DURGA SHANKER

Staff Engineer, Data Analytics Engineering
SanDisk | Western Digital

☎ 8123347938
✉ itisds.ai@gmail.com
in linkedin.com/in/itsds

PROFESSIONAL SUMMARY

☞ Having overall **8.2+ Years** of experience in **Big Data** and **Java/J2EE Technologies.**
☞ Good working knowledge in **Object Oriented** Application Development based on **SOLID Principle**, Maintenance life cycle process and **Agile SDLC**.
☞ **3+ Years** of working knowledge on **Big Data Technologies** like **Spark, Scala, Hadoop, MapReduce, Hive, Kafka, Airflow.**
☞ Knowledge in **Unix Commands** and **Hadoop Cluster Commands**.
☞ Very good Knowledge on **SQL** and work experience with **MySQL** and **Oracle**.
☞ Experienced in working with **AWS, GKE** on **GCP.**
☞ **Customer handling experience** in technical and business issues.
☞ Mentored junior employees on business and Technical functionalities.
☞ Team player with excellent team working skill and good at time and resource management.

SKILLS

Coursework	:	Data Structures, Design & Analysis of Algorithms.
Languages	:	Java , SQL, Scala, Python.
Big Data ES	:	Spark, Hive, HDFS, YARN, MapReduce.
Messaging System:		Kafka, Kafka Streams.
Orchestration tool	:	Airflow.
Web Services	:	RESTful with Jersey implementation of JAX-RS.
RDBMS	:	MySQL.
Monitoring Stack	:	Splunk, Grafana.
IDE	:	IntelliJ, Pycharm, SQL Workbench/J, Eclipse.
Operating System	:	MacOS, Windows, Linux.
Cloud	:	GCP-GKE, AWS.
Version Control	:	Git, GitHub, BitBucket, SVN.
Tools	:	Pandas, Kubernetes, Docker, Airflow, Anaconda, JIRA, Lens.

WORK HISTORY

SanDisk	Apr'19 - Present	Bangalore, India
Infosys	Jun'17 - April'19	Bangalore, India
HCL Technologies	Nov'14 - Jun'17	Bangalore, India

KEY PROJECTS

Data Pipeline Development	April 2019 – Present

Technology – Java**,** Hive, Spark, Kafka, Avro, YARN, Scala, DC/OS, Airflow, Git.

Description :

🖹 ETL library decodes & extracts data from logs, transforms the extracted data into tabular format & outputs result to facilitate fast data loading into Big Data.
🖹 ETL used by BDP pipeline to decode logs & ingest data into Redshift and Presto.
🖹 GHL2CSV CMD tool decodes logs & outputs data in CSV/JSON file using ETL.

Roles and Responsibilities :

- I am part of DAO responsible for development and maintenance of Western Digital's Big Data platform applications. An essential part of this Big Data system is writing, extracting and transforming (ETL) raw data into its final materialized format suitable for machine learning analytics. Work in a strong teamwork environment with Agile Scrum methodology to manage tasks assignments. I am responsible to design and build big data infrastructure and ETL processes in production.
- Developed pipeline using Kafka streaming where above materialized data will get populated to Kafka Topic and it will get consumed by different Team members.
- Developed Spark application where above materialized data will be reprocessed as per the ML Model requirement and then send processed data to Redshift warehouse.
- Written Airflow DAG to get Unsupported GHL Log from Hive Database and mailed those log to ETL Team members.
- Written Airflow DAG to Create Jira ticket by consuming JIRA REST API for all the Unsupported GHL Log from the above DAG.
- Implemented OAuth Authorization to consume JIRA REST API.

KEY PROJECTS	**Horizon's App Development & Monitoring**	**April 2019 – Present**

Technology - Java, Kafka, Python, Docker, Kubernetes, GCP, GKE , BitBucket, Splunk, Grafana, Mesos DC/OS.

Description – It is DAO's platform for bringing Big Data Analytics from cloud to real-time factory environment. So basically it is the platform where Applications Developer creates docker images, deploy and run their application in the streaming way. Horizon Platform running on **GCP** and leverages the concept of **Google Kubernetes Engine(GKE)**.This platform let us manages various parameter of the application like Manage Capacity, Manage uptime, Manage Releases, Handle Errors, See Status, Get the data you need, Restart your services etc. So, the goal horizon to provide an ecosystem to Application developer so that they don't have to do everything from scratch. Modules under this project are Horizon-Monitoring Stack, Horizon-Inspector, Horizon-dae-streaming, Horizon-Sieve etc.

Roles and Responsibilities :

- I am working as Team Lead as well as Developer of this project.
- Monitoring all the Horizon app through Grafana and our Monitoring tool.
- Developed and deploy various applications like dae-streaming ,inspector etc.
- Performed Memory and CPU consumption Analysis and Tuning.
- Effectively managed code released from Development to Staging and Staging to productions.

KEY PROJECTS	**SIEVE**	**May 2021 – Present**

Technology – Python, Kafka, Docker, Kubernetes, GKE, Pandas, DuckDB, Splunk.

Description – It is an streaming application that runs Paper Screen on the Manufacturing Data. It is means of keeping a HDD quality check before shipping it to customers. The purpose of this project is run screen on manufacturing data at volume and then send output to Manufacturing Team. They will use these output for RCA and

reprocess those drive again. We are leveraging concept of **Docker** and **Kubernetes** Technologies to run this project on **GCP** using **Google Kubernetes Engine(GKE)**. Development involves using **Kafka Real time streaming** input data along with automated user query inputs. We are using Splunk and Grafana to monitor behaviour about this application. It includes three Modules i.e **Sieve**, **Sieve-Screens** and **SVT**.

Roles and Responsibilities :

- Led and Developed Sieve Team enabling implementation in the Streaming way using Kafka Streaming and Python.
- My responsibilities includes gathering the requirements from Thailand Team and distribute/work on those requirements accordingly.
- Effectively managed code released from Dev to Stag & Stag to Production.
- Attend regular call with Thailand Team to discuss about current progress, new requirement or any issues/concerns they might have.
- Created an automated paper screening app for the global-internal Horizon platform also known as for the manufacturing HDD floor in Thailand.
- Created an Splunk dashboard to monitor it at regular interval.
- Enhance performance by replacing DuckDB with PandaSQL.

KEY PROJECTS	**TARS**	**June 2017 – April 2019**

Technology - Java**,** RESTful, Spring Boot, Hibernate, JPA, MySQL, Spring Core, Git.

Roles and Responsibilities :

- I worked on development of TARS application by using Java/J2EE tech. from scratch.
- Involved in the requirement study, design, development and documentation.
- Interacted with client to give demo for TARS application.
- Involved in integration of backend code with front end.
- Weekly call with client to discuss on the progress and status.
- Involved in the release procedures and prepared release documentation.
- Involved in Document Preparation i.e. System Architecture, Solution Requirement, Test Scripts.

KEY PROJECTS	**Admin Apps**	**Nov 2014 – June 2017**

Technology - Java**,** RESTful, Spring MVC, Hibernate, JPA, MySQL, Spring Core, Git.

Roles and Responsibilities :

- Worked on design and development of Admin Apps application from scratch. This Application is used to maintain associates details and travel maintenance system, which helps maintaining to travel details of the employee. And this is role based application, based on the role assigned to users. They can access the website menus/tabs.

EDUCATION

- JSS Academy of technical education, Bangalore, **B.E. in CSE**, 2014.
- Holy Mission High School, Samastipur, **HSC (10+2)**, 2009.
- North Point Children's School, Muzaffarpur, **SSLC**, 2007.

- Programming.
- Very good in Problem Solving and **Mathematics**, secured **95%** in **10+2**.
- Quick learner and eager to learn new Technologies.
- Got **Appreciation Award** from client for good performance in project since beginning.
- Certified **Nutrition & Fitness Consultant** from **INFS**.

DECLARATION

I hereby declare that the above-mentioned information is correct and true to best of my knowledge and belief and I take complete responsibility for the correctness.

Place: Bengaluru   **Sign: Durga Shanker**