ARNAB K. PAUL

Ph.D. Candidate in Computer Science at Virginia Tech

@ akpaul [at] vt [dot] edu | +1 (540) 998 - 1480 | Distributed Systems and Storage Laboratory, Virginia Tech
Blacksburg, VA, U.S.A. | https://arnabkrpaul.github.io/ | in arnabkrpaul

RESEARCH EXPERIENCE

Graduate Research Intern
Cray Inc.
Mentor: Cory Spitz, Nathan Rutman (Cray), and Scott White (LANL)
Jun. 2019 – Aug. 2019 | Los Alamos, NM, U.S.A.

- Scalable metadata indexing for large scale distributed file systems.

Graduate Student Summer Intern (Computation Scholar)
Lawrence Livermore National Laboratory
Advisor: Dr. Kathryn Mohror
May 2018 – Aug. 2018 | Livermore, CA, U.S.A.

- Analyze the metadata and job statistics of HPC I/O workloads.
- Build predictive models of I/O workloads based on the time series server data.

Graduate Student Summer Intern (Research Aide)
Argonne National Laboratory
Advisor: Dr. Ian Foster
May 2017 – Aug. 2017 | Lemont, IL, U.S.A.

- Built a scalable monitoring solution for arbitrary file systems.

Ph.D. Student @ Virginia Tech
Distributed Systems and Storage Laboratory
Advisor: Dr. Ali R. Butt
Aug. 2015 – PRESENT | Virginia Tech, Blacksburg, VA, U.S.A.

- Optimizing containers in HPC storage systems.
- Load balancing in large scale storage systems, such as Lustre.
- Auto-tuning of parallelism in Spark

Masters Student @ National Institute of Technology, Rourkela
Information Security and Data Communication Laboratory
Advisor: Dr. Bibhudatta Sahoo
May 2014 – May 2015 | NIT Rourkela, Odisha, India

- Dynamic virtual machine placement in cloud computing.

TEACHING EXPERIENCE

Instructor (Virginia Tech) *Aug. 2019 – Jan. 2020*

- CS 2505: Intro Computer Organization - *Fall 2019*

Graduate Teaching Assistant (Virginia Tech) *Aug. 2015 – Aug. 2019*

- CS 3214: Operating Systems - *Spring 2019*
- CS 5584: Network Security - *Fall 2018*
- CS 3114: Data Structures and Algorithms - *Spring 2018*
- CS 2506: Computer Org II - *Fall 2017*
- CS 2114: Software Design and Data Structures - *Fall 2016, Spring 2017*
- CS 1054: Introduction to Programming in Java - *Fall 2015, Spring 2016*

Graduate Teaching Assistant (NIT Rourkela) *Aug. 2014 – May 2015*

- CS 171: Computing Lab - *Autumn 2014, Spring 2015*
- CS 670: Data Mining Lab - *Spring 2015*

LIFE PHILOSOPHY

"Your best teacher is your last mistake." - Dr. A. P. J. Abdul Kalam

RESEARCH STATEMENT

I am a fifth year Ph.D. candidate in the Department of Computer Science at Virginia Tech. I work in the Distributed Systems and Storage Laboratory headed by Dr. Ali R. Butt. My research interests include *cloud computing, distributed systems, distributed file systems and big data APIs.*

EDUCATION

Ph.D. in Computer Science
Virginia Tech
Aug. 2015 – PRESENT (Expected: Aug. 2020)

- Advisor: Dr. Ali R. Butt
- GPA: 4.0/4.0

M.S. in Computer Science & Applications
Virginia Tech
Aug. 2015 – May 2018

- GPA: 3.85/4.0

M.Tech. in Computer Science & Engineering
National Institute of Technology, Rourkela
Aug. 2013 – May 2015

- Specialization: Software Engineering
- Advisor: Dr. Bibhudatta Sahoo
- GPA: 9.56/10.0

B.Tech. in Computer Science & Engineering
West Bengal University of Technology
Aug. 2009 – May 2013

- GPA: 9.02/10.0

SKILLS

Lustre File System | Apache Spark | Containers | Edge Computing | IoT

C | C++ | JAVA | Python | C# | SCALA | MySQL | HTML | CSS

UNIX | git | svn | latex | gnuplot

MENTORING EXPERIENCE

- Redwan Ibne Seraj Khan - *PhD, Virginia Tech, 2019 -*
- Debasmita Biswas - *MS, Virginia Tech, 2020 -*
- Subil Abraham - *MS, Virginia Tech, 2019 - 20*
- Arpit Goyal - *MS, Virginia Tech, 2016 - 17*

SELECTED PUBLICATIONS

Book Chapters

- **Arnab K. Paul** (2020). *Edge or Cloud: What to Choose?* Cloud Network Management: An IoT based Framework, Taylor Francis Group, CRC Press.
- **Arnab Kumar Paul** and Bibhudatta Sahoo (2017). *Dynamic virtual machine placement in cloud computing*. IGI Global, pp. 136–167.

Conference & Workshop Proceedings

- Abraham, Subil et al. (2020). "On the Use of Containers in High Performance Computing Environments". In: *Proceedings of the IEEE International Conference on Cloud Computing (Cloud)*. IEEE.
- **Arnab K. Paul**, Brian Wang, et al. (2020). "Efficient Metadata Indexing for HPC Storage Systems". In: *20th IEEE/ACM International Symposium on Cluster, Cloud and Internet Computing (CCGrid)*. IEEE/ACM.
- **Arnab K. Paul**, Ryan Chard, et al. (2019). "FSMonitor: Scalable File System Monitoring for Arbitrary Storage Systems". In: *2019 IEEE Cluster*. IEEE.
- **Arnab K. Paul**, Olaf Faaland, Adam Moody, Elsa Gonsiorowski, Kathryn Mohror, and Ali Butt (2019). "Understanding HPC Application I/O Behavior Using System Level Statistics". In: *Poster - SC 2019*.
- **Arnab K. Paul**, Olaf Faaland, Adam Moody, Elsa Gonsiorowski, Kathryn Mohror, and Ali R Butt (2019). "Improving I/O Performance of HPC Application Using Intra-Job Scheduling". In: *Work-In-Progress in Proceedings of the 4th Joint International Workshop on Parallel Data Storage & Data Intensive Scalable Computing Systems (PDSW-DISC'19) in conjunction with SC'19*.
- Sim, Hyogi et al. (2019). "Cslim: Automated Extraction of IoT Functionalities from Legacy C Codebases". In: *2019 8th International Workshop on Computing and Networking for IoT and Beyond (ComNet-IoT), Proceedings of the 20th International Conference on Distributed Computing and Networking*. ACM, pp. 421–426.
- Wadhwa, Bharti et al. (2019). "Resource Contention Aware Load Balancing for Large-Scale Parallel File Systems". In: *2019 33rd IEEE International Parallel and Distributed Processing Symposium (IPDPS)*. IEEE.
- **Arnab K. Paul**, Arpit Goyal, et al. (2017). "I/O load balancing for big data hpc applications". In: *2017 IEEE International Conference on Big Data (Big Data)*. IEEE, pp. 233–242.
- **Arnab K. Paul**, Steven Tuecke, et al. (2017). "Toward scalable monitoring on large-scale storage for software defined cyberinfrastructure". In: *Proceedings of the 2nd Joint International Workshop on Parallel Data Storage & Data Intensive Scalable Computing Systems*. ACM, pp. 49–54.
- **Arnab Kumar Paul**, Wenjie Zhuang, et al. (2016). "Chopper: Optimizing data partitioning for in-memory data analytics frameworks". In: *2016 IEEE International Conference on Cluster Computing (CLUSTER)*. IEEE, pp. 110–119.
- **Arnab Kumar Paul**, Sourav Kanti Addya, et al. (2014). "Application of greedy algorithms to virtual machine distribution across data centers". In: 11^{th} *IEEE India Conference INDICON 2014, Emerging Trends and Innovation of Technology*. IEEE, pp. 1–6.

ACHIEVEMENTS

2020

Bitshares Fellowship by CS @ VT
Member of Dean's Graduate Team @ VT

2019

Member of Association for India's Development
Member of CS Graduate Council @ VT
Travel Grant IEEE Cluster '19
Member of Dean's Graduate Team @ VT
Student Volunteer, SC '19

2018

Bitshares Fellowship by CS @ VT
Member of Dean's Graduate Team @ VT
Student Volunteer, SCiNet @ SC '18

2017

President of Bengali Students' Ass. @ VT
Travel Grant IEEE BigData '17

2016

Travel Grant IEEE Cluster '16
Student Volunteer, SC '16

2015

Gold Medalist, CS @ NITRKL

PROFESSIONAL SERVICES

PC Member

ICDCS 2020

Reviewer

IEEE Transactions on Parallel and Distributed Systems - 2019, 2020
Cluster Computing Journal - 2019, 2020
ASTESJ Journal - 2018
IJGHPC Journal - 2018, 2019, 2020
AUTOSOFT Journal - 2018
MGS Journal - 2017

External Reviewer

IEEE TSC Journal '18
BigData '17/'18, Cluster '17/'18, ECOOP '20, HPDC '17/'18/'20, IC2E '17, ICCD '19, ICDCS '17/'18/'19, ICS '17/'18, IPDPS '18/'19/'20

REFEREES

Dr. Ali R. Butt
butta@cs.vt.edu
Virginia Tech
Blacksburg, VA, U.S.A.