world-class research and thus create an impact. IBM Research provides an environment where interns experience a world class industrial research setting. We publish heavily at the top conferences & journals and encourage our interns to publish the great work they do with us as well. Our association goes beyond the internship and we always look at them as our future hires.

We accept interns at various levels in their academic life — from undergraduate students to PhD candidates — and from all over the world

Program timelines: (these a flexible)

Application Deadline:	Flexible
Selection Process:	Flexible
Offer Processing:	March, 2010
Internship program:	Flexible program of 3 – 4 months anytime from May until November

Applications along with your latest CV and the application form can be sent to Laura Friebel at friebell@unimelb.edu.au. Program enquiries can be direct to Ms Friebel on 8344 9335

Detailed overview of IBM's research groups and skills required for the projects is given in below.

IBMR

Internship Program 2010
IBM Research - India

High Performance Computing

The High Performance Computing (HPC) group at IRL is engaged in the design and analysis of cutting edge parallel programs and in improving the performance of engineering, scientific, and business applications on high performance platforms such as the IBM Blue Gene Supercomputer and Power-processor based clusters. The group is focused on areas based on performance on multi-core processors, performance on large-scale supercomputers, performance on clusters, medical imaging applications and parallel scalable algorithms for supercomputers. Examples of problems that we are addressing include:

- Optimization of benchmarks such as FFT and Transpose for novel supercoumputing (petaflop) architectures

- Optimization of HPC applications in the areas of molecular dynamics and quantum chemistry
- Algorithms for MPI collectives on large-scale Supercomputers/clusters.

- Parallel Algorithms for solving large, possibly ill-conditioned sparse matrices

Skills: Good knowledge of C, algorithms and data structures; Principles of parallel programming and computer architecture.

Level: PhD/M-tech/MS/B-tech

Location: Internship positions are available at Delhi.

Information Management

The Information Management and Analytics Group at IRL is focused on developing next-generation technologies in various areas such as advanced business intelligence and insight generation, context-oriented information integration, and extraction of semantic knowledge from unstructured data. These technologies are driven by IBM Research's goal of building intelligent solutions and services to address business problems in various industrial sectors, including financial, telecommunication, retail, and healthcare, among others.

We bring together the capabilities of information integration and data analytics to build next-generation integrated enterprise information management systems. This would encompass techniques for context extraction at the time of data upload and building new interfaces for on-demand access to information by enhanced business driven search and dynamic faceted browsing. We are also exploring the value of incorporating text data in various predictive analytic models for customer lifetime value (CLV), churn prediction, and targeted marketing.

The Information Management and Analytics team develops novel techniques for loosely-coupled structured and unstructured data through symbiotic and semantically-disambiguated information in an enterprise. This is achieved by viewing the structured data in the relational database as a set of predefined "entities" and identifying the entities from this set that best match a given document.

IBMR

We also focus on information extraction (IE) from unstructured data where we develop technologies, which involve the identification of entities such as organizations, places, product names and relationships among entities such as sellers and employees. To address the need for scalability in IE systems, we are developing innovative techniques that work on the inverted index of document collections. Some of the research challenges in this domain include techniques to deal with extremely noisy data (such as SMS, instant messenger logs, e-mail and automatically transcribed conversational data) and modeling and maintaining uncertainty and conflicts associated with information extraction.

We are seeking interested Ph.D./Master/BTech students with Database background to spend an existing research internship at our lab.

Location: Internship positions are available at Delhi and Bangalore.

IT Service Management

The IT Service Management Research team explores aspects around managing IT more efficiently. Our main focus is around three topics: First, IT optimization is the discipline of predicting how to optimally utilize resources to drive higher value. Our current focus is on analytics for data center consolidation and Green IT. Second, IT Systems Management explores new techniques to automate the management of systems. Our current focus is on virtualization and cloud computing management. Finally, IT Service Delivery focuses on the methods and tools to drive higher quality in the delivery of IT services, especially in Global Delivery. Our current focus is on business process management and related methods applied to various problems spaces around remote infrastructure management in global delivery centers.

We are seeking interested student in these three areas with the following pre-requisites:

1. IT Optimization: we are looking for students who have an interest in algorithms and are building usable tools to make these algorithms available to real world practitioners. Programming skills in Java or related languages are a must. B-Tech/BS or M-Tech/MS or PhD students are all welcome

2. IT Systems Management: we are looking for students with some knowledge or interest in virtualization and cloud computing. Fundamental knowledge of networks or storage would be helpful. Strong programming skills in Java required. Furthermore, familiarity with systems level concepts will be helpful. Advanced BS/B-Tech, MS or PhD students welcome.

3. IT Service Delivery: we are looking for students with an interest in data or process mining or business process management, both from a modeling as well as software engineering perspective. Strong programming skills are required. Advanced MS/M-Tech or PhD students are welcome to apply.

Location: Internship positions are available at Delhi and Bangalore.

IBMR

Next Generation Systems and Smarter Planet Solutions

The next generation systems and smarter planet solutions group explores 4G wireless infrastructure solutions, smarter planet infrastructure and solutions on next generation many-core hybrid systems. A broad category of existing projects and pre-requisites are:

I)   **Category:** Next Generation Telecom, Enterprise Service Infrastructures

   Pre-requisites:
   1. Wireless and IP communication network fundamentals
   2. Familiarity with network simulation and modeling tools
   3. Solid programming skills in any one high level language
   4. Parallel and many-core programming/architecture/design (a plus)

II)   **Category:** Intelligent public/private infrastructure management (smart city, smart water) / platforms/kernels for next generation many-core hybrid systems

   Pre-requisites:
   1. Data Mining and Machine Learning Techniques
   2. Solid programming skills in any one high level language
   3. Parallel and many-core programming/architecture/design (a plus)

III) **Category:** Smart Energy

As the world population is exploding, Energy is becoming one of the major problems around the world. Developed countries are not able to keep with the demand and emerging countries are not able to provide any access to energy to their people. To make matters worse, increasing energy production using conventional methods increases environmental pollution. To address this important problem, we are working on a wide variety of projects from context aware demand response systems to synchrophasor networks to electric vehicle modeling.

   Pre-requisites:
   1. Power systems
   2. Demand response
   3. Energy economics
   4. Power-aware computing
   5. Alternate energy sources

Skill Set/Level of Students:
Innovative B-Tech/BS, M-Tech/MS or PhD students, who can identify interesting problems, propose solutions and build models/simulations/prototypes to evaluate their solutions.

Location: Internship positions are available at Delhi and Bangalore.

IBMR

Internship Program 2010
IBM Research - India

Optimization, Data Mining, Text and Speech Analytics

The Services Information and Analytics group at IRL is involved in projects related to business optimization, predictive modeling, business intelligence, machine learning, natural language processing, information retrieval, information extraction, machine translation and speech analytics. We deal with information of various types including unstructured and noisy text, documents, e-mail, ticketing systems, on-line databases, team rooms, transcribed calls, and IT monitoring systems to generate optimal plans or insights that will improve the quality of service delivery. Examples of problems that we are addressing include:

- **Optimization**
    - Workforce Management – Decision support systems for assigning people to projects in global service delivery organizations
    - Emergency Management - Strategic and tactical resource deployment planning for managing emergency events
    - Portfolio Optimization – Large-scale stochastic programming modeling and optimization for asset and liability management problems.
- **Data Mining and Predictive Modeling**
    - RDMS and high volumes of data – exploring possible technology barriers of current RDBMS when managing terabytes of data
    - Targeting customers for promotions based on demographic and multi-channel interaction data
    - Portfolio risk modeling
- **Speech Analytics**
    - Search on audio data
    - Evaluating articulation and syllable stress to assess the quality of speech production
    - Automatically detecting and encrypting privacy information in audio files
    - Deep understanding of call center interactions (text/speech) for quality/monitoring purposes

- **Text Analytics and Machine Translation**
    - Statistical Machine Translation of documents from one language to another (e.g. from Hindi/Urdu to English)
    - Automatically extracting problem resolution information from online technical bulletin boards (e.g. SAP forums)
    - Extracting skills, skill levels and other structured information from resumes and job descriptions

Skills: Trained in one or more of the following areas: optimization, data mining, machine learning, text mining, natural language processing, machine translation and information retrieval. For internships in the optimization area, familiarity with techniques such as linear & nonlinear optimization and dynamic programming as well as tools/environments such as MATLAB, Arena, AMPL/CPLEX is desirable. In addition, knowledge of Java/C++ is highly recommended.

Level: PhD/M-tech/MS

Location: Internship positions are available at Delhi and Bangalore.

IBMR

Services Software Engineering

Our technical agenda is driven by two key trends. First, componentization and standardization are accelerating construction, and evolution of business solutions in a service-oriented fashion. Second, service providers are increasingly adopting a globally distributed paradigm for the entire lifecycle, for efficient, scalable and high-quality delivery of business solutions. Our ongoing research efforts focus on model-driven technologies for delivery of high-quality, industrial-strength business solutions, throughout their lifecycle, in a globally distributed fashion.

We broadly categorize our existing projects in following areas

Model Driven Solution Engineering: Businesses in today's flatter world are becoming collaboration-oriented both internally and externally. Deriving business benefits by operational agility and efficiency, superior supply chain management, faster time to market are now imperatives for addressing new market opportunities and competitive threats. There is also a significant momentum behind standardization across industries, enabling enterprises to componentize their business functions, underlying business processes and IT elements. We are investigating methodologies and tooling to synthesize service-oriented enterprise solutions, given business functions and associated transformation objectives. We are conducting research in formal models to explicitly define the structure and behavior of a business system, and leverage these models as "a recipe for construction" of business's IT systems. In this endeavor of closing business to IT gap, we are addressing several practical challenges associated with Model Driven Architecture (MDA). Key ones are model-driven composite application building and deployment, meta-models for incorporating business rules and policies, meta-models for human-interactions, change management issues, and acceleration of early lifecycle (e.g., blueprinting) via collaboration and reuse (of documents, diagrams, and models).

Distributed Development & Delivery: For delivery of service-oriented solutions, leading service providers have adopted the multi-site development paradigm. Multi-site development, however, bring with it several challenges, such as inadequate communication between teams, lack of information about remote sites, differences in processes, etc., which disrupt the intrinsically collaborative nature of software development. We have explored both methodological and tool-based approaches to address the challenges of multi-site requirements management. We believe that compared to collocated development, multi-site projects require more investment upfront in precisely communicating requirements to remote teams. We prototyped a tool called EGRET (Eclipse-based Global Requirements Engineering Tool) which extends conventional requirements management tools with rich support for formal and informal collaboration, along with knowledge-management capabilities. We also proposed an ontology-based-approach for seamlessly integrating various SDLC (software development lifecycle) tools that may be deployed in different multi-site teams.

We are now working to help streamline global delivery by leveraging proven IBM processes and best practices in a standardized automation environment. This is achieved through the notion of a self-contained work request, using which each work order is authored, transported, acted upon, and managed, as per the adopted delivery process, with rich traceability to issues, risks, defects, etc., thus providing for effective governance. We have designed algorithms for work-request scheduling, created tool adopters to seamlessly share

IBMR

work packets in a heterogeneous tool environment, and are exploring work assignment and capacity management solutions for remote delivery centers to efficiently execute the work packets. We are also investigating analysis of SDLC artifacts for providing 'business intelligence' to software development and delivery (the current focus on time and resource management).

Quality: As service-oriented solution construction, evolution, and delivery gains momentum, we anticipate that new challenges in maintaining the overall software quality will become commonplace. We are investigating techniques for making static-analysis-based bug-finding techniques consumable by developers. Static analysis tools are yet to enjoy wide adoption in development processes, mainly because of the large number of false positives that the tools generate, and a low rate of true positives. Moreover, for the true positives, developers are often not interested in learning about some of the bugs. Therefore, in addition to improving the accuracy of the analysis, assisting developers in identifying and understanding the bugs that they consider worthy of investigation is important too. We have developed new inter-procedural analyses to reduce false positives, while also increasing the true-positive detection rate. We are also developing automated techniques for identifying patterns of bugs that can help the developer in understanding and fixing a bug.

Knowledge Extraction: Evolution (or maintenance) of software constitutes a significant fraction of IT spending in IT-enabled enterprises. To enable software evolution for business transformation, we have developed analysis techniques and tools for recovering high-level logical models from legacy applications. Recovered logical models contain high-level abstractions of the data and functionality in the application. We have developed capabilities to reverse-engineer two kinds of logical models: process models and data models. A logical process model shows the ordering on a given set of interesting events occurring in an application. A logical data model is a value-added model of the data declarations in an application (resembling a UML class diagram) obtained by a semantic analysis of the code that uses the declared data. We are applying these in the financial domain where there is an increased focus on legacy evolution for multi-channel integration, customer relationship management, compliance, and also for estimating transformation effort etc.

Parallel Programming: Current object-oriented languages have revealed several drawbacks with respect to parallel (concurrent) programming at the level of unstructured threads with lock-based synchronization. IBM Research is developing X10, a modern object-oriented programming language designed for high performance with explicit programmer defined parallelism for realizing high productivity programming of parallel computer systems. The key features of X10 include explicit reification of locality in the form of places, support for a partitioned global address space (PGAS) across places, and lightweight activities embodied in *async, future, foreach*, and *ateach* constructs which subsume communication and multithreading operations in other languages. Our current focus is on static program analysis (e.g., May-Happen-in-Parallel analysis, Bad Place Analysis), compilation for C/C++, debugging for X10, and assessment and semi-automated migration of domain-specific serial code to emerging multi-core architectures, leveraging productive programming models and their variants (such as OpenMP, OpenCL).

Skill Set/Level of Students: We are seeking applications from B-Tech/M-Tech/MS/Ph.D students in Computer Science (or related field). Candidates with knowledge of one or more in empirical software engineering, metrics and estimation, software architecture, mining and

IBMR

analysis of software engineering artifacts for end-to-end SDLC, and collaborative software engineering for distributed development are particularly encouraged to apply. A strong interest in team-oriented industrial research and building large-scale software systems is a must. As an intern in Service Software Engineering group you will also have the opportunity to collaborate with some of the best minds in IBM, including researchers from IBM's world-wide research labs, work with India's premier research organizations, and deploy your methods and tools to the vast network of IBM India practitioners.

Location: Internship positions are available at Delhi and Bangalore.

Service Science

Collaboratory for Service Science
Global markets are increasingly becoming service-based economies. This is mainly because services are becoming important to traditional product companies. Traditional product companies are becoming major service organizations. In the competitive world, innovation around Services is needed to maintain differentiation in the market and hence, the profits. Service Science is the systematic study of organizations and networks that deliver services. This study would help to drive the innovation in improving various aspects of services including design, productivity, quality and scalability. By nature, Service Science is multi-disciplinary research that integrates various aspects of established fields like computer science, operations research, industrial engineering, business strategy, social and cognitive sciences, and legal sciences. **IBM Research – India** has recently started a Collaboratory for Service Science (CoSS) at the Indian School of Business at Hyderabad in pursuit of making fundamental contributions to the field of Service Science.

We invite application from Masters and PhD students in various disciplines in management and industrial engineering for a summer internship at CoSS. Students can apply with their CV and a brief write up on their current research. The disciplines of our interest include Organizational Behavior, Strategy, Operations, Economics, Information Systems and Services Marketing.

Example research problems of interest are:
- What are the individual and team behavior characteristics that are important to a service organization? How do they depend on the nature of service that is being delivered?
- How do you measure innovation in a service organization?
- How to analyze and optimize the performance of a service system?
- How do you optimize the resource utilization while increasing/maintaining the performance of a service organization?
- What are the best strategies to evolve workforce in a service organization?
- What are the new models of globally distributed service organizations?
- How do you characterize service quality in a B2B scenario?

Students would get an opportunity to interact with researchers at IBM and ISB, and work on real research problems originating in IBM Services Organizations.

Location: Internship positions are available at Hyderabad.

IBMR

Services Engineering

The service engineering group explores aspects of service design, service operations, and service management. A broad category of existing projects and pre-requisites are:

I)   **Category:** Service Design Tooling for Public Sector Services
        Pre-requisites:
        1. Familiarity with UML, EMF, and Eclipse plugins
        2. Solid programming skills in Java
        3. Interest in service design issues, human computer interfaces

Skill Set/Level of Students:

Innovative B-Tech/BS, or M-Tech/MS students who wish to participate in the design and implementation of service design tooling specifically for services in the public sector.
The modeling of a specific service or family of services in the public sector would be a possible outcome of the internship.

II)   **Category:** Service Simulation

The operation of services is an economically important subject – which has a large literature in the area of call centers but less in the area of more traditional services. We are working on a generic model of service and looking to investigate a variety of simulation paradigms which can be used to optimize service delivery and reduce the risk of service implementations.

        Pre-requisites:
        4. Knowledge of queueing theory a distinct advantage
        5. Interest in agent based simulation (a plus)
        6. Solid skills in object oriented design and Java programming skills

Skill Set/Level of Students:

Innovative M-Tech/MS or PhD students who can identify interesting problems, propose solutions and build models/simulations/prototypes to evaluate their solutions.

Location: Internship positions are available at Bangalore.

Telecom Research

A center that specializes in real life innovations, IRL is heavily focused on developing solutions for specific industries. The goal is to create groundbreaking applications and platforms that would help businesses and industries thrive. One of the key focus areas for IRL is the telecommunications space and it has established a Telecommunication Research and Innovation Center (TRIC) to address challenges in this industry.

TRIC is a core research group at IRL that executes innovative projects focused on the telecom industry. It is one of our industry-specific initiatives designed to tap relevant research technologies and expertise for collaborative projects with IBM clients and business partners. TRIC serves as both a global resource for the telecom industry as well as a conduit for telecom service providers into IBM Research at large. The key areas of TRIC research include:

IBMR

Internship Program 2010
IBM Research - India

1. Enabling IT for Emerging Economies utilizing the Mobile Platform
2. Telecom Analytics and Business Intelligence
3. Telecom Middleware
4. Application Development on Mobile Devices
5. Applications of Mobile in other industries for example Mobile Retail, Mobile Finance, etc.

Skills: Knowledge of one or more of Mobile Devices, Distributed Systems, Networking, Data Mining, Speech Processing; java/C++ for prototyping.

Level: PhD/M-tech/MS/B-tech

Location: Internship positions are available at Delhi.

IBMR