Kapil Arya

• kapil@ccs.neu.edu • http://www.ccs.neu.edu/home/kapil/

Academics

Northeastern University, Boston, MA August 2014

- Ph.D. in Computer Science
- Thesis: User-Space Process Virtualization in the Context of Checkpoint-Restart and Virtual Machines
- Advisor: Prof. Gene Cooperman (leads the High Performance Computing Lab)
- Research Interests: Operating systems, high performance computing and related areas

Jai Narain Vyas University, Jodhpur, Rajasthan, INDIA July 2004

- Bachelor of Science (Triple Major: Computer Science, Physics, Mathematics)

Work Experience

Northeastern University
Visiting Computer Scientist, College of Computer and Information Sciences Jan 2016 - Present

- Working on fault-tolerance and checkpointing related research as part of the High Performance Computing Lab.
- Mentoring junior PhD students.

Lecturer, College of Computer and Information Sciences Sept 2015 - Present

- Instructor for undergraduate- and graduate-level operating systems courses.
- Responsibilities include designing syllabus, quizes, homeworks and exams as well as assigning final grades.

Mesosphere, Inc. Aug 2014 - Present
Distributed Systems Engineer, Core Group

- Worked on designing and developing the module system for Apache Mesos.
- Worked on various networking and security related projects.

VMware, Inc. Summer 2008 - 2013
PHD Intern, Virtual Machine Monitor Group

- Developed the first direct and transparent solution for the Double-Paging problem in the Virtual Machines.
- Published a research paper and filed two patents.
- Won the Best Intern Project award.

Northeastern University
Research Assistant, College of Computer and Information Sciences Sep 2008 - Jul 2014

- Lead developer for the Distributed Multi-Threaded CheckPointing (DMTCP) project.
- Designed and developed the plugin architecture for DMTCP.
- Worked on developing techniques for the deterministic record-replay component for the Fast Reversible Debugger (FReD) project.
- Mentored junior Ph.D. students.

Instructor, Undergraduate Computer Organization Spring 2012

- Took over teaching responsibility when the course instructor suddenly fell ill.
- Designed the syllabus, homework assignments, midterm and final exams.
- Provided office hours, evaluated student progress, and assigned final course grades.

Teaching Assistant, College of Computer and Information Science Spring 2008

- Assisted in conducting an undergraduate class on General Purpose computing on GPUs.
- Helped students in designing parallel programs and debug them.

Avidyne Corporation Jan - Aug 2007
Software Engineer Coop, Systems and Engines Group

- Designed and developed Graphical Pages for mainstream application using SceneGraphs and DataGraphs in C++.
- Debugged and enhanced Checklist Editor Program, a utility software written in Visual C++.
- Contributed in Code Review during Final Design Review phase of product.
- Implemented part of product's configuration phase and miscellaneous system level operations.

Publications

- Jiajun Cao, Kapil Arya, Gene Cooperman, Rohan Garg, Khaled Hamidouche, Shawn Matott, D.K. Panda, Jonathan Perkins, Hari Subramoni, Jérôme Vienne. **System-level Checkpoint-Restart for Petascale Computing**. *IEEE Intl. Conf. on Parallel and Distributed Systems (ICPADS'16).*
- Kapil Arya, Rohan Garg, Artem Polyakov, Gene Cooperman. **Design and Implementation for Checkpointing of Distributed Resources using Process-level Virtualization**. *IEEE Intl. Conf. on Cluster Computing, (Cluster'16).*
- Kapil Arya, Tyler Denniston, Ariel Rabkin, Gene Cooperman. **Transition Watchpoints: Teaching Old Debuggers New Tricks**. *(under submission).*
- Rohan Garg, Jiajun Cao, Kapil Arya, Gene Cooperman, Jérôme Vienne. **Extended Batch Sessions and Three-Phase Debugging: Using DMTCP to Enhance the Batch Environment**. *Proc. of the 2016 XSEDE Conference: Scientific Advancements Enabled by Enhanced Cyberinfrastructure,* to appear.
- Kapil Arya. **User-Space Process Virtualization in the Context of Checkpoint-Restart and Virtual Machines (Ph.D. Thesis)**. *Northeastern University, Boston, MA. August, 2014..*
- Jiajun Cao, Gregory Kerr, Kapil Arya, Gene Cooperman. **Transparent Checkpoint-Restart over InfiniBand**. *ACM Sym. on High-Performance Parallel and Distributed Computing (HPDC'14).*
- Kapil Arya, Yury Baskakov, Alex Garthwaite. **Tesseract: Reconciling Guest I/O and Hypervisor Swapping in a VM**. *ACM SIGPLAN/SIGOPS Intl. Conf. on Virtual Execution Environments (VEE'14).*
- David Abdurachmanov, Kapil Arya, Josh Bendavid, Tommaso Boccali, Gene Cooperman, Andrea Dotti, Peter Elmer, Giulio Eulisse, Francesco Giacomini, Christopher D. Jones, Matteo Manzali, Shahzad Muzaffar. **Explorations of the Viability of ARM and Xeon Phi for Physics Processing**. *Proc. of Intl. Conf. on Computing in High Energy and Nuclear Physics (CHEP'13).*

- Kapil Arya, Gene Cooperman, Andrea Dotti, Peter Elmer. **Use of Checkpoint-Restart for Complex HEP Software on Traditional Architectures and Intel MIC**. *Proc. of Intl. Workshop on Advanced Computing and Analysis Techniques in Physics Research (ACAT'13).*
- Kapil Arya, Gene Cooperman. **DMTCP: Bringing Checkpoint-Restart to Python**. *Focus issue on Scientific Computing with Python (SciPy'13), Computational Science & Discovery (CSD'15).*
- Kapil Arya, Tyler Denniston, Ana-Maria Visan, Gene Cooperman. **Semi-Automated Debugging via Binary Search through a Process Lifetime**. *Workshop on Programming Languages and Operating Systems (PLOS '13).*
- Kurt L. Keville, Rohan Garg, David J. Yates, Kapil Arya, Gene Cooperman. **Towards Fault-Tolerant Energy-Efficient High Performance Computing in the Cloud**. *IEEE Intl. Conf. on Cluster Computing, (Cluster'12).*
- Ana-Maria Visan, Kapil Arya, Gene Cooperman, Tyler Denniston. **URDB: a universal reversible debugger based on decomposing debugging histories**. *Workshop on Programming Languages and Operating Systems (PLOS '11).*
- Jason Ansel, Kapil Arya, Gene Cooperman. **DMTCP: Transparent Checkpointing for Cluster Computations and the Desktop**. *IEEE International Parallel and Distributed Processing Symposium (IPDPS'09).*

Patents

- **Techniques for Reducing Read I/O Latency in Virtual Machines**.
  Yury Baskakov, Kapil Arya, Alex Garthwaite.
  VMware, Inc. (*Two separate applications filed on 12/02/2013*).
- **Detecting and Suppressing Redundant Input-Output Operations**.
  Alex Garthwaite, Maxime Austruy, Kapil Arya.
  VMware, Inc. (*Granted on 11/17/2015*).

Invited Talks

- Kapil Arya and Niklas Nielsen. **Mesos Gets Pluggable: Introducing Mesos Modules**. *MesosCon 2015.*

Current Research / Projects

Distributed Multi-Threaded Checkpointing (DMTCP): 2007 - Present

- DMTCP is a tool for transparently checkpointing the state of a distributed program spread across many machines without modifying the user's program or the operating system kernel.
- The checkpoint image can later be used to restore program in case of node/process failure or can be migrated to another homogeneous system.
- DMTCP works completely in user space and is implemented as a set of shared libraries.
- DMTCP's plugin architecture allows third-party plugins for modularity and extensibility.
- Open source software hosted at: `http://dmtcp.sourceforge.net`

Apache Mesos: 2014 - Present

- Apache Mesos is an open-source cluster manager provides efficient resource isolation and sharing across distributed applications, or frameworks.
- The software enables resource sharing in a fine-grained manner, improving cluster utilization.
- Open source software hosted at: `http://mesos.apache.org`

Previous Research / Projects

Fast Reversible Debugger (FReD): 2009 - 2013

- FReD is a new system that uses temporal search automatically over the process lifetime to rapidly travel back in time to an earlier point of interest.
- FReD also provides reverse-expression-watchpoints and works on multi-threaded applications.
- Two important components of FReD are deterministic replay and checkpointing. Deterministic replay is a prerequisite for such a system. Checkpoints speed up the search.
- FReD currently supports GDB, Perl, Python, and Matlab debuggers.
- It can be used to add reversibility to almost any existing debugger in less than a day.
- Open source software hosted at: `http://github.com/fred-dbg`

Static Performance Evaluation for Memory Bound Computing: Jan 2008 - Apr 2008

- Developed model for static evaluation of memory-bound general purpose programs written for NVIDIA CUDA architecture.
- Used to predict/optimize the runtime of a memory-bound program by pseudocode analysis.

Hand Written Hindi Character Recognizer: Oct 2005 - Jan 2006

- Developed Artificial Neural Networks based software to recognize hand written Hindi characters drawn on-screen with the help of mouse/stylus.

Scrabble Game Playing Robot: Mar 2005 - Jul 2005

- Created the interfacing circuit, device drivers and mounted various sensors on to a robot built to play the word game Scrabble with a human being.

Semi-Autonomous Robotic Car, The Survivor: Oct 2004 - Jan 2005

- Designed the mechanical structure and electronic control circuit of a semi-autonomous robotic car and wrote the device driver and software for controlling it using a PC.

Awards and Memberships

- Won the VMware Cambridge poster session award in 2011.
- Awarded prize for "Best Mechanism, Ideas of Implementation" in robotics event "Survivor - The international Challenge" at Techfest 2005, a technical festival, organized by IIT-Bombay.
- FIRST Position in INDIA in "Programming and Problem Solving Through 'C' Language" of DOEACC 'O' Level Examination held in July, 2003.
- Apache Mesos Project Management Committee.
- ACM