Morris Tai

GitHub / morristai.github.io / / morristai01@gmail.com

EXPERIENCE

Trend Micro

Sr. Software Developer

Ontario, Canada
08/2022 - Present

- Develop and maintain a data lake solution as the XDR platform backend, mainly including:
  - Implemented an OLTP + OLAP database that supports multitenancy, high throughput, and auto-scaling with a lead-follower architecture. Utilized Cosmos DB and Azure Data Explorer to deliver custom queries and sweeping APIs for downstream products, primarily for cyber threat hunting and investigation. Employed Databricks and serverless functions to construct cross-cloud ETL pipelines for data processing. Integrated additional log types and data sources into the pipeline, e.g., Telemetry/Detection of Endpoint, Secure Web Gateway, and Container-based platform.
  - Managed and iterated the pipeline schema with different teams. Wrote BDD tests to ensure the quality of the pipeline. Utilized Terraform and CloudFormation for cloud infrastructure management and automation, with an in-depth understanding of the public cloud & cloud native architecture.
  - Conducted research and integrated a Kubernetes security solution for the XDR platform using Falco and eBPF/Kernel Module. Integrated as a privileged container in a DaemonSet.
  - Enhanced AWS EMR with Apache Flink for model matching and filter rules to detect cyber threats in real-time.

Software Developer

Taipei, Taiwan
03/2021 – 07/2022

- Developed and managed a robust streaming platform using the Kafka ecosystem as a cloud app security backend. Implemented various Spring Boot microservices, including asynchronous services with Java Reactive Streams and services utilizing JNI for parsing metadata. Collaborated on expanding support for downstream products like SASE and Attack Surface Management. Managed and enhanced Avro and Protobuf schemas, integrating them with Kafka Schema Registry. Collaborated with the data science team to develop a detection pipeline using Apache Spark and Airflow.
- Migrated from a VM-based architecture to Azure Kubernetes, overseeing the implementation of internal deployment tools for seamless service deployment across diverse projects, regions, and environments. Managed the transition from Ansible to Helm, focusing on parameterization and version control. Additionally, introduced advanced tracing and logging capabilities through Loki and Tempo, replacing Fluentd in a sidecar. Demonstrated understanding of the container ecosystem's low-level intricacies through thorough research on namespaces, cgroups, OCI, Shim, MicroVM, etc.

Intel

Intern. Software Engineer (full-time)

Taipei, Taiwan
08/2019 – 09/2020

- Assisted the team in debugging graphics drivers using WinDbg and memory dumps. Wrote a parser tool capable of extracting Panel EDID and other information produced by the driver. Studied the graphics source code in User/Kernel Mode to understand how SSE/AVX SIMD works.

- Developed a dashboard using Flask and Plotly that displayed data related to our driver issues. Enabled users to set custom criteria for aggregating and filtering data from multiple sources. Visualized data with different types of charts, including bar, pie, and line charts. Implemented a notification system that could send emails to users when the data met the specified criteria.

EDUCATION

National Changhua University of Education 09/2016 to 06/2020
Bachelor of Management Information System

PERSONAL PROJECT

Apache OpenDAL Committer April 2023 to Present

- Contribute to the Rust core library and assist in implementing various services and layers, such as limiting IO throughput and server-side encryption. Developed Rust code that natively supported Databricks FileSystem, OpenStack Swift, and HuggingFace FileSystem. Implemented concurrent stat to speed up listing. Managed the release process, and more.
- My PRs: https://github.com/apache/opendal/pulls?q=is%3Apr+author%3Amorristai

Cathay / Microsoft Big Data Competition (Championship) media Sep 2016 to Jan 2017

- As the main analyst and leader of the team, I analyzed insurance data and built predictive models based on supervised learning and stacking model. Through detailed research, we discovered the relationship between insurance and cancer.

SKILLS

- Programming Languages: Multilingual (not limited to any specific language), especially experienced in Rust, Python, and comfortable with Go, Java, C, and C++ (in random order).
- DevOps: Experienced in Terraform, Ansible, Airflow, Kubernetes, Helm, Grafana, and Prometheus.
- Data: Kafka, Spark, Database, Lakehouse, Avro, Parquet, Protobuf, Structured-data machine learning.
- System Programming: Experienced in Unix-like OSes, Container/Kubernetes ecosystem, and WASM container. Understand the general architecture of cloud systems.
- Rust: Familiar with Cargo, understand common traits and procedural macros, Rust asynchronous runtimes, and unsafe usage. Contributed to Rust repositories and participate in the community in my free time.