# JASON BAK

jasonbak.co | jason.j.bak@gmail.com | Seattle, WA

## EXPERIENCE

**HubSpot** | Senior Software Engineer                                                           2020 - 2025
Infra Costs                                                                                            Seattle, WA

*Worked within Platform Infrastructure, building from scratch AWS cost optimization and efficiency systems. Achieved significant cost savings across a company with **$210M** in annual cloud spend.*

- Identified waste patterns, advocated for prioritization, and implemented an automated resource optimization system across high-cost resource types, **eliminating $2M in accrued waste** with **zero operational disruptions** spanning all organizations.
- Built a multi-layered automated enterprise cost allocation tagging system (event-driven workers, scheduled remediation jobs, infrastructure-level injection), achieving **>99% cost coverage**. This is the foundation of our cost allocation datasets, alerting infrastructure, and optimization tools.
- Architected configurable cost anomaly alerting system with custom grouping dimensions, multiple threshold types (percentage, scalar, statistical), alert muting, and team-based routing, **preventing $1M in runaway costs** and mentoring teammates on implementation.
- Drove **$56M in annual cost attribution improvements** via Spark jobs, event processors, and cost allocation logic changes in an ETL data pipeline.  Included EC2 storage attribution calculations, Kubernetes CPU/GPU cost breakdowns, and Elasticsearch indexer team mappings.
- Built scalable cost tracking infrastructure with Kafka listeners for usage sampling, event processors, and Spark jobs with custom filtering to enable granular cost allocation **across 400 engineering teams**.
- Built a centralized AWS resource metadata service, **tracking 2M resource states**, enabling historical state queries for cost data processing and simplifying data retrieval across internal sources and AWS.
- Collaborated directly with infrastructure and product teams to resolve their cost spend issues (waste, inefficiency, runaway costs, "unexplained" costs), **reducing costs by $2M**.

**Intuit** | Software Engineer Intern                                                             2019
Data Platforms                                                                                       Mountain View, CA

- Automated EMR cluster rotation workflow for an internal data infrastructure console, **reducing an 11-minute manual process to one click**, through orchestrated termination, polling, and recreation.

**Amazon** | Software Development Engineer Intern                                    2018
AI Algorithms                                                                                         Seattle, WA

- Enhanced anomaly detection by capturing temporal context in time-series data for the Random Cut Forest C++ algorithm library. Included Python and Java bindings for cross-language interoperability.

## EDUCATION

**Carnegie Mellon University** | B.S. in Statistics and Machine Learning                 2019
Minor in Computer Science                                                                          Pittsburgh, PA
TA: Intro to Machine Learning, Data Structures (2x)
RA: Institute for Software Research (Machine Learning)

## SKILLS

**Languages**: Java, Python, C++, C, R, JavaScript
**Technologies:** Spark, MySQL, Kafka, Elasticsearch, AWS (SQS, EventBridge, Athena, S3, DynamoDB), Maven, Grafana, Prometheus, Flask, Snowflake, PostgreSQL, React, Spring Boot, Git