

# Ashish Sasi

Cloud & DevOps Specialist | DataOps Engineer | AWS Data Engineer

West Melbourne, VIC

0422731509 | ✉ [sasiashish1995@gmail.com](mailto:sasiashish1995@gmail.com)

[www.linkedin.com/in/ashish-sasi/](https://www.linkedin.com/in/ashish-sasi/)

## PROFESSIONAL SUMMARY

- Cloud and DevOps Engineer with 5+ years of experience delivering secure, scalable AWS infrastructure across data, analytics, and application workloads
- Skilled in managing microservices on ECS/EKS, building CI/CD pipelines with GitLab and GitHub Actions, and provisioning infrastructure using Terraform and CloudFormation
- Knowledgeable in ETL orchestration and streaming using Airflow, Glue, Kafka, and Kinesis, with a strong focus on monitoring, observability, and post-incident RCA to ensure data pipeline reliability
- Known for driving automation, enabling self-service through well-documented tooling, and implementing compliance and governance frameworks in enterprise cloud environments

## PROFESSIONAL EXPERIENCE

### DEVOPS/CLOUD ENGINEER - Versent Pty Ltd | Melbourne, Australia

May 2021 – Present

- Provide 24x7 AWS operational support, release management and infrastructure automation across enterprise client environments, with a focus on data processing and analytics workflows
- Developed and enhanced Python-based CLI automation tools for streamlined multi-client AWS account access control, reducing operational overhead by 80% across managed environments
- Created a YAML-templated CloudWatch alarm deployment system using Python scripts, standardizing infrastructure monitoring and reducing manual setup time by 90–95%

#### Key Client Projects:

##### **Solstice Energy (formerly TasGas) - Data Infrastructure & Automation Engineer**

- Executed database migrations using AWS RDS and DMS with full schema replication and validation
- Maintained and monitored ETL pipelines using AWS Glue, PySpark, and Lambda to automate reporting workflows across raw, curated, and trusted zones following Medallion data model
- Monitored and maintained integration of DynamoDB tables linked to Cognito-authenticated users into AWS Glue Data Catalog, ensuring governed analytics access, schema change tracking, metadata governance, and CloudTrail-based audit visibility through IAM-controlled access
- Implemented identity and data access workflows using IAM and Cognito, integrated with PostgreSQL and Snowflake for secure data consumption
- Built data quality monitoring dashboards and alerting systems using CloudWatch and Python to ensure pipeline reliability

##### **Flybuys - Microservices & Infrastructure Platform Engineer**

- Managed multi-account AWS environments using CloudFormation for seamless application and data pipeline deployments
- Provisioned and maintained autoscaling configurations for EC2 instances and EKS clusters, providing operational support during promotional spikes to ensure application stability and performance
- Monitored and troubleshooted real-time and batch data pipelines running on ECS/EKS, including Spark Structured Streaming jobs consuming Kafka and Kinesis streams, with downstream ingestion into Redshift for near real-time offer redemption and campaign analytics for data teams
- Led day-to-day support for Control-M data jobs, analysing CloudWatch logs, Datadog dashboards, EKS/EC2 metrics, and message queues to identify bottlenecks and resolve transaction failures across Oracle-based systems
- Automated recovery of failed Control-M workflows using Bash, Python, and SQL; conducted structured root cause analysis (RCA) and post-incident reviews to improve observability and reduce recurring failures by 80%
- Monitored and supported Apache Airflow DAGs driving scheduled transformations, while ensuring performance, observability, and IAM compliance across batch workflows and Redshift ingestion pipelines powering customer engagement dashboards
- Designed and maintained CI/CD pipelines using GitLab CI/CD for deploying microservices and provisioning SageMaker notebooks and AWS Workspaces for data analysts, following branching strategies, peer review workflows, and safe rollout practices
- Implemented secrets management, network policies, and IAM roles to enforce secure service-to-service communication

##### **Toll Group - DevOps & Infrastructure Engineer**

- Monitored and supported large-scale AWS infrastructure for a high-availability CargoWise environment, ensuring uptime and performance of production data workflows
- Handled EC2 patching, WAF and Network Firewall rule updates, and log ingestion pipelines feeding into Splunk and Athena
- Conducted reliability reviews and collaborated on cost-optimization efforts using AWS Cost Explorer and Well-Architected principles
- Maintained and updated CI/CD pipelines using GitHub Actions and Azure DevOps, supporting deployment integrity, rollback testing, and environment drift resolution
- Managed S3-based data lake for backup, archival, and log analytics workflows, implementing lifecycle policies and cross-region replication to ensure data retention and compliance aligned with ISO standards

##### **Colonial First State - Cloud Infrastructure & Reporting Specialist**

- Supported CFS mobile app and reporting platform by managing secure, scalable infrastructure and ensuring uptime across data-driven services
- Provisioned and managed SSL certificates using AWS Certificate Manager (ACM) to secure internal endpoints and mobile API integrations
- Automated Tableau dashboard deployments and data refresh workflows, replacing manual processes and improving reporting efficiency
- Built Lambda-based scheduled data delivery systems with Amazon SES to automate weekly business reporting

**DATA SCIENCE INTERN - Removify | Melbourne, Australia**

Feb 2021 – May 2021

- Performed text classification and sentiment analysis on Google and Glassdoor review data using Python-based machine learning pipelines
- Built end-to-end NLP workflows with NLTK, spaCy, and ML models (SVM, Random Forest), orchestrated via Apache Airflow for automated execution and monitoring
- Developed and trained convolutional neural networks (CNNs) for content scoring and relevance assessment

**JUNIOR CONSULTANT - TIBCO Software Pvt. Ltd.**

Oct 2017 – Jul 2018

- Developed REST/SOAP API integrations using TIBCO BusinessWorks for Citibank financial data processing systems and performed analysis using TIBCO Spotfire.
- Collaborated on enterprise data integration projects using TIBCO EMS for high-throughput data messaging environments
- Enhanced system reliability and performance through middleware optimization and data flow analysis

**TECHNICAL SKILLS****DevOps & AWS Services:**

- AWS Cloud Services: EC2, S3, Lambda, Glue, EMR, API Gateway, IAM, CloudWatch, Step Functions
- Infrastructure as Code (IaC): Terraform, AWS CloudFormation
- CI/CD & Automation: Git version control, GitHub Actions, GitLab CI/CD, AWS CodePipeline, Azure DevOps
- Containerization & Orchestration: Docker, Kubernetes, Helm
- Monitoring & Observability: Datadog, CloudWatch, Splunk
- Observability & Incident Management: CloudWatch Logs, AWS Athena (log analysis), Splunk, custom dashboards, alerting, SLAs tracking, Lambda log processors, PagerDuty

**Programming & Analytics:**

- Languages: Python, JavaScript, R, SQL, Bash – basic scripting for automation and routine operations
- Python Libraries: Pandas, NumPy, PySpark, Scikit-learn, Boto3, FastAPI, Flask
- Web Development: HTML, CSS, JavaScript, React
- Data Analysis & Visualization: Jupyter Notebook, Tableau, D3.js, QuickSight, SageMaker
- Statistical Techniques: EDA, regression, classification, data profiling

**Data Engineering & Big Data:**

- ETL & Pipeline Orchestration: AWS Glue, Apache Airflow, Step Functions, PySpark, and Dagster (foundational knowledge through self-paced learning and experimentation)
- Big Data Processing: Apache Spark, partitioning, bucketing
- Data Lakes & Warehouses: S3, Redshift, Lake Formation, Snowflake
- Data Modeling & Design: Star/Snowflake schema, schema evolution, normalization
- Data Streaming: Apache Kafka, Kinesis, real-time data processing
- Data Cataloging: AWS Glue Data Catalog, schema and metadata management, data lineage tracking

**Databases & Storage:**

- Cloud Databases: RDS, Aurora, DocumentDB
- Relational Databases: PostgreSQL, MySQL, Oracle SQL
- NoSQL & Semi-Structured: MongoDB, DynamoDB, S3 (as Data Lake)
- Query Engines: Athena, Redshift Spectrum
- Storage & Access Patterns: Indexing, partition keys, compression, lifecycle policies

**EDUCATION****Master of Data Science | Monash University, Melbourne**

2018 – 2020

*Focus: Machine Learning, Deep Learning, NLP, PySpark, AWS, Statistics***Bachelor of Computer Engineering | Pune University**

2013 - 2017

*Core: Data Structures, DBMS, Network Security, Operating Systems***CERTIFICATIONS**

- **AWS Certified Data Engineer – Associate** (July 2025)
- **AWS Certified Solutions Architect – Associate** (Apr 2025)
- **AWS Certified Cloud Practitioner** (Feb 2022 – Feb 2025)

**KEY PROJECTS****• ECKIDZ Climate Data Platform (2020)**

Developed AWS-hosted web application with integrated machine learning backend for environmental data analysis. Built data processing pipelines for food impact analysis and CO2 emission calculations using Python and scikit-learn. Implemented automated data collection and processing workflows using AWS Lambda and S3. Integrated backend with a web interface, enabling dynamic CO<sub>2</sub> footprint estimation and food impact scoring.

**• Sentiment Analysis for Predicting Restaurant Review Ratings (2019)**

Understanding Neural Networks to implement Convolutional Neural Networks and Recurrent Neural Networks with TensorFlow and Keras for performing Sentiment Analysis and generating labels for restaurant reviews (supervised learning). Performed Text filtering and Data Cleaning using Python for over 650,000 records and label prediction for 250,000 records.

**• Sales Forecasting & Inventory Optimization (2019)**

Built an Excel-based BI model to forecast sales and optimize inventory for a retail chain. Used Solver and time series forecasting to reduce stockouts by 25% and minimize holding costs. Automated reporting with macros for faster decision-making.