SHANTANU LALITKUMAR JAIN

shantanu.jain3597@gmail.com | linkedin.com/in/shantanu-jain/ | shantanu-jain-2142.github.io/ | +1(646)-764-4875

EDUCATION

Columbia University, New York, NY May 2022
M.S. in Computer Science | GPA: 4.15
Coursework: Distributed Systems, Advanced Databases, Data Analytics, Cloud Computing, Machine Learning, Deep Learning

Savitribai Phule Pune University, Pune, India Jul 2019
B.E. in Computer Engineering | GPA: 3.82
Coursework: Data Structures, Discrete Mathematics, Object Oriented Programming, Machine Learning, Database Systems

TECHNICAL SKILLS

Programming Languages: Python, C++, Java (incl. Android Development), SQL
Web Technologies: Amazon Web Services, Google Cloud Platform, JavaScript, NodeJS, HTML, CSS
Software Tools/Frameworks: Docker, Kubernetes, Kafka, Spark, Terraform, Terragrunt, Fluentd, Hadoop, MongoDB, Git, CI/CD

PROFESSIONAL EXPERIENCE

Cohesity Inc., *Software Engineer* - San Jose, CA Jul 2022 – Jun 2023

- Collaborated as a member of the data archival team, overseeing archival and restore to and from the cold tier targets, utilizing efficient compression, encryption and deduplication strategies.
- Developed a feature to validate API permissions for external archival targets registered on the Cohesity platform, such as AWS, GCP, Azure, NAS, and QStar targets, with a customer focused approach.
- Contributed to the development of a transactional system in C++ that synchronized local and cloud data, implementing fallback mechanisms and utilizing intents to ensure consistency between the two states.
- Enhanced Cohesity's SaaS platform (Helios) by enabling the expansion and contraction of Cohesity Clusters across multiple cloud platforms, including AWS, GCP, and Azure.
- Ensured multi-tenancy and scalability by implementing Kafka queue and service worker threads.
- Responsible for end-to-end design of feature, from design documentation to unit and integration testing.

Siemens PLM Software, *Software Engineer Associate* - Pune, India Jul 2019 - Dec 2020

- Deployed and maintained the backend infrastructure on AWS as part of DevOps in the product research team.
- Developed a serverless architecture for automated maintenance of the entire MongoDB infrastructure.
- Migrated the infrastructure to various AWS regions using Terraform with Terragrunt.
- Created a deployment pipeline using Gitlab CI/CD to allow automatic detection of security vulnerabilities in the codebase.

ACADEMIC PROJECTS

CodeNote: Convenient snippet storage for developers Sep 2021 – Dec 2021

- Developed a python-flask webapp for easy storage, access, search, share and linting of code snippets.
- Employed GitHub actions for CI and CD, with unit testing, API testing, linting and deployment to AWS EC2.
- Utilized S3, Dynamo DB and Elasticsearch for storage along with Cognito for 2FA.

Flagged Post Analysis: Stackoverflow Sep 2021 – Dec 2021

- Analyzed 18 textual, code-based and user-based features for automatic identification of low-quality posts.
- Deployed the entire dataset of 96GB on BigQuery with data processing, feature extraction and ML pipeline on Spark.
- Used LSTM Encoder-Decoder model for labelling, and logistic regression for classification with an accuracy of 73%.

TeleEasy: Patient Portal Mar 2021 – Apr 2021

- Created a distributed subscription-based teleconsultation platform deployed entirely on Amazon Web Services.
- Utilized the theory behind buffet economics to subsidize consultation costs for the patients.
- Employed best practices for security (2FA) and scalability (SQS, ES and DynamoDB) for ensuring seamless experience.

Voice Controlled Photo Album Mar 2021 - Apr 2021

- Created a web application hosted on S3, for uploading and searching photos using text and voice inputs.
- Quick search access through elasticsearch, with authentication and authorization configured through AWS Cognito.
- Optimized the workflow, by using AWS Codepipeline for continuous integration and deployment.

E-Voting using Blockchain Feb 2021 - Mar 2021

- Designed a scalable and secure online voting system based on blockchain and smart contracts in Python.
- Synchronized blockchain across states with a combination of Merkle trees and proof-of-work consensus algorithm.
- Presented blockchain voting system as part of a peer-to-peer network of miners and users.