

# Sultan Mahmud Sajal

✉ smsajal116@gmail.com

🌐 smsajal.github.io

☎ +1-814-380-3595

## Education

- 2018 - Ongoing    📖 **Ph.D. in Computer Science and Engineering**  
**The Pennsylvania State University**  
**Thesis topic:** Improving the Fidelity of Trace-Driven Experiments in Cloud Computing Systems  
**Advisors:** *Timothy Zhu* and *Bhuvan Urgaonkar*
- 2013 - 2017    📖 **B.Sc. in Computer Science and Engineering**  
**Bangladesh University of Engineering and Technology**  
**Thesis topic:** An Empirical Study on the Growth of New Languages and Their Users in Stack Overflow  
**Advisor:** *Rifat Shahriyar*

## Professional Experience

- May, 2022 - August, 2022    📖 **Research Intern** at Cloud Operations Research (CORE), Microsoft Research.  
- *Mentors:* Luke Marshall, Beibin Li, and Ishai Menache.
- May, 2021 - August, 2021    📖 **Research Intern** at Gray Systems Lab (GSL), Microsoft.  
- *Mentors:* Abhishek Roy and Joyce Cahoon.
- August, 2018 - Present    📖 **Graduate Research Assistant** and **Graduate Teaching Assistant** at The Pennsylvania State University.  
- *Advisors:* Timothy Zhu, Bhuvan Urgaonkar.  
- *Collaborator:* Siddhartha Sen (*Microsoft Research*).
- October, 2017 - July, 2018    📖 **Junior Software Engineer** at Reve Systems.  
- *Manager:* Golam Md Muktadir.

## Skills

- Programming Languages    📖 Java, Python, C++, Scala, C, R
- Databases    📖 MySQL, PostgreSQL
- Technologies    📖 AWS Services and SDK, Azure Services and CLI, Apache Spark, Kubernetes(K8s), Docker, Git, Nginx, Varnish Http Cache, Memcached
- Scripting    📖 Bash, HTML, CSS,  $\LaTeX$
- Specialized Skills    📖 Distributed Systems, Microservice, Analysis and Testing of Systems, Performance Evaluation

## Recent Projects

- 📖 **Facilitate Isolated Experimentation for Reproducible Results in Cloud Systems** [2023 - Ongoing]  
- Motivate the need for isolated experimentation for realistic and reproducible experiments through realistic experiments  
- Using synthetic and real-world applications with *Kubernetes* as orchestration service
- 📖 **Efficient and Scalable Cloud Admission Control in Azure** [May, 2022 - Dec, 2022]  
- Developed novel admissions control techniques to guarantee SLAs for both allocated and reserved resources while maximizing resource efficiency.  
- Extended existing simulator ( written in *C++*) and generated synthetic capacity reservation requests using *Python3* to complement real-world trace, stored in *PostgreSQL*

## Recent Projects (continued)

- **Upscale Workloads to Evaluate Cloud Systems at High Load** [2020 - 2023]
  - Developed novel upscaling techniques for real workloads to enable faithful systems experimentation under varying loads.
  - Deployed a stateful replicated *DeathStarBench Social Network benchmark*, using *Varnish Http Cache* as front-end reverse proxy cache and *Nginx* as load balancer, and deployed in *Azure*
  - Deployed a stateless 16-node distributed Mediawiki application using *MySQL*, *Memcached*, and *Nginx* load balancer, and deployed in *Azure*
- **Development of Flight Simulator for Spark Jobs** [May, 2021 - Aug, 2021]
  - Analyzed anonymized *Azure HDInsight* telemetry data to generate synthetic database and queries to facilitate realistic performance experiments using synthetic data
  - Developed in *Apache Spark* using *Scala*
- **Improvements of the Python based PickleDB** [2021]
  - Investigated effects of different design choices: (1) presence of cache, (2) cache size, (3) write-through vs write-back, and (4) choice of Join algorithms
  - Performed evaluation using *IMDB* dataset
- **Downscale Workloads from Cloud System for Realistic Experimentation and Prototyping** [2018 - 2020]
  - Developed novel techniques to downscale cloud workloads while preserving important characteristics such as arrival process and performance to facilitate realistic systems research and industry prototyping.
  - Deployed two different application systems: (1) distributed *Elgg* and (2) *MediaWiki* with autoscaling
  - Both uses *MySQL* database and *Nginx* load balancer and are deployed as *Docker* containers in *AWS*
- **Evaluation of Different Clustering Methods** [2019]
  - Studied the performance of *K-means clustering*, *Gaussian Mixture Model*, *PCA*, *tree-structured classifier*, *a boosting method* and *SVM* on real-world dataset using *R*
- **Development and Evaluation of Replicated Linearizable Key-Value Storage** [2018]
  - Evaluated performance of two variants of key-value storage providing linearizability: (1) *blocking protocol* and (2) *A non-blocking ABD algorithm-based protocol*, implemented in *C++*
- **Implementation and Evaluation of LSM Tree** [2018]
  - Evaluated performance of a *2-level Log-Structured Merge Tree-based key-value storage*, implemented in *C++*
- **Discrete Event Simulator** [2018]
  - Implemented a discrete event simulator consisting of task dispatching, queueing networks, and task scheduler using *C++*
- **Implementation and Evaluation of Key-Value Storage** [2018]
  - Implemented and compared different types of IO for key-value storage implemented with *C++*
  - The evaluated approaches are: (1) multi-threaded blocking IO, (2) single event-driven thread non-blocking IO with – (2a) polling-based event notification, and (2b) signaling-based event notification

## Publications

- 1 **Sajal, Sultan Mahmud**, T. Zhu, B. Urgaonkar, and S. Sen, “*TraceUpscaler: Upscaling Traces to Evaluate Systems at High Load*,” in 19th European Conference on Computer Systems, ( **Eurosys '24** ), 2024, (Accepted to Appear).
- 2 **Sajal, Sultan Mahmud**, L. Marshall, B. Li, *et al.*, “*Kerveros: Efficient and Scalable Cloud Admission Control*,” in 17th USENIX Symposium on Operating Systems Design and Implementation, ( **OSDI '23** ), 2023.
- 3 **Sajal, Sultan Mahmud\*** and Hasan\*, Rubaba, T. Zhu, B. Urgaonkar, and S. Sen, “*TraceSplitter: A New Paradigm for Downscaling Traces*,” in Proceedings of the Sixteenth European Conference on Computer Systems, ( **EuroSys '21** ), *\*Equal Contribution*, 2021.