**Varun Bachawal**

***Azure Data Engineer***

**| Varun.bach@gmail.com | 313-651-4603 |**

**PROFESSIONAL SUMMARY:**

* Around **10+ years** of professional experience in developing and maintaining web and client/server applications using Microsoft Technologies.
* Hands-on experience in **Azure Cloud Services (PaaS & IaaS), Storage, Web Apps, Active Directory, USQLS, Application Insights,** and **Logic Apps.**
* Experience of working with Azure Monitoring, Data Factory, Traffic Manager, Service Bus, Key Vault.
* Experienced in **Azure Data Factory** and preparing **CI/CD scripts** for the deployment and development experience on **cloud platforms (Azure).**
* **Solid experience** on **building ETL ingestion flows** using **Azure Data Factory.**
* Experience in building **Azure stream Analytics ingestion** spec for data ingestion which helps users to get sub second results in Realtime.
* Experience in building **ETL (Azure Data Bricks) data pipelines leveraging PySpark and Spark SQL.**
* Experience in building the **Orchestration on Azure Data Factory** for scheduling purposes.
* Hands-on experience in **Azure Analytics Services – Azure Data Lake Store (ADLS), Azure Data Lake Analytics (ADLA), Azure SQL DW, Azure Data Factory (ADF), Azure Data Bricks (ADB) etc**.
* Experience working with **Azure Logic APP Integration tool.**
* **Experience in implementation of Azure log analytics** providing Platform as a service for **SD-WAN firewall logs**.
* **Experience in building the data pipeline** by leveraging the **Azure Data Factory.**
* Expertise on working with databases like **Azure SQL DB, Azure and SQL DW.**
* **Orchestrated data integration pipelines** in ADF using various Activities like Get **Metadata, Lookup, For Each, Wait, Execute Pipeline, Set Variable, Filter, until, etc.**
* Have knowledge and experience on **Basic Admin activities** related to **ADF like providing access to ADLs** using **service principle, install IR, created services like ADLS, logic apps etc.**
* Good experience on **polybase external tables in SQL DW.**
* Experience on major components in **Hadoop Ecosystem like Hadoop Map Reduce, HDFS, HIVE, PIG, Hbase, Zookeeper, Oozie and Flume.**
* Expertise in setting up processes for **Hadoop based application design and implementation.**
* Experience in **importing and exporting data using Sqoop from HDFS to Relational Database** and vice-versa.
* **Experience in managing and reviewing Hadoop log files.**
* Experienced in processing **big data on the Apache Hadoop framework using MapReduce programs.**
* Excellent **understanding and knowledge of NOSQL databases like HBase and Mongo DB.**
* Profound understanding of Partitions and Bucketing concepts in Hive and designed both Managed and External tables in **Hive to optimize performance**.
* **Proficiency in SQL across several dialects (MySQL, PostgreSQL, Redshift, SQL Server, and Oracle).**
* Extensively worked with **Teradata utilities Fast export, and Multi Load to export and load data to/from different source systems including flat files.**

**TECHNICAL SKILLS:**

**Cloud** : Azure (Data Factory, Data Lake, Data Bricks, Logic App, ARM, Azure SQL)

**ETL Tools** : Azure Data Factory(V2), Azure Data Bricks, SSIS, SSAS, IBM IIS DataStage

**Big Data**  : Apache NiFi, PySpark, Apache Kafka

**Reporting Tools** : Power BI, SAP Business Objects 4.1, & QlikView, Tableau Desktop

**Automation Tools** : Azure Logic App, Datastage Director, Crontab, Control-M, UiPath

**Code Repository Tools** : GitLab, SVN – Tortoise, TFS, StarTeam

**Database** : Oracle (10g, 11g, 12c), SQL Server (2008 R2), Db2, Netezza, MySQL, PostgreSQL

**Operating System** : Windows, Sun-Solaris, Red Hat Linux, Ubuntu

**Other tools** : Visual Studio, MS Office, WinSCP and Putty

**WORK EXPERIENCE:**

**McDonalds | Chicago, IL Jun 2021 to Present**

**Cloud/Azure Data Engineer**

**Responsibilities:**

* Involved in SDLC Requirements gathering, Analysis, Design, Development and Testing of application using Agile Methodology.
* Responsible for the execution of big data analytics, predictive analytics, and machine learning initiatives.
* Created **Linked Services for multiple source system** (i.e.: **Azure SQL Server, ADLS, BLOB, Rest API**).
* **Created Pipeline’s to extract data from on premises source systems** to azure cloud **data lake storage**, extensively worked on copy activities and implemented the copy behavior’s such as **flatten hierarchy, preserve hierarchy and Merge hierarchy**, Implemented **Error Handling concept** through copy activity.
* Exposure on Azure Data Factory activities such as **Lookups, Stored procedures, if condition, for each, Set Variable, Append Variable, Get Metadata, Filter and wait.**
* Configured the logic apps to handle email notification to the end users and key shareholders with the help of **web services activity.**
* **Created dynamic pipeline** to handle **multiple sources extracting to multiple targets** and extensively used **azure key vaults** to configure the **connections in linked services.**
* Configured and implemented the **Azure Data Factory Triggers** and **scheduled the Pipelines** and monitored the scheduled **Azure Data Factory pipelines** and **configured the alerts to get notification of failure pipelines.**
* Extensively worked on **Azure Data Lake Analytics** with the help of **Azure Data bricks** to implement **SCD-1, SCD-2 approaches.**
* **Created Azure Stream AnalyticsJobs** to replication the **real time data to load to Azure SQL Data warehouse.**
* Implemented **delta logic extractions for various sources** with the help of control table; implemented the **Data Frameworks to handle the deadlocks, recovery, logging the data of pipelines.**
* Understand the latest features like (**Azure DevOps, OMS, NSG Rules,** etc..,) introduced by Microsoft Azure and utilized it for existing business applications.
* Worked on migration of data from **On-prem SQL server to Cloud databases** (**Azure Synapse Analytics (DW) & Azure SQL DB).**
* Deployed the codes to multiple environments with the help **of CI/CD process and worked on code defect** during the SIT and UAT testing and provide supports to data loads for testing; Implemented reusable components to reduce manual interventions.
* **Developing Spark (Scala) notebooks** to **transform** and **partition the data and organize files in ADLS.**
* Working on **Azure Data bricks** to run **Spark-Python Notebooks through ADF pipelines.**
* Using **Data bricks utilities called widgets** to pass parameters on run time from **ADF to Data bricks.**
* **Created Triggers, PowerShell scripts and the parameter JSON files** for the deployments.
* Worked with **VSTS for the CI/CD Implementation.**
* Reviewing individual work on **ingesting data into azure data lake** and provide feedbacks based on **reference architecture, naming conventions, guidelines, and best practices.**
* Implemented **End-End logging frameworks for Data factory pipelines.**

**Environment:** Azure Data Factory, Azure Data Bricks, PolyBase, Azure DW, ADLS, Azure DevOps, BLOB, Azure SQL Server, Azure DW, Azure Devops, Azure synapse.

**Sun Trust Bank | Atlanta, GA Oct 2019 to May 2021**

**Azure Data Engineer**

**Responsibilities:**

* Created Pipelines in **ADF** using **Linked Services, Datasets and Pipeline** to **Extract, Transform, and load** data from different sources like **Azure SQL, Blob storage, Azure SQL Datawarehouse, write-back tool, and backward.**
* Extracted, Transformed and Loaded data from Sources Systems to **Azure Data Storage services** using a combination of **Azure Data Factory, Spark SQL, and U-SQL Azure Data Lake Analytics.**
* Data is Ingested to one or more **Azure Services - (Azure Data Lake, Azure Storage, Azure SQL, Azure DW) and processing the data in In Azure Databricks**. Worked on **Azure Services like IaaS, PaaS** and worked on storage like Blob (Page and Block), **SQL Azure.**
* Implemented OLAP multi-dimensional functionality using **Azure SQL Data Warehouse**. Retrieved data using **Azure SQL and Azure ML** which is used to build, test, and predict the data.
* Worked on Cloud databases such as **Azure SQL Database, SQL managed instance, SQL Elastic pool on Azure, and SQL server.**
* 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**).
* Responsible for estimating the **cluster size, monitoring, and troubleshooting of the Spark data bricks cluster.**
* Designed and developed Azure Data Factory pipelines to Extract, Load and Transform data from difference sources systems (Mainframe, SQL Server, IBM DB2, Shared Drives, etc.) to **Azure Data Storage services** using a combination of **Azure Data Factory, Azure Databricks (PySpark, Spark-SQL), Azure Stream Analytics and U-SQL Azure Data Lake Analytics**. **Data Ingestion into various Azure Storage Services like Azure Data Lake, Azure Blob Storage, Azure Synapse Analytics (formerly known as Azure Data Warehouse).**
* Configured and deployed Azure Automation Scripts for a multitude of applications utilizing the **Azure stack (including Compute, Web & Mobile, Blobs, ADF, Resource Groups, Azure Data Lake, HDInsight Clusters, Azure Data Factory, Azure SQL, Cloud Services, and ARM), Services and Utilities focusing on Automation.**
* Involved in Migrating Objects from Teradata to Snowflake and created Snow pipe for continuous data load.
* Increased consumption of solutions including **Azure SQL Databases, Azure Cosmos DB** and **Azure SQL.**
* Created **continuous integration and continuous delivery (CI/CD) pipeline** on Azure that helps to automate steps in the software delivery process.
* Deploying and managing applications in **Datacenter, Virtual environment, and Azure** platform as well.
* Involved in converting Hive/SQL queries into **Spark transformations using Spark RDD's and PySpark.**
* **Log data Stored in HBase DB** is processed and analyzed and then imported into Hive warehouse, which enabled end business analysts to write HQL queries.
* **Handled importing of data from various data sources**, performed transformations **using Hive, and loaded data into HDFS.**
* Design, development, and implementation of performant **ETL pipelines using PySpark and Azure Data Factory.**

**Environment**: Azure Data Factory(V2), Azure Data Bricks (PySpark, Spark SQL), Azure Data Lake, Azure BLOB Storage, Azure ML, Azure SQL, Hive, Git, JIRA, HQL, Snowflake, Teradata.

**UHG | Minneapolis, MN Mar 2016 to Sep 2019**

**Azure Data Engineer**

**Responsibilities:**

* **Designed and Configured Azure Cloud relational servers** and databases analyzing current and future business requirements.
* Worked on migration of data from On-prem SQL server to Cloud databases**(Azure Synapse Analytics (DW) & Azure SQL DB).**
* Have good experience in setting up separate application and reporting data tiers across servers using Geo replication functionality.
* Implemented Disaster Recovery and Failover servers in Cloud by replicating data across regions.
* Have extensive experience in creating pipeline jobs, scheduling triggers, Mapping data flows **using Azure Data Factory(V2) and using Key Vaults to store credentials.**
* Good experience in **creating Elastic pool databases** and schedule Elastic jobs for executing **TSQL procedures.**
* For Log analytics and for better query response used Kusto Explorer and created alerts using Kusto query language.
* Worked on creating tabular models on **Azure analysis services** for meeting business reporting requirements.
* Have good experience working with Azure BLOB and **Data Lake** storage and loading data into **Azure SQL Synapse analytics (DW).**
* Designed SSIS Packages using **Business Intelligence Development Studio (BIDS)** to extract data from various data sources and load into **SQL Server database** for further Data Analysis and Reporting by using multiple transformations.
* Worked on creating correlated and non-correlated sub-queries to resolve complex business queries involving multiple tables from different databases.
* Developed business intelligence solutions using SQL server data tools 2015 & 2017 versions and **load data to SQL** and**Azure Cloud databases.**
* Perform analyses on data quality and apply business rules in all layers of data extraction transformation and loading process.
* Perform validation and verify software at all testing phases which includes Functional Testing, System Integration Testing, End to End Testing, Regression Testing, Sanity Testing, User Acceptance Testing, Smoke Testing, Disaster Recovery Testing, Production Acceptance Testing and Pre-prod Testing phases.
* Have good experience in logging defects in Jira and **Azure Devops tools.**
* Involved in planning cutover strategy, go-live schedule including the scheduled release dates of Portfolio central Datamart changes.

**Environment**: Microsoft SQL Server, SSDT-2012 &2015, Azure Synapse Analytics, Azure Data Lake & BLOB, Azure SQL, Azure data factory, Azure analysis services.

**Spinix Solutions | Hyderabad, IN Apr 2012 to Feb 2016**

**Cloud/Azure Data Engineer**

**Responsibilities:**

* Worked on new Data warehouse design for ETL and **Reporting projects using SSIS and SSRS.**
* **Created various kinds of reports using Power BI and Tableau** based on the client's needs.
* **Implemented indexes** such as **clustered index, non-clustered index, covering index appropriately on data structures to achieve faster data retrieval.**
* **Worked with Data Governance tools** and extract-transform-load (ETL) processing tool for **data mining, data warehousing, and data cleaning using SQL.**
* **Optimized SQL performance, integrity, and security of the project’s databases/schemas.**
* Performed ETL operations to **supportincremental, historical data loads and transformations** using **SSIS.**
* **Created SSIS packages to extract data from OLTP to OLAP systems** and scheduled jobs to call the packages.
* Implemented **Event Handlers and Error Handling in SSIS packages** and notified process results to various user communities.
* **Designed SSIS packages to import data** from multiple sources to control **upstream and downstream of data into SQL Azure database.**
* Used various advanced **SSIS** functionalities **like complex joins, conditional splitting, column conversions** for better performance during package execution.
* **Developed impactful reports** using **SSRS, MS Excel, Pivot tables and Tableau** to solve the business requirements.
* Creating and deploying parameterized reports using **power BI and Tableau.**
* Maintained the physical database by **monitoring performance, integrity and optimize SQL queries** for maximum efficiency using **SQL Profile**.
* Worked on **formatting SSRS reports using the Global variables and expressions.**
* Created Power BI Reports using the **Tabular SSAS models in Power BI desktop** and published them to Dashboards using cloud service.
* Created designs and process flow on to how standardize **Power BI Dashboards** to meet the requirements of the business. Made changes to the existing **Power BI Dashboard** on a regular basis as per requests from Business.
* Created data maintenance plan for backup and restore and update statistics and **rebuild indexes.**

**Environment**: SQL Server 2016, T-SQL, SQL Profiler, SSIS, SSRS, SSAS, TFS, SSAS, MS SQL Server, Oracle10g, Oracle WebLogic Server, Query Analyzer, Power BI, Power Pivot, Windows, MS Excel.