

Pratik Vagyani

Software Development Engineer

PROFESSIONAL EXPERIENCE

Oracle America | Software Developer Engineer III Feb 2020 - Till

Log Parser

- Contributed to design and development of distributed log parser to process **200TB** of data every day from **40+** regions.
 - Built highly available, fault tolerant system on Kubernetes cluster **80 nodes**
 - Prepared robust & adoptable code to support **5+** different log formats.
 - Contributed in CI/CD setup and platform for unit, E2E tests.
 - Implemented monitoring with Alarms, Metrics and Grafana dashboard.
 - Coordinated with partner teams for data access and log formats.
 - Mentored team members to onboard project, handle operational efforts.
- Kubernetes, Java, Docker, SQL, MQL, Grafana, OCI, OKE, Identity

Unified Data Processor

- Designed and developed Spark apps to analyze cloud services data and provide insight on performance of all services with **300 TB** of data size.
 - Optimized data storage to enhance Spark job runtime by **7 hrs (90%)**.
 - Solved complex & data intensive problems with elegant, optimized spark logic.
 - Build business insight for executives by relating different big datasets.
 - Build pipelines for **10+ reports** with efficient sql procedures to help customers visualize service and client tool efficacy faster and in multiple perspectives.
 - Collaborated with Customers for requirement gathering and onboarding
- Java, Python, Spark, SQL, OCI DataFlow

Symantec, India | Software Engineer II Jan 2015 - July 2018

Signer Classification

- Developed Spark app to classify **10 million** digital certificates which saved **200 hr/week** of human work.
 - Scaled spark app to reliably process **10 TB** data in 1 hour on a daily basis.
 - Reduce query execution time by creating a stage table out of **50 TB** data.
 - Improved performance on Hadoop up to **30%** by optimizing spark queries.
 - Implemented producer to publish malicious certificates to Kafka topic.
- Java, Scala, Hadoop, Spark, Hive, Oozie, Data Analytics, Shell, Linux

Digital Certificate Classifier

- Implemented a Controller-worker system on AWS for file processing on a large scale to detect malicious patterns in file.
 - EC2 auto-scaling & SQS used to operate cost-effectively (25%).
 - Developed python script to update tools & email daily stats for visualization.
 - Performed additional role of scrum master for an Agile team.
 - Collaborated with DevOps team for rapid infrastructure development.
- Python, AWS, SQS, S3, Kafka, Unit-test

Academic Project

ExtraShift — Distributed publish-subscribe

- Implemented distributed pub-sub system to solve the issue of open shift by connecting employees and campus restaurants.
 - Built a network of 4 docker containers and established communication between dockers by **socket programming** in python.
 - Setup MySQL database as data backup for disaster recovery.
- Python, Docker, Socket, flask, JavaScript, MySQL, github

716-430-2274
pratik.vagyani2nov@gmail.com
Seattle, WA
[linkedin.com/in/pratik-vagyani](https://www.linkedin.com/in/pratik-vagyani)
github.com/pratik027

EDUCATION

University at Buffalo, SUNY

M.S in Computer Science
Jan 2020 CPA - **3.58**

Walchand College of Engg, India

B.Tech in Information Technology
May 2015 GPA - 7.68/10

SKILLS

Programming: Python, Java, C++, C

Cloud: OCI, DataFlow, OKE, AWS, SQS, SNS, EC2, S3

Big Data: Hadoop, Spark, Kafka, Hive, HDFS, Oozie, Sqoop, Data Analytics, Parquet, ETL pipelines

Web: JavaScript, NodeJS, Django, Bootstrap, PHP HTML, CSS, RestApi

Database: MSSQL, MySQL, PostgreSQL, MongoDB, Hive, SQL

Other: Kubernetes, Docker, VM, Git, Jenkins, Maven, Jira, Spring, JSON

OS: Linux, Windows

SDLC: Agile, Scrum

PATENT & AWARDS

• Patent granted. **US10250588**

System and Methods for Determining Reputation of Digital Certificates Signer

- 2 Awards : "Passion & Focus" for excellent work done on multiple projects at Symantec

- Accolade: "Openness & Collaboration" for partnering with multiple teams at Symantec.

Course

- Algorithms
- Distributed Systems
- Data Intensive Programming
- Data Structures
- Databases
- Machine Learning
- Software Engineering

Hobbies

