Caleb Carlson

inf0rmatiker.github.io

ccarlson355@gmail.com

(512) 934-3355

Summary

Ambitious and motivated Systems Engineer with solid leadership, teamwork, and creativity. Eager to learn cutting-edge technologies and passionate about HPC, distributed systems, big data, high-speed networking, machine learning and AI.

Industry Experience

- **Hewlett Packard Enterprise** — Fort Collins, CO
  HPC Storage - Systems Software Engineer II — *June 2022 – Present*
  - Created an automated, user-friendly, and scalable software upgrade procedure for DAOS storage cluster
  - Designed and implemented cluster bootstrapping process for DAOS storage product, using HPCM as a foundation
  - Triaged Kubernetes management issues in several customers' legacy storage products
  - Contributed bugfixes and LNet kernel module tuning improvements to community Lustre filesystem project
  - Benchmarked HPC Lustre filesystems against GPUDirect Storage and POSIX clients
  - Created Jenkins pipeline that builds DKMS-enabled Lustre clients for a variety of distros and kernels
  - Streamlined add-on data mover node discovery and software mapping process
- **Cray, Inc.** — Longmont, CO
  HPC Storage - Software Development Intern — *May 2020 – May 2022*
  - Built Jenkins pipelines, Go unit tests, and Slack notifications for team's repositories
  - Improved gRPC communication strategy between data path microservices
  - Automated Lustre filesystem provisioning based on discovered NVMe/SCSI drives
- **Data Ductus, Inc.** — Longmont, CO
  Software Development Intern — *May 2019 – Aug 2019*
  - Implemented administrative RBAC for Verizon APIs
  - Reduced end-to-end testing time from over 20 minutes to 6 minutes

Education and Research

- **Colorado State University** — Fort Collins, CO
  M.S. in Computer Science; GPA: 4.000 — *Aug 2020 – May 2022*
- **Colorado State University**
  B.S. in Computer Science, Math Minor; GPA: 3.984 — *Aug 2016 – May 2020*
- **Graduate Research Assistant**
  Pallickara Lab — *Jan 2020 - May 2022*
  - Published two papers to ACM and IEEE conferences, see *Projects*
  - Established procedures for ingesting and sharding datasets into MongoDB, and managing clusters of 200 nodes
  - Helped develop LSTM recurrent neural networks for point-cloud datasets generated from agricultural LIDAR equipment
- **Teaching Assistantships**
  Data Structures and Algorithms, Software Engineering, Operating Systems — *Aug 2017 - Dec 2020*
  - Facilitated learning as lead TA in multiple classes, teaching object-oriented design patterns, data structures, algorithms, operating systems, Agile software development, and version control

Projects

- **Omniscient** Distributed resource monitoring for memory pressure, CPU usage, storage I/O, Ethernet and InfiniBand throughput. [github.com/inf0rmatiker/omniscient]
- **Validation Service** Validates spatiotemporal models on large datasets in a distributed fashion and visualizes metrics as a geospatial heatmap. [github.com/Project-Sustain/validation-service]
- **Research Papers**
  - ACM BDCAT 2021 — *Distributed Orchestration of Regression Models Over Administrative Boundaries*
  - IEEE ICCBE 2022 — *Resource Efficient Profiling of Spatial Variability in Performance of Regression Models*