Mohanad Elsafty

Dublin, Ireland
+353-870002865
Email: mohanad.elsafty@outlook.com
LinkedIn: linkedin.com/in/mohanad-elsafty
Online Resume: mohan3d.github.io/resume

Employment

Software Development Engineer II (L5) | Amazon Web Services (AWS), Dublin — DynamoDB

Sep 2024 – Present
Dublin, Ireland

DynamoDB MemDS (Java, Rust, C++, Python)

- Engineered core **in-memory metadata service (Java)** for DynamoDB, serving **billions of requests/sec** with ultra-low latency.
- Designed and implemented **sharding architecture** to distribute metadata across hosts, reducing memory footprint and improving scalability.
- Built **shard discovery protocol (Rust)** for clients, including heartbeat and failure-detection mechanisms; delivered full end-to-end implementation.
- **Led deployment of sharding into production**, enabling dynamic metadata distribution and improved system reliability.
- Automated **deployments and scaling** across AWS regions; benchmarked new features and optimized low-level performance (ENI, RSS tables, IRQ).
- **Handled on-call operations**, resolving critical production incidents and troubleshooting infrastructure issues.

Technical Lead | Singtel

Aug 2023 – Sep 2024
Singapore, Singapore

Data Fabric (Apache Spark, Apache Hadoop, Apache Airflow, Amazon AWS)

- **Led the formulation of the project's strategic objectives** and aligned technical efforts with the business goals, ensuring the Data Fabric relevance to Singtel's long-term vision.
- Designed the foundational architecture for the Data Fabric, integrating **cutting-edge and cloud-native technologies and best practices** to create a scalable and adaptable data fabric solution.
- Orchestrated **cross-functional** collaboration among diverse teams.

Data Security Enhancement Initiative (Java, Scala, Python, Google Tink)

- **Developed an innovative** encryption solution to ensure the secure transfer of sensitive data from on-premise servers to the cloud. This solution encrypts all data before it is moved to the cloud, ensuring compliance with security policies and regulations.
- **Integrated with existing frameworks**, including **Apache Spark, by providing User-Defined Functions (UDFs)** in Java for encrypting, decrypting, and HMAC operations on DataFrame columns.
- **Developed a Kafka plugin** to encrypt or HMAC messages during transmission, ensuring data security from source to sink. Decryption is performed at the receiving end, maintaining data integrity throughout the data pipeline.
- Utilized **Google Tink library, to align with Singtel's security policies** and ensure robust data protection.

Senior Data Infrastructure Engineer | Shopee

Jan 2021 – Aug 2023
Singapore, Singapore

Apache Spark (Scala, Java, JavaScript)

- **Dramatically enhanced** Spark history server for handling millions of records effortlessly, resulting in an **80% performance improvement** with larger datasets.
- Enabled search and filter on the client side, reducing server load by **45% and enhancing** user experiences.

Apache Hadoop/HDFS (Java, Airflow, Kubernetes)

- Developed logs aggregation for Yarn/HDFS logs, **leading to more efficient monitoring and saving 3 hours weekly** .
- Designed and implemented a low-memory fingerprint for extended offline analysis of large size FSIMAGE, allowing **the biggest FSIMAGE to fit within just 30% of its previous memory footprint**.
- Engineered a block-based InputFormat for DistCp (MapReduce application), optimizing workload distribution among mappers based on blocks count exploiting **data locality and shortening data transfer time by 20%**.
- Enhanced DFSClient and NameNode to effectively mitigate the impact of sluggish DataNodes on read/write operations.

Apache Ozone (Java, C++, Golang, JavaScript)

- Engineered a sophisticated petabyte-scale data Retention Service akin to AWS S3 lifecycle configurations, **saving 15% of wasted** storage. Created AWS CLI/SDK compatible s3gateway endpoint for easier integration with existing clients.
- **Architected** a universal s3gateway solution, **empowering user access to multiple clusters** without compromising performance. Devised an efficient fair call queue for s3gateway, **ensuring fairness for all client**s. Additionally, Applied Quality of Service mechanisms for s3gateway, **safeguarding impeccable service delivery while preventing any misuse or exploitation by client**s.

Data Engineer | Global Fashion Group

Jan 2019 – Dec 2020
Singapore, Singapore

ZMM (Markdown recommendation system) (Airflow, Python, SQL, Pandas/NumPy, Notebooks)

- Managed the process of revamping and refactoring the existing system into python and Jupyter notebook.
- Employed containerization, restructured, and fine-tuned ZMM, enabling efficient and auto-scalable deployment and cost reduction.
- **Resulted in an impressive 41% reduction in runtime and 70.6% cost savings.**

ZMA (Media attribution model) (Python, R, SQL, Pandas/NumPy)

- Pioneered a highly effective media attribution model based on the **Hidden Markov model, improving marketing decision-making by 20%.**
- **Collaborated with management to visualize model outcomes using Apache Superset for informed decision-making.**

CPA (Automated prices prediction) (PySpark, Python, pandas, Catboost model)

- Led the development, enhancement, and optimization of an automated price prediction system.
- Designed a novel data pipeline, leveraging Pyspark SQL interface and built-in ML features **improving training time by 25% over the existing pipeline .**
- Devised and implemented highly efficient data ingestion techniques, data bridges, and optimizing cluster size and operational costs.

Full Stack Developer | SigmaProIT

Dec 2017 – Sep 2018
Cairo, Egypt

Experto (App to App VOIP connects clients with experts) (Python, Django, NodeJS, Angular/Ionic)

- Developed a VOIP android application with multiple authentication methods along with statistical views.
- Dedicated webservice to keep track of voice calls lifecycle.

Ziruna (Expert booking app) (Python, Django, Angular/Ionic)

- Participated in mobile application development and **optimized the restful backend API response time by 20%.**

Software Engineer | Cloudypedia

Apr 2016 – Nov 2017
Cairo, Egypt

(Python, Django, GoogleAppEngine, ElasticSearch, OpenEDX)

- Constructed a certification module to support Arabic language for enhanced accessibility.
- Customized the registration module to integrate specific OAuth providers and domains.
- **Built filter and tagging module to optimize search capabilities on OpenEDX search module, reducing search time by 30%. Designed a googleAppEngine based web application slashing email migration time by 40%.**
- Crafted python wrappers for G-Suite APIS to enhance application integration.

Education

Cairo, Egypt
Ain Shams University
2011 – 2015
B.Sc. in Computer Science and information systems (Scientific computing department).

Languages and Technologies

- Python; Rust, Java; Scala; Golang; JavaScript; C++;
- Apache Spark; Apache Hadoop; Apache Ozone; RocksDB; Apache Kafka;
- SQL; PostgreSQL; Django; Flask; Webapp2; Angular; ReactJS; Ionic;
- AWS services; GoogleAppEngine; Heroku; Docker; Firebase; Apache Hadoop;
- Apache Ozone; Apache Spark; Distributed systems; Pandas; NumPy