Morris Tai

(Tai, Ping Chuan)

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

EXPERIENCE

Trend Micro

Sr. Software Developer Ontario, Canada. 08/2022 - Present

Software Developer Taipei, Taiwan. 03/2021 - 07/2022

Development and Operation of XDR Platform Backend and Data Lake.

ETL Pipeline:

- Works on a pipeline that processes real-time logs from upstream products within the Databricks ecosystem, optimizing file sizes and counts to reduce database load. Integrates Azure Data Explorer for time-series OLAP DB, optimizing resource usage by segmenting workload groups and ensuring high availability with leader/follower database patterns.
- Constructs and manages the ETL pipeline infrastructure(Terraform, Helm, etc.), enabling multi-tenancy, abstracting region/environment configurations, and incorporating CI/CD processes. Enhances observability by establishing metrics that monitor throughput for each component and measure latency at different stages within the pipeline. Collaborates with threat research teams to develop data schemas.
- Developed and managed a streaming platform using the Kafka ecosystem as an identity detection security backend. Implemented various Spring Boot microservices, including asynchronous streaming with the Java reactive runtime. Utilized C++ through JNI for parsing metadata. Expanded support for downstream products. Managed and iterated Avro and Protobuf schemas, integrating them with the Kafka Schema Registry. Worked with the threat research team to develop a detection pipeline using Airflow and Apache Spark.
- Migrated the architecture from a VM-based system to Kubernetes. Managed the transition from Ansible to Helm, focusing on parameterization and performance tuning. Replaced Fluentd with a sidecar, introducing tracing and logging capabilities through Loki and Tempo.

API Development:

- Develops APIs for downstream teams, incorporating serverless functions, Cosmos DB, and asynchronous APIs for efficient querying and handling of large raw logs in the Data Lake.
- Implements dynamic sliding windows, designs custom queries for APT detection, and improves search efficiency with cross-team search/index and pre-filter mechanisms.

Security and Research:

- Researched and integrated container security solutions with Falco and eBPF/Kernel Module. Gained a good understanding of the low-level aspects of the container ecosystem, such as namespaces, cgroups, OCI, Shim, MicroVM, etc.
- Advocated for the Rust language since 2021, guiding the team through its features and examples. Rewrote some of the serverless functions in the pipeline using Rust.
- Addressed container vulnerabilities to comply with FedRAMP/PCI standards and conducted architectural reviews for SOC/ISO audits.

Intel

Taipei, Taiwan

Intern. Software Engineer (full-time) 08/2019 – 09/2020

- Studied the graphics driver source code in User/Kernel Mode to understand how SSE/AVX SIMD works. Assisted the team in debugging graphics drivers using WinDbg and mem dumps. Wrote a parser tool capable of extracting Panel EDID and meta data produced by the driver.
- Developed a dashboard with mini ETL using Flask and Plotly for driver issues, enabling data aggregation and filtering for the team.

EDUCATION

National Changhua University of Education Sep 2016 to Jun 2020

- Bachelor of Management Information System (GPA: 3.8 /4.0)

OPEN SOURCE

OpenDAL Committer | Apache Software Foundation April 2023 to Present

- OpenDAL aims to provide efficient data access from various storage services and is known for its distributed, low latency, and high performance. I contributed to the Rust core library and implemented various services and layers, such as limiting IO throughput and server-side encryption. I developed Rust code that natively supports Databricks FS, OpenStack Swift, and Hugging Face FS. I designed concurrent stat to speed up listing, among other features.
- PRs: https://github.com/apache/opendal/pulls?q=is%3Apr+author%3Amorristai

CERTIFICATE

CKAD: Certified Kubernetes Application Developer | The Linux Foundation

- ID Number: LF-he54r5sc2f Mar 2024 to Mar 2027

COMPETITION

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

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

AIME: American Invitational Mathematics Examination certificate Mar 2014

- Invited for being in the top 2.5% of participants worldwide in the AMC 12

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).
- **Rust**: 2+ years with Rust. Familiar with Cargo, understand common traits and procedural macros, Rust asynchronous mechanism(refer to my PR), and unsafe usage. Currently studying for advance profiling. Contributed to Rust repositories and participate in the community in my free time.
- **System Programming**: Experienced in Unix-like OSes, Container/Kubernetes ecosystem, and WASM container. Understand the general architecture of cloud systems.
- **DevOps**: Terraform, Ansible, Kubernetes, Helm, Grafana, Prometheus, Loki, Tempo, Docker, GitHub Action, etc.
- **Data**: Kafka, Spark, PostgreSQL, S3, Blob, ClickHouse, Avro, Parquet, Protobuf, Structured-data machine learning.

MISCELLANEOUS

Personal Blog

- Blog: https://morristai.github.io/
- My Note on Rust and its related aspects: Link

Algorithm / Data Structure

- I wrote Rust solutions without recursion to help others get a better grip on ownership: https://leetcode.com/Morris_Tai/