

Dhruv Matani

700 Health Sciences Drive, Chapin J 2142 Bx, Stony Brook, NY 11790 (USA)

Landline: (631) 403-7653

<http://dhruvbird.com/>

dhruvbird@gmail.com

Systems Experience

Programming Languages:	C, C++, Object Pascal, Basic
Scripting Languages:	Python, Javascript, PHP, bash, Perl, VBScript
Presentation technologies:	HTML, LaTeX
Platforms:	Linux, MS-DOS, Windows
DBMS:	MySQL, SQLite, PostgreSQL, Oracle
RDBMS Language:	SQL

Educational Qualifications

Masters in Computer Science at Stony Brook University (Currently Pursuing; August 2011 - present). Expected Graduation: Fall 2012.

Subject taken in Fall 2011

Analysis Of Algorithms

Computational Biology

Project: **Genome (diff) Compression** - Finding and storing the difference between 2 human genomes stored as FASTA files. Achieved *139x - 895x* 2012 compression ratios for various genome pairs. i.e. *3GiB* is compressed to between *22MiB* & *3.3MiB* ([link to paper](#))

Operating Systems

Project: **Process Migration** - Migrating a live process from one machine to another over the network. Implemented a system call *migrate()* that performs process migration in the JOS kernel.

University of Mumbai, B.E. Computer Engineering, DJSCOE, Mumbai: 62.6% June 2006

Work Experience: 5 years

Directi Pvt. Ltd. (Senior Software Engineer)

[Jan 2009 - June 2011]

Authored, tested, and integrated an **XMPP BOSH Server** for use in the talk.to project. <https://github.com/dhruvbird/node-xmpp-bosh>

March 2011 - present

Responsible for **problem-setting and training interviewers** for the IIT graduate hiring process.

Oct - Dec 2010

Update mechanism for .pw desktop chat client. This handles creation, shipping and fail-safe application of differential updates (deltas) which are required to save the user's network bandwidth

Feb 2010 - Feb 2011

Stress testing of user facing web-applications using UI testing tools and Titanium Desktop

August 2010

Setup, administration and question setting for the Directi Online Test (DOT) which is used for recruitment of software developers and engineers.	Oct 2009 - Feb 2010
Developed the BLL layer for administering the chat server and related services	Oct 2009 – April 2010
An application to enable translations of various string data for internationalizing web applications.	May 2009 – Sept 2009
Email Context Analysis Engine(CAE): Developed an email context analyzer to determine the most relevant commercial keywords/phrases for a given email.	Jan – April 2009
Mukesh Patel School of Technology Management & Engineering	[Nov 2008 - Jan 2009]
Taught a course on Systems Architecture & Programming	Nov 2008 – Jan 2009
Taught a course on Operating Systems-II	July - Nov 2008
Calsoft Pvt. Ltd, Pune (Senior Development Engineer)	[July 2006 - April 2008]
File Download Tool: Developed a file download tool which is being used by 250+ employees. This is used by the company to monitor file downloads and allow file sharing (multiple downloads of the same file).	Feb - Apr 2008
HTTP Traffic Generator: Developed an HTTP Traffic Generator to generate different types of HTTP Traffic with different User-Agents, access patterns and varying HTTP Headers for Umber Media systems.	Jan 2008
HTTP Caching Proxy: Developed an HTTP Caching Proxy to cache youtube flash video content for Umber Media systems	Sept - Nov 2007
RTSP Caching Proxy: Developed an RTSP Caching Proxy to cache youtube 3gp video content for Umber Media systems. Also performed re-encoding and down sampling of the video and audio content. This proxy also let you insert advertisements at various places in the original video. It was used as a POC for an ad-serving proxy for video content	July - Sept 2007
DiskImage: Developed the DiskImage tool which is used by ScaleMP for automated installation of linux on machines from a remote TFTP server. This tool allows selecting multiple Disk Images from the remote TFTP server, and choosing a partitioning scheme or using one from a pre-provided list.	April - June 2007
Ingres Optimization: Optimization of the Ingres Database Server's statistics collection module for ScaleMP's VSMP architecture	Feb - June 2007
Log Parser Optimization: Optimization and bug-fixing of the logpar(log parser) application for ScaleMP, which is used for parsing the performance logs for their architecture.	Dec 2006 - Jan 2007
Technology Research Laboratories' Undergraduate Summer Research Internship	[2004]
Multi-thread aware allocators: For concurrent high-response applications that are memory intensive and have many parallelly running threads.	2004

Major Personal Projects

algorithm-js: A collection of Data Structures & Algorithms in Javascript. https://github.com/dhruvbird/algorithm-js	March 2011 - April 2011
lib-face: A very fast auto-complete server; to be used for as-you-type search suggestions. https://github.com/duckduckgo/cpp-libface	October 2010 - Present
pymq: A Python based Message Queuing system which is meant to scale to millions of queues. http://code.google.com/p/pymq/	April 2010
gaeQuery: Query the appengine datastore from your browser. http://gaequery.appspot.com/	February 2010
p2p-fs: p2p-fs is read-only peer-to-peer file-system, meant to be used on high-speed LANs, or a local network. It provides the features of a P2P system, and the ease of use of a traditional file system for sharing files. http://p2p-fs.sf.net/	June - July 2007
LibLyric: A tool for downloading song lyrics from the internet using an approximate intersection algorithm to avoid relying on any one web-site for the song lyrics. http://dhruvbird.com/liblyric.html	March - June 2007
TDDb: The Distributed DataBase: A new database designed from ground-up for supporting Distributed Transactions, and Distributed Join Processing. http://tddb.sourceforge.net/	Jan 2005 - June 2006
Memory Allocators: A collection of memory allocators which can be used either as general purpose memory allocators, or tunable for specialized applications. http://nstdlib.sf.net/	2004 - 2005
Equation Cruncher: A tool for solving a system of linear equations using multiple compute nodes. Includes failover capabilities. It uses Cramer's rule to break down the computation over different machines.	June - Nov 2004
MemBoost: Memory booster for Windows. (more than 40k downloads) http://memboost.50g.com/	2003 - 2004

Minor Personal Projects

Developed a chat bot to search Duck Duck Go (duckduckbot@bot.im). This can be used via most chat providers since it uses the infrastructure provided by http://www.imified.com/ .	December 2010
A very flexible and completely web-based testing framework for webapps. This multifarious framework can be used for unit testing, functional testing as well as stress testing of webapps and uses Titanium desktop extensively. More powerful than Selenium, which is the most popular webapp testing framework available as of this writing.	July - August 2010
Developed a CAPTCHA image generating program (image from text) to differentiate humans from computers. Being used at http://www.bakesmith.co.in to prevent automated comment posting.	November 2008
Javascript Image Viewer: Developed an image viewing application using Javascript for rich client functionality and PHP as the backend processing engine. Capable of supporting a large number of images.	November 2008
MapReduce Simulator Framework: Developed a simple MapReduce framework for educational purposes to test my own understanding of the concepts and issues involved. This sample implementation runs multiple threads to take advantage of multiple CPUs and/or I/O latency.	October 2008
Developed a File Download Tool and an HTTP Proxy server which uses the HTTP Range header to split HTTP requests to server so that large files can be downloaded from networks that restrict downloading of large files.	November 2007

Added a progress bar to the UNIX cp(1) utility to show a file's copy progress every second. There has been much debate over this kind of a feature on Usenet groups, but I find it quite useful, esp. for large files.	October 2007
Parallel Hash-Tables: Hash Tables for use in a concurrent environment. These have minimal locking for maximum concurrency.	2004
Computerised Buzzer system: A buzzer system linked to the computer for use in determining first hits and disambiguating potential ties. This system was used in the Protocol 2004 Quiz event at DJ Sanghvi College of Engineering.	2004
x86 assembly language interpreter: Supports a small subset of the Intel x86 assembly syntax.	2003

Extra Curriculars

XMMS bug-fix and feature addition: Fixed a bug in XMMS which caused it to crash when you cleared a playlist which had songs that were queued to play. Added a feature for allowing a song to be pre-queued and to allow a songs to be queued multiple times.	2007
Top 1000 in India in the Code-For-Bill contest.	February 2006
Top 250 in Google India Code Jam, 2006.	2006
Submitted a new allocator; bitmap_allocator to libstdc++ , which is part of gcc3.4 through the current gcc release. Assisted in further developing __mt_alloc , and submitted various patches to the standard containers of the STL , which are all currently in mainline.	2004 - 2006
Stood 1st in Open Software Competition in 2004-05 in events held at: DJ Sanghvi College of Engineering, Saint Francis Institute of Technology, and Vartak College of Engineering.	2004 - 2005
Stood 2nd in C-coding competition held at DJ Sanghvi College of Engineering.	2004

Presentations

Conducted a session on Introduction to programming at Mukesh patel School of Technology Management & Engineering	November 2010
Conducted a session on Algorithms and Data Structures at DJ Sanghvi College of Engineering	August 2010
Conducted a session on Python programming at Directi Pvt. Ltd.	June 2010
Conducted a session on Map-Reduce at Mukesh Patel School of Technology Management & Engineering	2010
Conducted a session on Algorithms and Data Structures at Directi Pvt. Ltd.	2009
Conducted a session on Map-Reduce and Hadoop at Directi Pvt. Ltd.	2009
A session on An Introduction to Bash scripting at Calsoft Inc.	2007
A presentation on Distributed Databases, Distributed Transaction processing , and the the working of TDDB at DJ Sanghvi College of Engineering.	2006
A presentation on Distributed Systems - DHT, CHORD & Bit Torrent at DJ Sanghvi College of Engineering.	2006
A presentation on the design and implementation of a User-level Multi-threaded memory allocator for Capital OS at TRILabs.	2004

Papers & Publications

- | | |
|---|------|
| A distributed approach for solving a system of linear equations.
(http://sciencepub.net/) | 2004 |
| A technique for extracting song lyrics from web pages without knowing their structure (http://dhruvbird.com/liblyric.pdf) | 2006 |
| An $O(1)$ algorithm for LFU (Least Frequently Used) cache replacement.
(http://dhruvbird.com/lfu.pdf) | 2010 |
| An $O(k \log n)$ algorithm for prefix based ranked autocomplete
(http://dhruvbird.com/autocomplete.pdf) | 2011 |

Professional Associations

Student Member, Computer Society of India [CSI]

Other Interests

Professional Baker. <http://www.bakesmith.co.in/>

Cooking, Sweet making, Electronics Hobbyist, and an Amateur Carpenter

References available on request