Aviral Srivastava

• Github: kebab-mai-haddi • LinkedIn: /in/sriavi • Website: aviralsrivastava.com
• Email: hi@avisri.com • +1-617-283-3811 • Blog

EDUCATION

Boston University, Boston, MA.   Sep 2019 - Dec 2020
- MS, Computer Science. CGPA : 3.77/4
  - **Subjects**: Graduate Distributed Systems (MIT), Advanced Analysis of Algorithms, etc.

Vellore Institute of Technology, Chennai, India.   Jul 2014 - May 2018
- B.Tech, Computer Science and Engineering. CGPA : 3.56/4

SKILLS

- **Languages**: Python, Scala, C++, Go, Java, PHP, Rust, SQL, MySQL, PostgreSQL
- **Technologies**: Spark, Dask, Aurora, Mesos, Kubernetes, Docker, Vagrant, Argo, Airflow, Kafka, Pulsar, Celery, Flink, Snowflake, Cassandra, RocksDB, PrestoDB, KyotoDB, Redis, Django.
- **Platforms**: AWS, GCP, Azure, On Premise Data Centers

WORK EXPERIENCE

Twitter, Software Engineer II, Graph Storage(Platform/Real Time Storage), Seattle   Apr 2021 - Present
- We are replacing the decade old Flock (Graph Storage service) with a new service (Flock v1.5) that will scale the write pipeline, improve data consistency, and reduce human ops.
- Led setting up v1.5 as a multi-Data Center setup.
- Proposed, designed, and developed Backup Verification in **Scala**: a service to verify backups as they (backups) are used in v1.5 to serve live traffic.
- Performed oncall operations such as rebalancing shards, restoring servers during heatwave, modifying mirrorsets to cater fast product iterations, etc.
- Led the team for building a managed graph database offering for analytics purposes, graph machine learning, and applications. Reduced onboarding time from 1 week to minutes, for data teams at Twitter.
- Led Requirements Engineering and drove the Production Readiness Review.
- Supervised request based authorization for Flock data, mentored two interns.

Redhat, Software Engineering (Data) Intern, Internal Data-Hub team, Boston   May 2020 - 20201217
- Re-partitioned the schema of SOS reports (**10 Trillion+ rows**, **100+ TeraBytes** data) to ensure high data availability and low latency. **Reduced** the count of S3 objects by more than **90%**: from 15M+ objects to < 1M objects.
- Standardized data workflows for the internal teams using **Argo, Spark, JupyterHub, Kubernetes** that led to scalable visualization and alerting dashboards.

Boston University, Research Associate (Software Engineering), Boston   Sep 2019 - May 2020
- Got funded by MIT-ICORPS for commercializing the research for improving software quality.

Atlan, Software Engineer(Data)   May 2018 - Jul 2019
- Atlan is a data democratization startup that helps data teams collaborate frictionlessly on data projects.
- Ensured the development and on-premise deployments of dashboards like United Nations SDGs monitoring, Ujjwala Dashboard using **Python3, Kubernetes, Docker and AWS S3**.
- Integrated serverless JupyterHub: a multi-user hub that spawns, manages, and proxies multiple instances of the single-user Jupyter notebook server. Stack: **papermill, Python3, Docker, Helm, Kubernetes**. Used by companies like **Unilever, Mindshare, McKinsey & Co.**
- Built and deployed Data Repository: a version control system for tabular data, including features like 'git diff', rollback, at a scale of over **billion rows in one cycle**. Stack: **Python3, S3, RocksDB, KyotoDB, Hive metastore and PrestoDB**. This scaled the production releases by **300%**: 1 project release to 3 projects release per quarter.

Centify, Co-founder, CTO   Mar 2018 - Jul 2018
- Audience Engagement tools which enable digital media platforms to integrate social media reactions with video or textual content.
- Built the complete system for processing social media content and applying emotional as well as sentimental analysis. Got featured on IBM's developer blog

PROJECTS

- List of all Projects