Shiva Tarun Soka
starun.9515@gmail.com | +1 (605) 371-6031
https://www.linkedin.com/in/shiva-tarun-soka-
496aa0279?lipi=urn%3Ali%3Apage%3Ad_flagship3_profile_view_base_contact_details%3BeE4PS%2FaFQmmjYrRdjstIcA%3D%3D

Summary:

- Over 5 years of professional experience in in Python programming and hands-on experience with Apache Spark, Scala, and Hadoop.
- Worked on streaming data pipelines using Apache Kafka and Spark Streaming for real-time analytics.
- Extensive experience with HDFS, Hive, Pig, Sqoop, MapReduce, and Yarn for large-scale data processing.
- Designed and implemented ETL workflows for data ingestion, transformation, and storage in HDFS and data lakes.
- Proficient in databases like MongoDB, Cassandra, HBase, and SQL-based databases for data storage and retrieval.
- Extensive experience with cloud computing platforms such as AWS, including EC2, S3, RDS, Lambda, EMR, and Cloud Formation for deploying scalable solutions.
- Implemented serverless architectures using AWS Lambda, API Gateway, and DynamoDB to optimize application performance and reduce infrastructure costs.
- Experience with AWS services like EC2, S3, EMR, RDS, and Lambda for scalable data solutions.
- Developed RESTful APIs using Python frameworks like Django and Flask for data processing applications.
- Built end-to-end data pipelines integrating Kafka, Spark, and Hadoop for efficient data processing.
- Hands-on experience with Docker and Kubernetes for containerization and orchestration of microservices in cloud environments.
- Optimized Spark jobs, SQL queries, and data pipelines for improved efficiency and lower latency.
- Managed AWS IAM roles and policies to ensure secure access control for cloud-based applications.
- Experience with version control (Git), Jenkins, and automation tools for deployment and monitoring.
- Strong knowledge of unit testing, test-driven development (TDD), and debugging techniques in Python and Spark applications.
- Worked in Agile environments, collaborating with cross-functional teams using Jira and Confluence.
- Ability to troubleshoot complex data issues, optimize workflows, and enhance system performance.
- Excellent teamwork skills, technical documentation expertise, and ability to convey complex concepts effectively.

Technical Skills:

Programming Languages	Python, Scala, Java, C, C++
Big Data Technologies	Apache Spark, Hadoop, Hive, Pig, Sqoop, MapReduce, HDFS, Yarn
Streaming & Messaging	Apache Kafka, Spark Streaming
Databases	MySQL, MongoDB, Cassandra, HBase, Oracle
Web Technologies	HTML, CSS, JavaScript, AngularJS, Bootstrap, JSON

Frameworks	Django, Flask, Spring, Hibernate
Cloud & DevOps	AWS (EC2, S3, EMR, RDS, Lambda), Docker, Kubernetes, Jenkins, Puppet
Testing & Automation	JUnit, Cucumber, Jemmy, Selenium, Automated Testing
Parsing & Data Processing	SAX Parser, DOM Parser, PySpark, Pandas, NumPy
Version Control & CI/CD	Git, Subversion (SVN), Jenkins
Agile & Project Management	Agile, Scrum, Jira, Confluence
Other Tools	UML, Rational Rose, cProfile, API Development, Debugging, Troubleshooting

WORK EXPERIENCE:

Python Developer   **Aug 2024 – Till Date**
Quantum Capital Group
Responsibilities:

- Involved in requirement gathering, system analysis, design, development, testing, and deployment of scalable applications.
- Developed web-based applications using Python and Django for handling large datasets and data-driven decision-making.
- Built data pipelines using PySpark, MapReduce, and Hive for efficient data transformation, cleaning, and processing in Hadoop ecosystems.
- Optimized Spark jobs developed Python scripts for automation and used NumPy for numerical analysis in financial applications.
- Designed and managed MySQL databases, implemented data validation scripts, and connected databases to Spark for real-time processing.
- Designed and developed responsive user interfaces using HTML, CSS, JavaScript, AngularJS, Bootstrap, and JSON.
- Worked in a DevOps environment using Git, Jenkins, Puppet, Subversion, and AWS CloudFormation for continuous integration and deployment.
- Configured and deployed applications on AWS EC2, integrated S3 for scalable storage, and used RDS for database management.
- Utilized AWS Lambda for serverless computing, reducing operational overhead and improving cost efficiency.
- Implemented security best practices using AWS IAM, CloudTrail, and CloudWatch for monitoring and access control.
- Configured and deployed applications on Amazon EC2 and integrated AWS services for scalable data solutions.
- Collaborated with teams in Agile environments, participating in sprints, stand-ups, and managing tasks using Jira.
- Developed unit tests, automated test scripts, and regression tests to ensure high-quality code and system stability.

- Implemented logging using Python's logging library, profiled applications using cProfile, and monitored system performance.
- Worked cross-functionally with developers, testers, and business analysts to troubleshoot issues, optimize workflows, and improve system performance.

Environment: Python, C, C++, MySQL, Linux, HTML, AWS, CSS, JavaScript, Django, Java Servlets, Beans, SAX Parser, DOM Parser, Oracle Database, Cucumber, Jemmy, JUnit, AWS (EC2, S3, RDS, Lambda, EMR, Cloud Formation), Docker, Kubernetes, Automated Testing, Web Development, Data Parsing, API Development, Debugging, Troubleshooting.

Python Developer   **Feb 2020 – Dec 2023**
ICICI Bank, Hyderabad, IN
Responsibilities:
- Strong understanding of Object-Oriented Analysis and Design (OOA/OOD) principles with experience in UML diagrams like Use-Case, Class, and Object diagrams.
- Hands-on experience with Python, C, and C++ for developing software components and implementing APIs.
- Built user interfaces using HTML, CSS, and JavaScript, ensuring responsive and interactive web applications.
- Designed and managed MySQL databases, writing queries for efficient data storage and retrieval.
- Created and managed business logic using Java Servlets and Beans for web applications.
- Used SAX/DOM Parsers to extract and process data from various sources into an Oracle database.
- Developed step definition code in Cucumber, utilizing Jemmy and JUnit for automated testing.
- Built web applications using the Django framework, following best practices in Python and MySQL integration.
- Developed and deployed cloud-based applications using AWS services like EC2, S3, RDS, and Lambda.
- Utilized Docker and Kubernetes to containerize and manage applications in a cloud-native environment.
- Monitored and resolved data loading issues in applications and improved automated tools with feature enhancements and bug fixes.
- Worked effectively in a team environment, demonstrating strong analytical skills and a proactive approach to learning new technologies.

Environment: Python, C, C++, MySQL, Linux, HTML, CSS, JavaScript, Django, Java Servlets, Beans, SAX Parser, DOM Parser, Oracle Database, Cucumber, Jemmy, Junit, AWS (EC2, S3, RDS, Lambda), Docker, Kubernetes.

Education:
- Graduated in Computer Science from ST. Martin's Engineering College.
- Masters in Computer Science from University of South Dakota.