# Konstantinos Konstantinidis

Ph.D., Iowa State University

515-357-2016 | k.m.konstantinidis@gmail.com | kkonstantinidis.github.io | github.com/kkonstantinidis

## Skills

**Proficient**: Python • Java • AWS • MATLAB • PyTorch • MPI • MapReduce
**Good**: C++ • SQL • Hadoop • NumPy • Bash • Git • Jenkins

## Education

**Iowa State University** — Ames, IA
*Ph.D., Electrical and Computer Engineering, GPA: 3.94/4, Advisor: Prof. Aditya Ramamoorthy.* — *Dec. 2022*

**Technical University of Crete** — Chania, Greece
*Diploma (5-year program), Electrical and Computer Engineering, GPA: 3.4/4, Advisor: Prof. George Karystinos.* — *Dec. 2016*

## Industry Experience

**Software Engineer at C3 AI** | ***Java, Python, IntelliJ IDEA, OpenSearch.*** — Redwood City, CA | 09/2022 – Present
I am a member of the Platform - Data team developing C3.ai's *feature store* system that stores machine learning data. My work involves alleviating bottlenecks identified through algorithmic analysis and profiling (e.g., *YourKit*) across the data pipeline, hence improving the latency of feature serving and enabling the system to scale distributively across multiple nodes.

**Software Engineer Intern at Meta (Facebook)** | ***Python, PHP, SQL, C++.*** — Menlo Park, CA | 05/2022 – 08/2022
Developed debugging tools for machine learning feature authoring used in Facebook Marketplace. The implemented framework categorizes errors during feature compilation, generates alerts, and assigns tasks to the appropriate team; I integrated it with the CI/CD. Another end product of my work was a UI tool to retrieve and transform feature values from low-latency storage.

**Software Engineer Intern at C3 AI** | ***Java, Python, AWS, Jenkins, Splunk.*** — Redwood City, CA | 06/2021 – 08/2021
Implemented a framework for cluster failure prediction; its first component is the data pipeline which loads cluster health metrics, handles missing data, and creates a training data set. The second component is the ML pipeline which trains a model and predicts the cluster's state using streaming data. Followed the process of continuous integration / continuous deployment (CI/CD).

## Research Projects

**ByzShield: Robust distributed learning** | ***Python, PyTorch, AWS, MPI, Bash.*** — 01/2020 – 04/2021
Developed a defense for distributed learning in which computing devices may return erroneous or malicious gradients. The method achieves a **20**% increase in top-1 accuracy on the **CIFAR-10** dataset and a **36**% reduction in the fraction of corrupted gradients.

**CAMR: Aggregated MapReduce over multiple jobs** | ***Python, AWS, MPI, Bash.*** — 12/2018 – 04/2020
Proposed a method named *CAMR*, which reduces the MapReduce communication overhead when the desired functions can be aggregated (amenable to deep learning applications). It achieves state-of-the-art communication load but with an exponentially smaller requirement on the minimum number of jobs. Its speedup is **4.3**× over the baseline approach.

**Staggler mitigation in matrix multiplication** | ***Python, MPI, AWS.*** — 07/2018 – 01/2019
Proposed and implemented a technique to tolerate the presence of servers that suffer from slow computation. Our algorithm can alleviate a higher number of slow servers and requires **80**% of the time needed by prior methods on actual AWS EC2 simulations.

**MapReduce communication load reduction** | ***C++, AWS, MPI, Hadoop, HDFS, Bash.*** — 05/2017 – 12/2018
Proposed an algorithm to reduce **MapReduce** communication load; it uses *Single Parity Check* codes and design theory to assign tasks to servers and splits files less finely than prior work. Tweaked the **TeraSort** algorithm (to sort data sets in **HDFS**). The method uses **MPI** for communication. It supersedes the state-of-the-art by **2.6**× and the baseline approach by **4.7**× on AWS EC2 clusters.

## Selected Awards

- **Best Student Poster Award**, Midwest Machine Learning Symposium (MMLS), June 2019 (link).
- **Teaching Excellence Award**, Iowa State University, May 2019.

## Selected Publications (Google Scholar)

- **K. Konstantinidis** and A. Ramamoorthy, "Aspis: Robust Detection for Distributed Learning," *IEEE International Symposium on Information Theory (ISIT)*, June 2022.
- **K. Konstantinidis** and A. Ramamoorthy, "ByzShield: An Efficient and Robust System for Distributed Training," *Machine Learning and Systems (MLSys)*, April 2021.
- **K. Konstantinidis** and A. Ramamoorthy, "Resolvable Designs for Speeding up Distributed Computing," *IEEE/ACM Transactions on Networking (ToN)*, August 2020.
- L. Tang, **K. Konstantinidis** and A. Ramamoorthy, "Erasure coding for distributed matrix multiplication for matrices with bounded entries," *IEEE Communication Letters (COMML)*, November 2018.
- **K. Konstantinidis** and A. Ramamoorthy, "Leveraging Coding Techniques for Speeding up Distributed Computing," *IEEE Global Communications Conference (GLOBECOM)*, December 2018.