

Li Jun Huang

🏠 lijunhuang.com
✉ lhuang210@gmail.com
🌐 [linkedin.com/in/hlivt](https://www.linkedin.com/in/hlivt)

EDUCATION

Virginia Tech

BS, Computer Science
Minor Math
Phi Beta Kappa
Grad. May 2019
GPA: 3.89

Georgia Tech

MS, Computer Science
Grad. May 2022
GPA: 4.00

SKILLS

Azure, Synapse, Databricks
Event Hubs
Apache Spark, Delta
Scala, C#, Python, SQL

EXPERIENCE

Microsoft

Senior Software Engineer Software Engineer II

Incentives Data Platform

Sep 2024 - PRESENT
Oct 2021 - Aug 2024

- Led the platform API facilitating cross-team communication on data movement & completeness, enabling external teams to go from schedule trigger to event driven data processing model with comprehensive insights on data lineage.
- Collaborated cross teams to develop & deploy data system for Service Account Management using Apache Spark Streaming & Event Hubs, improve E2E performance from 72h to 15m for averaging daily 1M transactions.
- Developed near real time data processing system for direct pay partners and sellers with Apache Spark Streaming, Delta Lake, and Event Hubs to processing monthly 5M earning transactions and 300M earning amounts and improve E2E performance from 38h to 1h.
- Introduced Spark Scala client library for both producer & consumer teams to access data platform in unified manner. Improved overall platform standardization and external teams' on-boarding & development process.
- Manage data infra and pipelines to support MSFT partner & seller financial data processing, storage, and querying layer, totaling 334M daily transactions, with Spark, Scala, Event Hubs, Azure Synapse, Data Explorer, Delta table, & Data Lake.

Software Engineer

Internal Financial Reporting

Aug 2019 - Oct 2021

- Developed system & data pipeline that continuously process & analyze TBs of financial data to generate End-of-Period report to meet MSFT business need with Spark, Scala, U-SQL, Databricks, Azure Data Factory & Data Lake. Improved E2E performance for monthly closing period SLA from 10+ hours to 2+ hours.
- Developed domain context reconciliation, data quality check & triangulation to improve E2E report accuracy & compliance.
- Developed JEDI compliant CICD & infra automation to support Microsoft Federal financial data ingestion.

Engineer Intern, Viasat

May 2018 – Aug 2018

- Developed Kafka internal streaming platform API using Dropwizard & Java.
- Enhanced API scalability with http2, cache, object pool, metrics & health check.
- Built corresponding Jetty API client compatible with Avro, JSON, and POJO.

Undergraduate Research Assistant, Virginia Tech

Jul 2017 - May 2018

- URA for T. M. Murali's Computational Systems Biology research group.
- Co-authored two F1000 Research papers on PathLinker-Cytoscape.
- Led the PathLinker-Cytoscape and CyGraphSpace major releases using Java.
- Created the PathLinker-Cytoscape REST API using the Swagger framework.