Chaman Thapa

Cambridge, MA, USA
✉ chaman.thapa@gmail.com
🖮 www.chamanthapa.com

Career Summary

- 12+ years of experience in designing and developing software systems with complex backend architecture following agile software development practices.
- Hands-on experience finding and resolving security flaws in web applications (SQL Injection, XSS, CSRF, XXE, DOS).
- Hands-on experience resolving performance bottlenecks and scaling complex storage and distributed database systems.
- Excellent communication and teamwork skills.
- Enthusiastic about learning new technologies with a passion for building secure software systems.

Skills

Languages and Frameworks
Ruby on Rails, Python, Advanced SQL, C, C++, Java, Scala, PHP, PL/SQL, C#, Bash, Javascript, XML, HTML, Stimulus.js, OpenPGP.js, Qt

Databases
Oracle, MySQL, SQLite, PostgreSQL, MongoDB, Cassandra, Snowflake

Tools & Technologies
Git, Github, Airflow, RabbitMQ, Kafka, Flink, Docker, Brakeman, Rubocop, Datadog, Docker, Jenkins, GoCD, Chef, VMware, Wireshark, NLTK, Weka, Memcache, Redis, Blockchain, Private/Public-key cryptography

Work Experience

11/2019 - 10/2024

Senior Software Engineer II, *Instacart*, Remote.
(Snowflake, Airflow, Python, Flink, Scala, SQL, Ruby on Rails, Postgres, Event sourcing, gRPC).

- Enhance and maintain key features for Instacart's Catalog, focusing on data ingestion, product identity, outsourcing flow, enrichment, and engagement across a 4-sided marketplace for CPG brands, retailers, consumers, and shoppers.
- Led a cross-functional initiative involving Catalog, ML, Operations and Fulfillment teams to improve product identity accuracy. Designed and implemented a semantic similarity model using embeddings and LLM for effective duplicate detection.
- Led the design and development of a comprehensive whole-file validation system to proactively detect and prevent inaccuracies in retailer pricing and availability data.
- Made significant contributions in migrating a high-volume legacy batch system, processing ∼4.5 billion price and availability events per day, to a Flink-based real-time system. This initiative delivered ∼$3M/year in cost savings.
- Designed and developed a suite of self-service tools for catalog ops, enabling them to effectively address ad-hoc production issues and significantly reduce on-call engineering ticket volume.
- Led Instacart's first internationalization (i18n) project for Catalog coordinating with cross-team stakeholders. This launch enabled new GTV opportunities and laid the groundwork for future i18n efforts.
- Facilitated team success by unblocking members, ensuring high standards for code quality and system architecture during reviews, leading incident response efforts, and collaborating with EM and PM to refine roadmap plans.

04/2017 - 10/2019

Senior Software Engineer IV, *Cisco Systems - Cisco Secure Endpoint*, Remote.

- Revamp endpoint policy module responsible for controlling endpoint malware scanner.
- Led the re-design and performance improvement of device trajectory module that allows visibility of endpoint activities for deeper malware analysis and identify entry-point of a malware.
- Led the re-architecture of CI infrastructure using GoCD. Apply distributed techniques to decrease build time by multiple folds. Better manage builds by adding dockerized containers, one-click and auto deploys.
- Led Rails 4 to Rails 5 migration and Ruby upgrade projects using incremental dual-boot approach.
- Introduce zero downtime migration using pt-online-schema-change and partition billion rows mysql table for 10x read performance.
- **Hackathon Winner**: Compute similarity score between two malwares based on their behavioral pattern.

04/2014 - 05/2016	**Senior Software Engineer**, *Cisco Systems - Cisco Secure Endpoint*, Calgary, AB, Canada. *(Ruby on Rails, JRuby, MySQL, MongoDB, RabbitMQ, Memcache, AWS, Cassandra, Jenkins, Chef, Cloud Security, Datadog)*.

- Designed and developed distributed database infrastructure for high availability and scalability involving clusters with 60+ Cassandra nodes and 15+ RabbitMQ nodes. Responsible for schema design, developing asynchronous firehose client and resolving performance bottlenecks to achieve write speed handling more than 2 billion events per day.
- Developed modules to apply Cisco's custom Indication of Compromise (IOC) rules to endpoint malware scanner and allow users to maintain list of IOC rules, schedule scan and view detailed scan results triggered by the IOC files.
- Led the integration with ThreatGrid (acquisition) for auto-sandboxing of executable files. Built custom rest api client gem specific to ThreatGrid API.
- Developed support-portal features that allows Tier3 support to view business metrics.
- Developed modules to identify vulnerable programs running on endpoints and alert users with detailed report.

02/2013 - 11/2013	**Software Engineer**, *Datagardens Inc, (acquired by CenturyLink)*, Edmonton, Alberta, Canada.

- Involved in developing backend modules for virtual machine replication and synchronization software responsible for disaster recovery and threat management for private and public clouds. (*C, pthreads*)
- Involved in doing performance tests of the existing system. Fixed issues in the existing linux block device driver and iSCSI configuration to improve the performance of the system. (*iostat, scstadmin, Vmware*)
- Involved in making architectural decisions for snapshot technology (Redirect on write, Copy-on write).
- Automated complex backend processes involving editing of Windows registry hive and converting Windows dynamic disk to basic. (*Python, Hivex, Testdisk*)
- Created proof of concepts for security project - live honeypotting with preserved memory state. (*Vmware FT, Snort, Ossec, iptables, Metasploit, Bash*)

Summer 2011,2012	**Research Assistant**, *University of Alberta, Canadian Obesity Network*, Edmonton, Canada.

- Prepared a dataset of obesity-related resources present in Canada. Built a crawler capable of extracting company names, URLs and addresses from Google search results and crawling web pages from the company websites. (*Python, cURL, lxml*)
- Built a system to classify websites into topical and non-topical categories using Support Vector Machine. (*Python, Weka, NLTK*)

02/2009 - 08/2010	**Software Architect**, *ITCorn Pvt. Ltd*, Kathmandu, Nepal.

- Involved in the system design and development of an interactive educational software. Added support for an elaborate set of mathematical and scientific symbols used in formulas and equations. Implemented an installer using InnoSetup and added security so that the software requires CD in the CD-ROM at all time to run. (*C#, .NET, SQLite, LaTeX*)

11/2007 - 01/2009	**Senior Software Engineer**, *D2Hawkeye Services (acquired by Verisk Health)*, Kathmandu, Nepal.

- Analyzed health care data for US-based insurance companies to generate financial and health related custom reports (HEDIS). (*Oracle, PL/SQL, SQL tuning*)
- Award: ***"Most Promising and Outstanding Employee"***

Academics

09/2010 - 09/2012	**Master of Science in Computing Science**, *University of Alberta*, Canada.
	Research: Data mining, Information Extraction, Machine Learning
07/2003 - 07/2007	**Bachelor of Engineering in Computer Engineering**, *Kathmandu University*, Nepal.
09/2010 - 04/2012	**Teaching Assistant**, *Department of Computing Science, University of Alberta*, Canada.

- Conducted labs as the Primary TA, marked assignments and proctored exams. Helped students during active learning sessions in the lecture. (*Programming, Boolean logic, Circuits*)

2012	**Publications**.

- **C. Thapa**, O. Zaiane, D. Rafiei, and A. M. Sharma. **Classifying Websites into Non-topical Categories**. In Proceedings of the 14th International Conference on Data Warehousing and Knowledge Discovery, 2012.