HARSHAD DESHMUKH

Phone: 267-240-6354

http://pages.cs.wisc.edu/~harshad
hbdeshmukh@google.com

EDUCATION

- **PhD**, Computer Sciences (Database Systems)　　　　　　　　　(Aug 2012 - Aug 2018)
  – Advisor: Prof. Jignesh M. Patel
  – University of Wisconsin - Madison
- **Master of Science**, Computer Sciences　　　　　　　　　　　(Aug 2012 - May 2014)
  – University of Wisconsin - Madison
- **M.Sc.** Mathematics and **B.E.** Computer Science　　　　　　　(Aug 2006 - Jun 2011)
  – Birla Institute of Technology and Science (BITS), Pilani, India

PUBLICATIONS

- **Adaptive Concurrent Query Scheduler For Quickstep Database System** (IEEE BigData Congress 2017)
  – Developed an adaptive, and responsive database scheduler for concurrent queries
  – The scheduler separates resource sharing policies from mechanisms using a probabilistic framework
  – Demonstrated the effectiveness of policies such as fair, highest priority first and proportional priority
  – Winner of the best student paper award

- **Quickstep: A High Performance Data Platform based on a Scaling-Up Approach** (VLDB 2018)
  – Described the Quickstep data platform developed with a scaling-up design philosophy
  – Showcased the adaptive nature of the query scheduler by flexibly changing CPU resources while the query is in progress
  – Demonstrated query progress monitoring to get statistics such as degree of parallelism of relational operators and their execution times
  – **Contributions:** Co-author with Quickstep colleagues, designing and performing experiments, profiling query execution, and optimizing query execution performance

- **Data Partitioning for In-Memory Systems: Myths, Challenges, and Opportunities** (CIDR 2019)
  – Studied the problem of data partitioning in realistic settings
  – Presented a partitioning benchmark to facilitate future studies in this area

- **Revisiting Pipelining for In-Memory Database Systems** (Under submission)
  – Revisiting the notion of pipelining in the in-memory settings
  – Compare pipelining and non-pipelining performance through empirical and analytical methods

- **Survey of Scheduling in Analytic Database Systems** (Under submission)
  – A comprehensive survey of how scheduling is performed in various database systems
  – Presented a taxonomy of scheduling and classification of systems under this taxonomy

- **Determining Data Locality in a Distributed System using Aggregation of Locality Summaries**
  – Patent pending: US Patent App. 15/249,138, 2017.

Professional Experience

- *Software Engineer*, Google, Madison (Since Aug 2018)
  – Helping advance Google's compute infrastructure
  – Working with Google's networking team to deliver a highly performant and reliable network

- *Research Assistant*, Database Systems Group, UW - Madison (5.5 years from Jan 2013)
  – Worked on Apache (incubating) **Quickstep** database system with Prof. **Jignesh M. Patel**
  – Designed and built a dynamic **Query Scheduler** for multi-threaded execution of relational queries

- *Research Intern*, CISL, Microsoft, Sunnyvale, California (3 months from May 2017)
  – Worked on designing **self regulating Stream Processing systems** that can take smarter scheduling decisions using **Machine Learning** techniques
  – Collaborations with the Microsoft Dhalion project team and Apache (incubating) Heron project team

- *Intern*, Pivotal Inc., Palo Alto, California (3 months from June 2015)
  – Joined as a part of the acquisition of Quickstep Technologies by Pivotal Inc.
  – Continued development of Quickstep system working jointly with the employees of Pivotal Inc.

- *Research Intern*, Samsung Semiconductors, Milpitas, California (3 months from May 2014)
  – Worked on optimizing **SSDs** (Solid-state drives) for **Data Management Systems**
  – Studied intricate details of SSD functionalities and applied them for efficient data management
  – Precise details not mentioned due to confidentiality agreement

- *Engineer IT*, Cisco Systems, Bangalore, India (11 months from Jul 2011)
  – Developed **Web Services** based on **Service Oriented Architecture** methodology
  – Used Apache CXF framework, J2EE and SQL for back-end development
  – Contributed in developing an Ontology application using SPARQL, Oracle RDF data store

- *Student Intern*, RSA Security at EMC, Bangalore, India (6 months from Jan 2011)
  – Worked on Digital Certificate importing modules in **RSA Data Protection Manager**
  – Implemented Certificate Management Protocol in **Java** using RSA BSAFE APIs

Awards and Recognitions

- **Best student paper** at IEEE BigData Congress 2017, Honolulu, Hawaii
- Multiple **Cisco Star** and **Spark** awards for excellent technical contribution at Cisco Systems
- Ranked in top 30 students in **Indian Mathematics Olympiad** from Maharashtra and Goa states in India
- Certified in **Oracle Database 11g** - Oracle Certification 1Z0-051 on SQL Fundamentals
- **Summer Research Fellowship** from Indian Academy of Science
- **Merit cum Need Scholarship** for consistent academic performance at BITS Pilani

Presentations

- **Introduction to Apache Quickstep and its scheduler**, various venues
  – SRI Labs, Menlo Park, California
  – Hortonworks, Santa Clara, California
  – Meetup organized by H2O.ai in Mountain View, California
  – SCI Labs, Computer Sciences, UW-Madison

- **Dynamic Rerouting in Vehicular Ad-Hoc Networks** at **IBM Research Lab**, New Delhi
  – Proposed a Dynamic Rerouting algorithm using Ant Colony Optimization to beat traffic congestion

- **Road Traffic Simulation using Cellular Automata** and **Swarm Intelligence** at APOGEE
  – APOGEE is the annual technical festival held in BITS Pilani
  – Simulated road traffic scenarios in Python based GUI.

ACADEMIC PROJECTS

- **Complex Event Processing**
  – Analyzed the scalability of and developed SQL like queries for Storm from Twitter
  – Evaluated the throughput and memory footprint of the execution

- **Indexing and Regular Expression Search for Text Documents in *Quickstep***
  – Facilitated bulk loading of text documents and construction of an inverted index
  – Implemented keyword and regular expression search for the documents

- **Database System Implementation using XML**
  – Analyzed the merits of XML in representing data in an object-relational model and using a generic model for storing object-relational data
  – Designed a new database system's concurrency and crash recovery management features

- **Simulation of Traffic Flow using Cellular Automata Model**
  – Studied various Cellular Automata models for road traffic
  – Designed simulations using Python based GUI to demonstrate rules given by Cellular Automata about acceleration, velocity and displacement for vehicles

SKILL SETS

Programming languages	C++, Python, Shell scripting, Java, C
Data management	MySQL, MonetDB, HBase, LevelDB
Development environments	Vim, CLion, Jupyter
Web technologies	JSP, JavaScript, CSS/HTML, PHP, ASP.NET, Django
Publishing	LaTeX, Microsoft Word, Open Office

EXTRA CURRICULAR

- **Seminar Coordinator**, Database Group, UW Madison (`http://database.cs.wisc.edu/seminar.html`)
- **Moderator**, BITS2MSPhD (`http://groups.yahoo.com/group/bits2msphd/`) - A forum for BITS students interested in Masters/PhD
- **Member of Steering Committee**, Award for Excellence in Computer Science - An award for undergraduates in BITS Pilani (`http://aecs.bits-csa.org/`)
- **Coordinator**, Communication Cell, BITS Embryo (`http://embryo.bits-pilani.ac.in`)
- **Vice President**, Maharashtra Mandal - a cultural group in BITS Pilani
- **BITSAA International Leader** - for BITS Alumni Association (BITSAA)