**SUMMARY**

* Over **11+ Years’** experience as **Azure Data Engineer/Azure Migration Engineer, Azure Data Factory (ADF) V2. Managing Azure SQL Database, Azure Data Platform services - Azure Data Lake (ADLS), Data Factory (ADF), Data Lake Analytics, Stream Analytics, Azure SQL DW, HDInsight/Databricks, NoSQL DB), SQL Server, Oracle, Data Warehouse etc.**
* Expertise in Microsoft SQL Server, Azure PaaS components (Azure Data Factories, Data Bricks, Azure Logic Apps, Azure Data lake Analytics using USQL, Azure App Services, Geo-Replication, Azure Application insights and SQL Data Warehouse).
* **Strong development skills with Azure Data Lake, Data Bricks, Azure Data Factory, SQL Data Warehouse, Azure Blob, Azure Storage Explorer.**
* Experience managing **Big Data platform deployed in Azure Cloud**.
* Experience **in Azure Data Lake Gen1/Gen2**. Build how the data will be received, validated, transformed and then published.
* Experience in implementing hybrid connectivity between Azure and on - premise using virtual networks, VPN and Express Route.
* **Experience managing Azure Data Lakes (ADLS) and Data Lake Analytics and an understanding of how to integrate with other Azure Services.**
* Experience in troubleshooting and resolving architecture problems including database and storage, network, security and applications.
* Design and implement database solutions in Azure SQL Data Warehouse, Azure SQL.
* Implement ad-hoc analysis solutions using Azure Data Lake Analytics/Store, HDInsight.
* Design & implement migration strategies for traditional systems on Azure (Lift and shift/Azure Migrate, other third-party tools).
* Design and implement end-to-end data solutions (storage, integration, processing, visualization) in Azure.
* Experience managing Azure Data Lakes (ADLS) and Data Lake Analytics and an understanding of how to integrate with other Azure Services. Knowledge of USQL and how it can be used for data transformation as part of a cloud data integration strategy.
* Establish database standards for operations, upgrades, migrations and onboarding new applications and/or customers.
* Performed migration of Microsoft SQL server to Azure SQL database.
* Created Azure data factory Pipelines for moving data from databases and calling stored procedures using stored procedure activities, task is to extract the client product data as per the requirement of the client from cosmos clusters using Cosmos Scope scripts(USQL) and dump the data into Azure data lake and then cook the data through USQL Scripts and dump the data again to Azure Data Lake and then to Data Ware house using data factories.
* Develop workload migration plan in conjunction with other technical teams.
* Design and implement streaming solutions using Kafka or Azure Stream Analytics.
* Code and automation create to New System for scaling deployment mission critical operation of Cloud.
* **Azure Analysis Services (SSAS) and SQL Server Reporting Services (SSRS) and Familiarity with Excel PowerPivot, PowerView and PowerBI.**
* Creates and maintains Disaster Recovery procedures and participates in periodic disaster recovery testing
* Deploying Azure Resource Manager JSON Templates from PowerShell.
* Created Azure SQL database and performed monitoring and restoring of Azure SQL database. Performed migration of Microsoft SQL server to Azure SQL database.
* Advise on cost efficiency for future usage and cost optimization for current infrastructure.

**TECHNICAL SKILLS**

**Operating Systems:**RHEL/CentOS, Ubuntu/Debian/Fedora, Windows Server 2016/2012 R2

**Languages:**C# .NET, ASP.NET MVC, ASP.NET Web API

**CI Tools:**Visual Studio 2015, SSMS, Adobe Analytics, Asimov DataGrid, Asimov X-flow, PowerShell, Power BI, Microsoft Excel, Adobe Analytics

**Databases:**SQL Server, MySQL

**Scripts:**Shell scripting and PowerShell script

**SDLC:**Agile, Scrum, Waterfall Methodologies

**Monitoring Tools:**Azure Application Insights, Cloud Watch

**Microsoft Technologies:**ASP.NET 2.0/3.5/4.0, WCF 3.5 MVC, REST API’s, Microsoft Azure Components like Azure Data Factories, Logic Apps, Azure Function Apps, Azure Data lake, Azure Data lake Analytics, Azure App Services, Azure SQL Databases, SQL Data Warehouse, Azure Management Tools, Azure Bot Framework, Azure Databricks.

**PROFESSIONAL EXPERIENCE**

**Client :Dun & Bradstreet, Florham Park, NJ Oct 2024 – Till date**

**Sr Data Engineer**

**Responsibilities:**

* Designed and developed Security Framework to provide fine grained access to objects in **AWS S3** using **AWS Lambda, DynamoDB.**
* Set up and worked on Kerberos authentication principals to establish secure network communication on cluster and testing of **HDFS, Hive, Pig** and **MapReduce** to access cluster for new users.
* Performed end- to-end Architecture & implementation assessment of various **AWS services** like **Amazon EMR, Redshift, S3.**
* Implemented the machine learning algorithms using python to predict the quantity a user might want to order for a specific item so we can automatically suggest using kinesis firehose and **S3 data lake**.
* **Used AWS EMR to transform and move large amounts of data into and out of other AWS data stores and databases, such as Amazon Simple Storage Service (Amazon S3) and Amazon DynamoDB.**
* Created Pipelines in **ADF** using Linked Services/Datasets/Pipeline/ to Extract, Transform, and load data from Azure SQL, Blob storage, and **Azure SQL** Data warehouse.
* Designed and implemented end-to-end data pipelines using Azure Data Factory to facilitate efficient data ingestion, transformation, and loading (ETL) from diverse data sources into Snowflake data warehouse.
* Orchestrated robust data processing workflows utilizing **Azure Databricks** and **Apache Spark** for seamless large-scale data transformations and advanced analytics improving data processing speed by 14%.
* Developed real-time data streaming capabilities into **Snowflake** by seamlessly integrating **Azure Event Hubs** and **Azure Functions**, enabling prompt and reliable data ingestion.
* Created pipelines to run the notebooks using **ADF** on a scheduled basis.
* Stored streaming data in, Azure **CLI**, and **Azure Data Lake** Storage for decoupling ingestion from processing.
* Employed **Azure Blob Storage** for optimized data file storage and retrieval, implementing advanced techniques like compression and encryption to bolster data security and streamline storage costs.
* Designed and implemented data processing **workflows** using **Azure Databricks**, leveraging **Spark** for large-scale data transformations.
* Worked on Spreadsheet dataset to create a data load template for Workday integration.
* Demonstrated expert-level technical capabilities in areas of Azure Batch and Interactive solutions, and operational zing Azure Data Lakes (ADLS), Azure IoT, Copilot end-to-end, **MS** **Fabric**, and Azure Cloud Analytics solutions.
* Experience in developing CI/CD (continuous integration and continuous deployment) and automation using Jenkins, Git, docker, **Kubernetes** for ML model’s deployment.
* Integrated **ADF** with email Azure logic apps for sending email notifications.
* Extract Transform and Load data from Sources Systems to Azure Data Storage services using a combination of **Azure** **Data Factory, data bricks, PySpark, Spark SQL** and **U-SQL**, **Azure CLI**, **Azure Data Lake Analytics**.
* Understanding of structured data sets, data pipelines, ETL tools, data reduction, transformation and aggregation technique, Knowledge of tools such as **DBT**, DataStage
* Created and managed different types of tables in Snowflake, such as transient, temporary, and persistent tables.
* Experience as Azure Data Engineer in Azure Cloud, Azure Data Factory, Azure Data Lake Storage, Azure Synapse Analytics, Azure Analytical services, Redis, Azure Cosmos, Copilot, Databricks, **MS** **Fabric**, **NO** **SQL** DB, Big Data Technologies (**Hadoop** and **Apache** **Spark**), and Data bricks, Azure IoT, migration.
* Develop ELT data pipeline to migrate applications using **DBT** and Snowflake framework.
* Built scalable and optimized Snowflake schemas, tables, and views to support complex analytics queries and reporting requirements.
* Designed and implemented data pipelines for ingesting, transforming, and loading large datasets into BigID. (Highlight specific tools used like Apache Spark or Airflow)
* Developed data ingestion pipelines using **Azure CLI**, **Azure Event Hubs,** and **Azure Functions** to enable real-time data streaming into **Snowflake**.
* Utilized **AWS SQS** for decouple data ingestion from processing for scalability and reliability.
* Worked with **Azure BLOB** and **Data Lake storage** and loading data into **Azure SQL** Synapse analytics (DW).
* Developed real-time data streaming capabilities into Snowflake, enabling prompt and reliable data ingestion.
* Integrated Snowflake seamlessly with Power BI and Azure Analysis Services to deliver interactive dashboards and reports, empowering business users with self-service analytics capabilities.
* Designed and implemented a real-time data streaming solution using **Azure EventHub**.
* Leveraged Snowflake's Time Travel feature, ensuring optimal data management and regulatory compliance.
* Proficient in Snowflake integration, integrated with different data connectors, **Rest API’s** and **Spark**.
* Conducted performance tuning and optimization activities to ensure optimal performance of Azure Logic Apps
* Developed a Spark Streaming application to process real-time data from various sources such as **Kafka**, and **Azure Event Hubs.**
* Demonstrated hands-on experience in **Azure Cloud Services**, including **Azure Synapse** Analytics, **SQL Azure**, Data Factory, Azure Analysis Services, Application Insights, **Azure Monitoring, Key Vault**, and **Azure Data Lake**.
* Created batch and streaming pipelines in **Azure Data Factory (**ADF) using **Linked Services, Datasets,** and **Pipelines** to efficiently extract, transform, and **load data**.
* **Built a Spark** Streaming application for real-time analytics on streaming data, leveraging **Spark SQL** to query and **aggregate data** in real-time and visualize the results in **Power BI** or Azure Data Studio.
* Developed **JSON** Scripts for deploying the Pipeline in Azure Data Factory (ADF) that processes the data using the SQL Activity. Performed data flow transformation using the data flow activity. **Implemented Azure, self-hosted integration** runtime in ADF. Developed streaming pipelines using with Python.
* Deployed and developed various data pipelines extensively and intricately harnessing the power of C# .NET, MS Fabric, Azure Data Factory.
* Working as Kubernetes Administrator, involved in configuration for web apps, Azure App services, Azure Application insights, Azure Application gateway, Azure DNS, Azure traffic manager, App services.
* Transformed and copied data from JSON files stored in **Data Lake** Storage into Azure Synapse Analytics tables using **Azure Databricks,** ensuring accurate and efficient data migration.
* Leveraged Azure Databricks, Azure Storage Account, and other relevant technologies for source stream extraction, cleansing, consumption, and publishing across multiple user bases.
* Knowledge of **PDI's** features for data cleansing, validation, and enrichment.
* Familiarity with scheduling and monitoring ETL jobs in Pentaho Data Integration.
* Build streaming **ETL pipelines** using **Spark Streaming** to extract data from various sources, transform it in real-time, and load it into a **data warehouse** such as **Azure Synapse** Analytics.
* Used tools such as, Azure CLI, Azure Databricks or HDInsight to scale out the Spark Streaming cluster as needed.
* Developed **Spark API** to import data into **HDFS** from Teradata and created **Hive tables**.
* Familiarity with data security and privacy practices, especially relevant to **BigID** focus on data governance.
* Created Partitioned and **Bucketed Hive** tables in Parquet File Formats with Snappy compression.
* Involved in running all **the Hive scripts** through Hive on Spark and some through **Spark SQL**.
* Using the **JSON** and **XML** SerDe's for serialization and de-serialization to load JSON and XML data into **HIVE.**
* Involved in complete **Big Data flow** of the application starting from data ingestion from upstream to HDFS, processing and analyzing the data in HDFS.
* Used **DBT** to debug complex chains of queries. They can be split into multiple models and macros that can be tested separately.
* Developed Spark core and Spark SQL scripts using Scala for faster data processing.
* Performed **data profiling** and transformation on the raw data using **Python**.
* Mentor and guide analyst on building purposeful analytics tables in **dbt** for cleaner schemas.
* Implement **RDD**/**Datasets**/Data frame transformations in Scala through **Spark Context** and **Hive Context**
* Used **Jira** for bug tracking and **Bitbucket** to check in and check out code changes.
* Designed and implemented data archiving and data retention strategies using, **Azure CLI, Azure Blob Storage** and **Snowflake's** Time Travel feature.

**Environment:** Azure Data Factory, Azure Databricks, Snowflake data warehouse, Azure Event Hubs, Azure Functions, Azure Logic Apps, Azure Data Lake Storage, Azure Blob Storage, DBT (data build tool), Power BI, Kubernetes ARM Templates, Terraform, Apache Purview, Apache Atlas, AWS Glue, AWS Sage Maker.

**Client: PG& E, San Francisco, CA Nov 2023 – Sep 2024**

**Sr Azure Data Engineer**

**Responsibilities:**

* Design and implement end-to-end data solutions (storage, integration, processing, visualization) in Azure
* Architect and implement ETL and data movement solutions using Azure Data Factory, SSIS
* Develop dashboards and visualizations to help business users analyze data as well as providing data insight to upper management with a focus on Microsoft products like SQL Server Reporting Services (SSRS) and Power BI.
* Engage with business users to gather requirements, design visualizations and provide training to use self-service BI tools.
* Developed and maintained multiple Power BI dashboards/reports and content packs
* Created POWER BI Visualizations and Dashboards as per the requirements
* Used various sources to pull data into Power BI such as SQL Server, Excel, Oracle, SQL Azure etc.
* Design and implement database solutions in Azure SQL Data Warehouse, Azure SQL
* Migrate data from traditional database systems to Azure SQL databases
* Implement ad-hoc analysis solutions using Azure Data Lake Analytics/Store, HDInsight
* Propose architectures that consider cost/spend in Azure and develop recommendations to right-size Azure data infrastructure.
* Work with similar Microsoft on-prem data platforms, specifically SQL Server and related technologies such as SSIS, SSRS, and SSAS
* Design & implement data security and governance policies and standards
* Migrate data from traditional database systems to Azure databases
* Recreating existing application logic and functionality in the Azure Data Lake, Data Factory, SQL Database and SQL datawarehouse environment. Experience in DWH/BI project implementation using Azure DF
* Interacts with Business Analysts, Users, and SMEs on requirements
* Designs Logical and Physical Data Model for Staging, DWH and Data Mart layer.
* Engage with business users to gather requirements, design visualizations and provide training to use self-service BI tools.
* Developed and maintained multiple Power BI dashboards/reports and content packs.
* Created POWER BI Visualizations and Dashboards as per the requirements
* Used various sources to pull data into Power BI such as SQL Server, Excel, Oracle, SQL Azure etc.
* Propose architectures considering cost/spend in Azure and develop recommendations to right-size data infrastructure
* Experience managing Azure Data Lakes (ADLS) and Data Lake Analytics and an understanding of how to integrate with other Azure Services. Knowledge of USQL and how it can be used for data transformation as part of a cloud data integration strategy
* Participating in Technical Architecture Documents, Project Design and Implementation Discussions
* Azure Automation through Runbook Creation, Migration of existing .PS1 scripts, Authorizing, Configuring, Scheduling
* Dev Ops tools, processes and best practices including automation of resource provisioning/maintenance
* Compute and Storage area in multi-platform environment such as Linux and Windows

**Environment:** Azure SQL, Azure Data Factory, Azure Data Lakes, Power BI, Azure Storage Explorer, Azure Blob, Shell Script, SSIS, SSRS, Windows.

**Client : Nationwide Insurance, Columbus, OH Jul 2021 – Oct 2023**

**Azure Data Engineer**

**Responsibilities:**

* Created and maintained continuous integration (CI) using tools Azure DevOps over different environments to facilitate an agile development process which is automated and repeatable, enabling teams to safely deploy code many times a day while ensuring operational best practices are supported.
* Design and document CI/CD tools configuration management.
* Responsible for orchestrating CI/CD processes by responding to Git triggers, human input, and dependency chains and environment setup.
* Build and Deployment CI/CD Pipelines.
* Managed User access control, Triggers, workflows, hooks, security and repository control in Bitbucket.
* Responsible for new and existing data onboarding from MAG to Splunk and troubleshooting every Splunk issue and optimizing performance.
* Configured Microsoft add on for Splunk to send data from Azure event hubs to Splunk.
* Normalizing Splunk data to ensure all fields are mapping with CIM and bringing data to Splunk ES and ITSI.
* Working closely with Business team and helping to solve Splunk problems within the business Environment.
* Provided regular support guidance to Splunk project teams on complex issue resolution and assisting with best practices, creating deployment runbooks to empower the development teams.
* Performed Configuration of Azure Web Apps, App services, Application insights,AzureDNS,AzureNetwork Watcher and Implementing Azuresite recovery,Azurestack, Backup and Automation.
* Experience in configuring ARM, VM Networking, Private Static IP Address and Public IP address inAzure.
* Created Virtual Networks inAzurefor hybrid connectivity with On-premise networks (Express Route, VPN Gateways).
* Worked on MicrosoftAzurefor Development, Test and UAT environments for all customers regardless of final Production platform to speed up.
* DevelopedAzure PaaS services including WebJobs, AzureSQL, Azure BLOB storage, Azure Application Insights, Azure Data Factory, Azure functions, Azure BOT Service, Azure Active Directory, Azure Stream Analytics, Azure CDN.
* Implemented Infrastructure automation through Terraform and Azure DevOps for auto provisioning, code deployments, software installation.
* Used Terraform Enterprise version and integrated that with VCS repository hosted on Git to automate deployment of infrastructure of check in for Dev, Test and UAT environments.
* Designed and created multiple deployment strategies using CI/CD Pipeline and configuration management tools, with remote execution to ensure zero downtime and shortened automated deployments.
* Deployed applications using Azure DevOps and troubleshooted build and release job failures.
* Build, deploy ASP.NET web application on using TFS Build Automation, Blob Storage and Portal.
* Experience in using Visual Studio .NET platform (C#, ASP.NET, ADO.Net), thorough knowledge of the architecture and the framework.
* Worked with scrum and development team in overseeing bug tracking, test prioritization, test automation and releases.
* Improved team performance and accelerated release cycle accelerate delivery of new features and increase overall product qualities.

**Environment:** Git, Apache Tomcat, Agile/Scrum, SDLC, Windows, Linux, Azure, PowerShell, Splunk, Shell script, Power BI

**Client : Abilo IT Solutions, India June 2017 – Aug 2020**

**Azure Data Engineer**

**Responsibilities:**

* Hands on Experience working with the MicrosoftAzureCloud Services and Deployed the servers throughAzure Resource Manager Templates orAzurePortal.
* Planning and implementing Disaster Recovery solutions, capacity planning, data archiving, backup/recovery strategies, Performance Analysis and optimization.
* Strong Experience on Administration and Troubleshooting ofAzureIAAS Components (VM, Storage, VNET, OMS, NSG, Site to Site VPN, RBAC, Load Balancers, Availability Sets).
* Experience in setting up cluster environment for load balancing and high availability of Windows and Linux servers.
* Experience in Creating the CI/CDAzure DevOpsPipeline for the .Net, java & UI based Web Applications.
* Good Exposure inAzureSite Recovery andAzureBackup and ConfiguringAzureBackup vault for protecting required VMs to take the VM level backups forAzureand On Premises Environment.
* Working knowledge onAzureCloud IaaS and PaaS Services. Having Experience of Creating and Managing the users and groups inAzureAD.
* Experience in managing Microsoft Windows server infrastructure and data - center operations by effectively planning, installing, configuring and optimize the IT infrastructure to achieve high availability and performance.
* Designed Network Security Groups (NSGs) to control inbound and outbound access to network interfaces (NICs), VMs and subnets.
* Worked on the PowerShell Script to automate the Windows patching and created the release in Azure DevOps pipelines.
* 24/7 monitoring theAzureResources usingAzureMonitor and Web Apps for Application Insights.

**Environment:** Azure, Azure DevOps, .Net, PowerShell, SQL Server 2012,2008, SQL 2005,2000, Windows Server 2012, Windows 2008 R2, /2000 Server (64bit)

**Client : : Cyient Systems, India Dec 2015 – May 2017**

**Azure Data Engineer**

**Responsibilities:**

* Analyze, design and build Modern data solutions using Azure PaaS service to support visualization of data. Understand current Production state of application and determine the impact of new implementation on existing business processes.
* Extract Transform and Load data from Sources Systems to Azure Data Storage services using a combination of Azure Data Factory, T-SQL, Spark SQL and U-SQL Azure Data Lake Analytics. Data Ingestion to one or more Azure Services - (Azure Data Lake, Azure Storage, Azure SQL, Azure DW) and processing the data in In Azure Databricks.
* Implemented Proof of concepts for SOAP & REST APIs
* REST APIs to retrieve analytics data from different data feeds
* Created Pipelines in ADF using Linked Services/Datasets/Pipeline/ to Extract, Transform and load data from different sources like Azure SQL, Blob storage, Azure SQL Data warehouse, write-back tool and backwards.
* Developed Spark applications using **Pyspark** and Spark-SQL for data extraction, transformation and aggregation from multiple file formats for analyzing & transforming the data to uncover insights into the customer usage patterns.
* Developed **PySpark** based pipelines using spark data frame operations to load data to EDL using EMR for jobs execution & AWS S3 as storage layer.
* To meet specific business requirements wrote UDF’s in Scala and **PySpark**.
* Creating pipelines, data flows and complex data transformations and manipulations using ADF and **PySpark with Databricks**.
* Developed mapping document to map columns from source to target.
* Created several Databricks Spark jobs with **PySpark** to perform several tables to table operations.
* Responsible for estimating the cluster size, monitoring and troubleshooting of the Spark data bricks cluster.
* Experienced in performance tuning of Spark Applications for setting right Batch Interval time, correct level of Parallelism and memory tuning.
* Developed JSON Scripts for deploying the Pipeline in Azure Data Factory (ADF) that process the data using the Sql Activity.
* Hands-on experience on developing SQL Scripts for automation purpose.
* Created Build and Release for multiple projects (modules) in production environment using Visual Studio Team Services (VSTS).

**Client – Qualcomm Inc,India Mar 2013 – Nov 2015**

**Data Engineer**

**Responsibilities:**

* Design and implement database solutions in Azure SQL Data Warehouse, Azure SQL.
* Architect & implement medium to large scale BI solutions on Azure using Azure Data Platform services (Azure Data Lake, Data Factory, Data Lake Analytics, Stream Analytics, Azure SQL DW, HDInsight/Databricks, NoSQL DB).
* Design & implement migration strategies for traditional systems on Azure (Lift and shift/Azure Migrate, other third-party tools.
* Engage with business users to gather requirements, design visualizations and provide training to use self-service BI tools.
* Used various sources to pull data into Power BI such as SQL Server, Excel, Oracle, SQL Azure etc.
* Propose architectures considering cost/spend in Azure and develop recommendations to right-size data infrastructure.
* Develop conceptual solutions & create proof-of-concepts to demonstrate viability of solutions.
* Technically guide projects through to completion within target timeframes.
* Collaborate with application architects and DevOps.
* Identify and implement best practices, tools and standards.
* Design Setup maintain Administrator the Azure SQL Database, Azure Analysis Service, Azure SQL Data warehouse, Azure Data Factory, Azure SQL Data warehouse.
* Build Complex distributed systems involving huge amount data handling, collecting metrics building data pipeline, and Analytics.