Joel Victor

Pune, Maharashtra, India | joelsvictor@gmail.com | 7038063827 | LinkedIn

Introduction

I'm a fast learner with extensive experience in big data architectures and distributed systems, skilled at designing scalable, maintainable software. I prioritize understanding business needs to deliver solutions that make an impact. Always eager to explore new technologies to tackle complex challenges.

Experience

Tech Lead, TeamOhana – Remote   July 2021 – Present

- Founding engineer at TeamOhana.
- Developed the backend from the ground up. This included designing and implementing a robust and scalable multi-tenant database schema in PostgreSQL.
- For login, authentication and authorization Auth0 was integrated into the webapp.
- Designed and built a robust ABAC & permissions system.
- Worked on integrating various HRIS (BambooHR, Workday), ATS (Greenhouse, Ashby, Lever), and Payroll systems (ADP) with TeamOhana
- Developed APIs for seamlessly Okta SCIM integration enhancing security and user management.
- Designed and developed an AI copilot module using GPT-3.5 and techniques like zero shot prompting and RAG.
- Evaluated software solutions to align with business requirements and documented outcomes.
- Involved in the process for upgrading PostgreSQL without any downtime.
- Responsible for mentoring and ensuring the team designs maintainable, scalable and testable software.

Senior Data Engineer, Helpshift – Pune, Maharashtra, India   June 2015 – June 2021

- Founding member of the Data Platform Team.
- Along with two other team members, rearchitected the analytics platform to scale. The new platform used AWS Redshift and then moved to Hadoop infrastructure for cost optimization. Managed the Hadoop infrastructure using Hortonworks Data Platform (HDP).
- Over my tenure, I was involved in various projects that involved building data pipelines to compute metrics for existing and new features. These metrics provided customers with insights about the value of adopting Helpshift into their product.
- Led the effort to optimize the data pipelines to reduce data processing latencies. Various strategies like using ORC file formats, improved partitioning of the data and tuning the parameters for the execution engines like Spark, Tez and MR. These improvements reduced the runtime of the data pipelines by 50%.
- Built and integrated APIs to provide analytics data for PowerBI or customer platforms.
- Built Content Packs in PowerBI. This involved building data models, dashboards in PowerBI and writing Power Query M code.
- Evaluated PowerBI Embedded for integrating PowerBI dashboards into the web app.
- Worked on uploading billing data into Oracle Netsuite using the Netsuite SOAP API.
- Involved in the effort of upgrading HDP without downtime and then later moving from HDP to AWS EMR.
- Responsible for mentoring and ensuring the team designed maintainable, scalable and testable software.

Software Development Intern, Amazon – Hyderabad, India   June 2014 – Nov 2014

- Worked in the team that built services for optimizing inventory allocation in Amazon's Fulfillment Centers.
- Worked on reducing the runtime of a data pipeline that calculated the Fulfilled unfulfilled Demand (FUD). This pipeline used to take more than 8 hours during peak times like Black Friday.
- This was not acceptable to business as it resulted in delayed deliveries which in turn meant unhappy customers.
- The implemented optimizations reduced the runtime by more than 70%.

Certifications

IBM Data Engineering Essentials April 2024

Certificate

IBM Python for Data Science and AI April 2024

Certificate

IBM Python Project for Data Engineering May 2024

Certificate

Technologies

Programming languages: Clojure, Java

Big data technologies: Apache Spark, Apache Hadoop (MapReduce & HDFS) Apache HBase, Apache Hive, Apache Pig, Apache Phoenix, Apache Storm, Apache Kafka, Hortonworks Data Platform (HDP), Apache Flink, Apache Zookeeper, Apache Airflow

AWS: S3, Redshift, RDS (PostgreSQL), API Gateway, Quicksight, EMR, SES, Kinesis, Lambda, EC2, SWF, OpenSearch

Databases: PostgreSQL (RDS & Self Managed), MongoDB, Hive, HBase, Phoenix, Redshift

BI Tools: Quicksight, PowerBI

Others: Linux CLI, BASH & ZSH (Scripting), Git, Jenkins (Groovy DSL), Docker

Education

Pune University, Computer Science Department, Masters of Computer Application 2012 – 2015

- GPA: 4.8 out of 6.0

Fergusson College, Bachelors of Science in Mathematics 2009 – 2012

Activities

IN/Clojure 2020 - Intermediate Workshop February 2020

Conducted an 8-hour workshop, guiding attendees on how to build web APIs in Clojure and covered advanced concepts like refs, STMs, futures and promises. Successfully guided 7 attendees, who provided positive feedback for the hands-on approach to complex Clojure topics.