Thomas Benoit

703-307-5534 | thomasgeorgebenoit@gmail.com | linkedin.com/in/tgbenoit

Experience

Software Engineer May 2022 – Present

Software Engineering Institute, Carnegie Mellon University

- Engineered and operationalized software foundational to the support of enterprise security monitoring tools, leading to promotion in January 2024 and numerous Spot Awards.
- Designed and implemented a novel VPC Flow Log optimization strategy across AWS Virtual Private Cloud environments by identifying compression opportunities through cloud infrastructure analysis. Implemented multiple automated ETL (Extract, Transform, Load) pipelines to perform the optimization, resulting in a 60 percent reduction in log size while preserving full analytic and security monitoring value. The approach was subsequently adopted for Microsoft Azure environments and presented at a research conference (Lambda function code). (*AWS, Lambda, Python, NumPy, Pandas, Splunk, CloudWatch*)
- Independently developed a standalone Python program to convert cloud flow logs into SiLK binary flows, taking ownership of the entire technical approach from algorithm design through deployment. Established integration standards for existing security tools, managed persistent configuration files for consistent sensor ID labeling, and collaborated with network analysts to align with operational security needs. Reduced storage requirements by 75 percent. (*Python, NetFlow, SiLK*)
- Designed, implemented, and deployed the first cloud-based instance of YAF for traffic mirroring streams in AWS. Engineering functionality to perform VxLAN-UDP decapsulation and extract key values, allowing the sensor to generate flow from AWS mirrored packet streams and integrate with enterprise cloud security infrastructure. Recognized for extending secure telemetry collection into the cloud. (*AWS, C, Lua, IPFIX, Wireshark*)
- Briefed the results of various work streams across technical and executive audiences. Leveraged specialized knowledge and deep understanding of network and cloud infrastructure to enrich cross-team collaboration. Provided bottom-line-up-front messaging of such work to ensure understanding across business units, conveyed the value of security decisions, and optimized work efforts across directorate teams.
- Architected enhancements to a flow sensor's configuration system to support analyst-customizable regex patterns for deep packet inspection (DPI), enabling users to update packet detection rules without modifying source code. Independently designed data types and structures to ensure flexible and maintainable configurations, making it easier for analysts to extend packet inspection capabilities to meet evolving security requirements. (*C, Lua*)

Projects

Chess Opening Success Analysis System

Architected a high-performance distributed system that processed and analyzed 125 million chess games at scale. Designed and implemented a scalable ETL pipeline using Apache Kafka middleware to decouple PGN parsing from PostgreSQL database operations, enabling horizontal scaling with multiple producers and consumers at 2,500 games/second throughput. Optimized query performance with custom PostgreSQL materialized views for precomputed statistics and developed a React frontend to visualize opening success rates across player levels. (*Python, Kafka, PostgreSQL, Docker, React*)

Education

George Mason University

Bachelor of Science in Computer Science

Publications

Demystifying the Shape of Traffic in the Cloud Speaker, 2024

Presented at FloCon 2024 on how AWS flow log sensor architecture introduces routing and security monitoring challenges, highlighting key differences between cloud and on-premise network telemetry.

Keeping Up with Cloud Security using Available Telemetry Author, 2025

Coauthored an article on cloud flow collection, focusing on flow generation nuances, sensor deployment, and architecture considerations for optimal data collection.