Saurabh Kadekodi

http://www.cs.cmu.edu/~saukad *saukad@cs.cmu.edu* *847-912-2652*

Interests File and Storage Systems, Fault Tolerance, Probability and Queueing in Operating Systems, Data Compression and Deduplication.

Education

Carnegie Mellon University *2014 - present*
Ph.D. in Computer Science
Advisor: Prof. Garth Gibson
Relevant Coursework - Advanced and Distributed Operating Systems, Storage Systems.

Northwestern University *2012 - 2013*
Masters in Computer Science
Advisor: Prof. Peter Dinda
Relevant Coursework - Operating Systems, Distributed Systems, Resource Virtualization, Independent Study (Compression in Checkpointing and Fault Tolerance Systems), Art of Multi-Core Concurrent Programming, Design and Analysis of Algorithms (Honors), Probability and Stochastic Processes, Information Theory, Graph Theory.

Pune Institute of Computer Technology (University of Pune) *2005 - 2009*
Bachelors in Computer Engineering
Relevant Coursework - Computer Organization, Operating Systems, Advanced Computer Architecture, Distributed Systems, Data Structures.

Research

Carnegie Mellon University

SMRfs - A file system for Shingled Magnetic Recording *(file systems)*
This research aims at building the first shingled-aware filesystem. Specifics of my work included ways to minimize the size of unshingled partitions (typically used for frequently updated data viz. metadata, small files, etc.) on shingled disks along with the analysis and implementation of inline and opportunistic cleaning algorithms. I also overlooked the design and development of a benchmarking framework for SMRfs. (code / wiki).

University of Wisconsin - Madison - *Graduate Research Intern* *Summer 2014*
Advisor: Prof. Remzi Arpaci-Dusseau

SSD over Infiniband *(file systems)*
This study aims at comparing the performance between a locally connected SSD and remotely connected SSD (over infiniband). Using the extremely efficient SCSI RDMA protocol (SRP) for communication, we are analyzing the costs in accessing remote SSDs. As a result, we hope to get insight into improving software architecture of next-gen data centers from the storage perspective.

Price of Ext4 *(file systems)*
This study revolves around measuring the software overhead of the Ext4 filesystem. In the era of microsecond storage devices, the bottleneck is definitely shifting to the software. This study should throw light on the costs of the various submodules within Ext4 and suggest optimizations to make it future-proof.

Projects

Space Maps in Ext4 *(file systems)*
Designed and developed an extent-based free-space management technique for the Ext4 filesystem, called Space Maps, along with an allocator that uses Space Maps for disk-space allocation. Consisting of a red black tree and a log, Space Maps enhanced the allocation speed by 30% and deallocation speed by 80% and aided in reducing file and free space fragmentation (publication).

Publication - Kadekodi Saurabh, Jain Shweta. *Space Maps in Ext4*. Proceedings of the Linux Symposium, Ottawa, 13-16 July 2010. Ed. Robyn Bergeron. Ottawa, 2010. 121 - 132.

Compression in Checkpointing and Fault Tolerant Systems *(fault tolerance, compression)*
Studied the hazy nature of compression algorithms used in checkpoint / restore systems, and went on to

suggest possible enhancements and future directions in library-level checkpoint compression for faster and more efficient checkpointing with reduced disk footprint (technical survey).

VM Co-Migration *(virtualization)*
Designed and developed a UDP based VM migration module in Palacios - an OS independent embeddable VM monitor. It supported multiple-source multiple-destination migrations specifically aimed at distributed applications in HPC environments (viz. supercomputers) to exploit page-sharing among participating nodes giving increased parallelism for migration (technical report).

DNA Compression *(compression)*
Explored a run-length based preprocessing scheme exploiting the power-law behavior of genomic data to reveal possibilities of Markovian compression and variable length encoding algorithms for higher compression ratio than provided by existing dictionary based compression algorithms (technical report).

NIC of Time *(networking)*
Designed and developed a tool for exploring the state space of all possible combinations of offloaded functionalities on the NIC vs their presence in the kernel. The tool performs extensive analysis of throughput and CPU utilization to suggest one or a group of features that should be offloaded to the NIC depending on the particular workload under consideration. (code).

Active Databases *(distributed systems)*
Implemented a proof-of-concept of decentralized active databases on top of Kademlia - a distributed hash table on a decentralized peer-to-peer network. Active Databases essentially mean event-driven databases following event-condition-action (ECA) rules (technical report).

Industry

Soft Corner - *Software Developer* *Oct 2010 - Sep 2012*

PatientScribe *(healthcare web application on ipad, fault tolerance)*
Worked on data reliability, fault tolerance and audit trails at the clinical level. Developed global data reliability using data analytics. Also developed a distributed data-storage algorithm for seamless remote backup and restore.

Project-X *(social network for academia)*
Designed and developed a custom stackable context-based framework for the social network. Developed a suggester module based on a graph-based traversal of the user connection web to help generate meaningful suggestions based on connected entities. Extended WWW SQL Designer by adding collaboration (using Node.js), versioning and a chat module to build a real-time block diagram and discussion tool.

Patent - Kadekodi Saurabh, Narayanan Shrikanth, Ranade Sanyogita, Patil Bharat. 2012. *Project-X*. India Patent Application 2024/MUM/2012, filed July 12, 2012 *(pending)*.

Spring Computing Pvt. Ltd. - *Software Engineer* *Jun 2009 - Oct 2010*

Emdebian Research *(file systems)*
Created a 9 MB JFFS2 filesystem image using Emdebian Crush (an online repository that helps create BusyBox based root filesystem images) and ported the resulting filesystem image on an ARM board.

Kernel Porting on Embedded Devices *(embedded systems)*
Participated in full cycle from porting Linux kernel 2.6.30 onto 3 customized ARM boards, testing and tweaking peripherals to adapt to the new kernel and making public releases of the boards.

Leadership

Product Owner / Architect of Project-X - a social networking solution to bridge the gap between students, universities and the industry.

President of the PICT Art Circle - a group of 50 theater enthusiasts. *Jun 2007 - May 2009*

Mentorship of Masters Students - in 4 independent studies performed in shingled disks. *Fall 2014*

Skills

Programming Languages - Proficient in C, JavaScript with some experience in PHP, Java, Shell Scripting. Beginner in C++.

Systems - Proficient in Linux kernel / userspace development, file system architecture and in web application development.