# Summary Qualifications Consultant Role with Prolific

* 19+ years of experience in IT, Healthcare, Telecommunication and Financial domain
* 10+ years in Data Analytics, including an end to end Solution Architecting, Data Design, and Integration architect
* 5+ years in Public Cloud Data Architecture, Cloud DataPlatform implementation, and Cloud Application Migration
* 4+ years in Azure/ Aws Modern Datawearhouse / cloud Artificial Intelligence /Machine Learning solutions development
* Strong hands-on in Databricks platform, Data ingestion, Data Extraction, Data Governance, Data modeling, Data security & Data Migration for **Transactional and OLAP systems**.
* Strong experience in Public cloud (Azure , AWS ) / Hybrid Cloud model & On-premises Data Analytics application development, various platform design and Implementation
* Microsoft Azure certified Solution architect expert, Azure certified Administrator, Cloudera certified Data Architect

Profile

A performance-driven accomplished Cloud-Data Architect professional with a wide range of experience in developing advanced data-driven Enterprise solutions, analytics solutions (Artificial Intelligence - Machine Learning), data products and data engineering capabilities from the ground up across various industry segments (e.g., Telecommunication, **Banking, Insurance**, Transportation and R&D). Well-versed in building an economically feasible and scalable modern data-driven platform using a wide range of cloud and on-premise technology stack and tools. Currently leading Enterprise modern DataPlatform architecting and migration, have also led various Cloud data application migration, End to end Data migration, Cloud Integration, Operation, and Data modeling.

Possess strong hands-on Architecture experience in Microsoft Azure cloud platform / Azure modern Data platform, Bigdata in AWS, Data/Application migration from On-prem to Cloud, Data ingestion, data storage, data analysis, data modeling and governance, Various Data security and integration experience in Azure, multi-cloud, **Databricks**, SPARK, NiFi data flow, Azure Data Factory, Azure SQL, Synapse, and Cloudera CDP.

**SNOWFLAKE:**

* Experience in migrating data into the Snowflake environment.
* Worked on data models to support analytics.
* Loading data from Azure to Snowflake.
* Copying parquet files from Azure environment to Snowflake.
* Worked in SnowSQL, SnowPipe

Technical Skills

|  |  |  |
| --- | --- | --- |
| **Microsoft Azure** | | Microsoft Azure,ADLS, ADF, Azure SQL, Synapse Analytics, Databricks, Azure AI service, Azure Data lake HDInsights, Event Hub , Azure DevOPS , **Informatica**, PowerBI, **Snowflake** ( SNOW SQL, Snowflake Pipe, Snowflake utility ), Servicenow |
| **AWS** | | AWS (storage, EMR, Kinesis, Data pipes, Ec2, ), Cloud break, Glue, S3, Athena, EMR, Redshift, DynamoDB |
| **BigData Hadoop** | | Cloudera (CDP), Ambari, HDFS, YARN, MapReduce, NiFI Realtime streaming, Stream set, Sqoop, Kafka, Spark Streaming, Mongo DB |
| **Programing Language & Other technology** | | SCALA, Python, NoSQL HQL, XML, shell scripting, SQL, LINUX, Webservice, JSON/XML, Ansible, Big SQL, Oracle/MySQL/PostgreSQL |
| **Machine learning & AI** | | Unsupervised learning (Clustering), supervised learning (Classification), K-means algorithm, NLP, Proactive technical analysis |
| **Architect:** | | Databricks platform, Cloud Data Architecture, Azure solution architecture, Big Data Application Architect, Modern Data Platform Architecture , Cloud Datawarehouse Architecture |
| **Industry standards :** | | eTOM, TOGAF 9.1, SID, PMP, Props’ C, ITIL v3, Agile Development methodology |
| **Security** | Kerberos, Ranger, Knox, Active directory LDAP integrations | |
| **Databases** | SQL server, Oracle, Postgres, MongoDB, MySQL | |

Education

* Master’s in Business Administrations
* Bachelor’s in Electronics and Communication Engineering, University degree

Certifications

* AWS certified cloud practitioner.
* Microsoft certified AZURE (Az-300 / Az-301) cloud Solution Architect Expert
* Microsoft Certified Azure (Az-103) Cloud Administrator Engineer
* Microsoft Certified Azure (Az-900) Cloud associate
* HDPCD- Hortonworks certified Hadoop Developer
* HDPCA-Hortonworks Certified Hadoop Administrator
* HCA- Hortonworks certified Hadoop associate

Work Experience

**Photon ( Walgreens Pharmacy – Chicago USA ) Aug 2022 – July 2023**

**Sr.Multi-Cloud ( Azure/Aws) Data Architect**

**Project 1:**

Working as a Modern Cloud Data Warehouse Enterprise architect, where I designed and delivered a multi-warehouse environment. Under this project, I performed the following task:

Designed and implemented Azure Datalake which handles multi-stream Data Ingestion, Data cataloging and Data sharing environments.

Designed and delivered Metadata driven Data ingestion framework.

Architected and designed DevOps integration with Data lake services.

Designed and developed a Data processing framework using Databricks.

* Designed end-to-end data processing architectures using Databricks, considering data sources, transformations, storage, and analytics requirements.
* Collaborated with stakeholders to define technical specifications and requirements for Databricks-based solutions.
* Created architectural diagrams and documentation to outline Databricks workflows, data pipelines, and integration points.
* Implemented network configurations, authentication mechanisms, and access controls for Databricks workspaces.
* Integrated Databricks with other Azure services and third-party tools, enabling seamless data flow and processing.
* Designed and developed ETL pipelines using Databricks notebooks, leveraging PySpark and Spark SQL for data transformations.
* Implemented role-based access controls (RBAC), fine-grained permissions, and data encryption to ensure data security and privacy.
* Implemented CI/CD practices for Databricks notebooks and pipelines, ensuring reliable and efficient deployment of code changes.

Maintained Onsite and Offshore resources for implementation.

**Project 2:**

Working as a Cloud migration Data Architect for Web & Mobile App migration projects. Where I am helping them to Data migration and Database migrations.As a part of this project, I performed the following task:

Worked as a Metadata architect and designed solutions for Data-driven catalog products.

Performed Data & Application Discovery

Extensive data management experience in a data warehouse and data lake environment

Performed Data modeling and Data profiling on MongoDB

Performed Dimension and Fact table modeling and integrated both Transactional and Operational data.

Designed and Architected just in time and Bulk migration data to external SAAS authentication service.

Architected and Designed AWS NoSQL DynamoDB migration from Mongo.

Strong Hands on experience in Enterprise Data lake Activityes ( Data Ingestion, Data processing and AI/ML )

**GWL INSURANCE Oct 2018 – Aug 2022**

**Sr. Azure Cloud Data Solution Architect (Azure cloud, AWS CDP, HDF)**

Working as a Sr. Azure cloud Data architect consultant for enterprise data lake, where I am helping them to transform their End to End Data strategy into an Azure cloud Data Analytics platform as a multi-phased approach. Also, I build the BigData reference Architecture patterns for various data ingestion, storage & computation. Also migrating theircurrent on-prem BigData platform & applications to Azure cloud platform :

* Architect and Delivered Microsoft Azure-based modern analytics platform High level & Low level Design documents, other Service integration solutions, various Data layer design, Data security, Alert and monitoring solutions .
* Participate in the review of project artifacts – User story, System Requirement Specification, System Architecture, System Design, etc.
* Develop and maintain Cloud Data Models for Enterprise Class Data Initiatives, Developing and enforcing data modelling standards and best practices
* Design and Implemented Data and Application Migration strategy from On-prem to Azure platform.
* Design and Implemented Batch & Event Data ingestion tools(ELT/ETL) like Azure Data factory, Synscort, Event Hub, Storage layers (Azure Data lake ADLS Gen2, Blob, Synapse managed storage), for data processing (Azure Data bricks, Synaose Analytics) , for Machine learning ( Azure ML).
* Design and Implemented Azure Data security using (Azure security, key Vault, ctive directory), Also closely worked with DevOPS using Azure DevOps service for CI/CD pipeline deployment.
* For another project, I designed and implemented Cloud-based (Azure) data lake using Microsoft HDInsight Platform-As-A-Service (PaaS) and migrated on-prem application to Azure HDInsight
* Design and Implemented the various levels of Data security like (Data in Motion SSL, at Rest Ranger KMS, different layers of access control), For Authentication perimeter using Knox and End to end Governance using Ranger.
* Lead the development of BigData reference architecture patterns for data ingestion, data processing, analytics and data science solutions in both On-prem and Azure Data lake.
* Design and Develop AWS Data lake for Data analytics.
* Ingest the Data to AWS using AWS Kinesis, runs analytics processing using EMR & Spark and persists the data for analytics.
* Design and develop NiFi flow for incremental data ingestion
* Implement Ranger authorization for data governance and Role-based grants in Big SQL
* Design and Develop end to end Incremental data ingestion including ETL process
* Design and develop Hive query language and Big SQL for Business users
* Design and develop the incremental implementation of the Data Governance in Entire data lake, metadata source. Integrate Atlas and Ranger for Data tagging
* Sound knowledge in data modelling concept such as snowflake schema,star schema modelling,facts and dimension tables.

**Technical Environment**:

AZURE( ADLS Gen2, Azure data factory, Synscort, Databricks, Synapse analytics, Event Hub, Blob, VNet, NSG, Azure Monitor, ARM templates), ServiceNow, AWS, Hortonworks (Ambari, NiFi, Hive, Sqoop, Atlas, Ranger, Knox, Spark, Scala, shell scripts, Avro, ORC), AKS, Docker, Azure board, Azure pipeline, GitHub, Jira

**ROYAL BANK OF CANADA Apr 2018 – Oct 2018**

**Senior Data Architect (Azure HDP/HDF)**

Worked as a Sr. Big Data platform engineer for enterprise data lake of size 400+ nodes, where helped them in the implementation of High availability Hadoop services, optimization and tuning of existing various applications and acted as a subject matter expertise in the enterprise data lake. Some key deliverables:

* Architect and implemented complex bigdata solutions. Developed highly scalable and extensible big data platforms which enable the collection, storage, modelling and analysis of massive data sets.
* Implemented Data lake security in wire and rest, SSL and RPC encryptions
* Implemented Informatica ETL data ingestion solution for cybersecurity
* Implemented the Data Governance in Entire data lake, metadata source. Integrated Atlas and Ranger for Data tagging
* Performed the Entire cluster application tuning, automation, blueprint setup
* Redesigned the existing 400+ node enterprise Big data lake, for effective utilization
* Re-engineered, optimized and performed improving on Spark applications
* Performed tuning on all Hive query and Sqoop jobs in Data lake
* Introduced the data zoning, layering in the data lake
* Performed POC in AWS cloud to offload cold data.
* Re-organized on already deployed application to yield maximum performance from the cluster
* Worked in Spark, Hive, Sqoop application and HDFS data storage

**Technical Environment:**

AWS Data Lake, Glue, S3, Athena, EMR, Informatica, AZURE(ARM template, ADF, ADLS) HD-Insights, Hortonworks (NiFi, Ambari, Hive, Sqoop, Atlas, Ranger, Knox, Kafka, Spark, Pig, Flume, YARN, Pepperdata, Scala, Shell Scripts, Avro, ORC), GitHub, Connect, Jira, AWS

**EQUIFAX FINANCIAL INSTITUTION Nov 2017 – Apr 2018**

**Senior Cloudera Consultant**

Worked as a Sr. Cloudera Hadoop consultant, where helped to develop an application currently running on the mainframe with petabytes of Data into Cloudera Hadoop cluster using, Avro data format and Spark 2 Scala HBase batch application. Accomplishment in these projects:

* Designed the end-to-end solution in Cloudera Hadoop cluster. In this design following components are used, Spark2, HBase, Solr, Avro, Drools
* Designed the HBase data model (Namespace, table and schema designed)
* Developed the Spark Avro parser
* Developed the Spark application end to end using Scala
* Developed the HBase reader and writer API’s
* Integrated the Spark application between, HBase, Solr & Drools

**Technical Environment:**

Cloudera (Cloudera Manager, Spark, HBase, Flume, Hue, Avro, Hive), Shell Script, Git, Scala

**CANADIAN NATIONAL RAILWAY Jun 2017 – Nov 2017**

**Principal Hortonworks Solution Architect**

Worked as a Sr. Hortonworks consultant where helped them to migrate ETL based report generation pipeline into HDP/HDF framework.

* Developed and optimized national railway regulation report generation using Hortonworks HDP 2.6 & HDF 2.2 component stack
* Designed and developed NiFi and Sqoop ingestion framework to ingest data from multiple data sources to the Data lake
* Developed Hive UDF to handle orACLe binary data type
* Implemented and integrated Bug deployment for continuous deployment builds
* Developed Oozie and Hive codes for the materialization of data enrichment in the data lake

**Technical Environment:**

Hortonworks (NiFi, Sqoop, Hive, Oozie, Ambari, Pig), Shell script, Jira, Confluence, BitBucket, Data zonings, Data ingestion framework, Data processing framework

**CONSULTANCY ORG Feb 2016 – Aug 2017**

**Senior Big Data Consultant**

**Client: Freedom Mobile Communications Company (Data Architect)**

* Performed Social media integration, scoring, categorization and enhancements and improvements in promotional campaigns for advertisements
* Designed and developed data ingestion framework to aggregate data from social media twitter logs, marketing channels, Teradata- EDW source and third-party demographic database using Apache NiFi, Sqoop, Oozie and Spark streaming
* Applied machine learning methods, principle component analysis and regression testing on the dataset to build a customer sentiment analysis model through Natural language processing and K-means algorithm using Spark (Scala) and Hive component
* Performed complete Hive Data modelling, Hive data enrichment, coded Hive DDL/DML. Coded Scala-based SPARK programming for data analysis and integrated to micro strategy visualization dashboard
* Installed, configured two nodes standalone NiFi cluster. Transformed the client testing Hadoop cluster into a production cluster by building codes, deploying data zoning and enabling Kerberos security
* Conducted cluster sizing, tuning, performance benchmarking. Introduced YARN queues, ACLs and quote allocations for multi-tenancy
* Worked closely with System Administrators, Data owners and key business leaders to establish SLAs and acceptable performance metrics for Hadoop as a service offering

**Technical Environment:**

Hortonworks (NiFi, Spark, Spark streaming, Oozie, Sqoop, Zeppelin, Ambari), Scala, Shell script. Cluster application administration, Data zonings, Data ingestion framework, Data processing framework, Data consumption framework

**Senior Big Data Consultant**

**Client: Global Data Aimia Advanced Analytics Company**

* Kicked off the Big data implementation with the client, understood their needs in data lake purpose, formulated into requirements and a proposed end to end Hadoop solution stack, platform details along with a detailed implementation strategy
* Performed HDP and HDF cluster sizing, storage sizing and multi-tenancy estimations
* Designed, installed and configured for 22 & 16 nodes HDP and HDF (NiFi) production and development clusters
* Performed YARN capacity scheduler, HDFS schema deployment, ACL enablement, levels of data governance, High availability setup, OS-level disk partitioning etc.
* Conducted cluster sizing, tuning and performance benchmarking on a multi-tenant HDP and HDF platform to achieve desired performance and storage metrics
* Involved in the Active Directory integration and configured Kerberos setup of the cluster
* Created High level / Low-level design and handover documents, along with run books and E2E solution documents
* Mentored the whole team for day to day cluster administration, operation and established complete end to end process flow for post-production support

**Technical Environment:**

Hortonworks (Ambari, HDFS schema, YARN Queue, ACL, AD integration, Kerberos setup, NiFi, Ambari Views, All service, Ranger), Postgres database, Linux, Data zonings, Data ingestion framework, Data processing framework

**Senior Big Data Consultant**

**Client: CPPIB Investment Banking**

* Offloaded ETL application from Informatica and Oracle into Hadoop cluster
* Worked closely with Business owners, Enterprise architects and Application owners to migrate total portfolio application into the HDP cluster
* Installed and configured the HDP 2.5 cluster for development and production environments separately. Also enabled High availability for Hadoop services
* Implemented Ranger, LDAP and Centrify based security solutions into the cluster
* Conducted cluster sizing, tuning and performance benchmarking on a multi-tenant HDP platform to achieve desired performance and storage metrics
* Developed a custom ELT framework to offload the existing Data Warehouse application in Golden gate EDW into a new Hadoop cluster. As a part of this, performed complete Hive data modelling, enrichment and automated the data flow from source to target
* Ingested data from Oracle RDBMS, Oracle Golden gate EDW and other CSV files using Sqoop incremental load, WebHDFS tool into Hadoop cluster
* Processed the data using Spark and Hive LLAP to generate custom reports in TABLEAU as per customer need. Increased usage efficiency of TABLEAU
* Developed a POC to generate near real-time stock options application using Kafka, Spark, Hive LLAP, Pig, HCatalog
* Administrated the cluster and application for 2 months to support day to day operations

**Technical Environment:**

Hortonworks (NiFi, Sqoop, WebHDFS, Hive, Hive LLAP, PySpark, HCatalog, Pig, Kafka, Ambari, HDFS schema, YARN Queue, ACL, Kerberos, Ambari Views, All service, Ranger) PostgreSQL, Oracle, Python, Linux, TABLEAU

**Hadoop Sr.Solution Architect**

**Client: RBC**

The scope of the project was offloading 25% of CPU cycles amounting to around 100 terabytes of data from existing Teradata to Hadoop cluster to save operational expenses.

* Prepared E2E solution design, implemented strategy and Data migration strategy
* Integrated HDP cluster with IBM stage server to import data
* Designed NiFi workflow and ingest data from different source systems in the enterprise. Developed Sqoop scripts to export data to data marts

**Technical Environment:**

Hortonworks (NiFi, Sqoop, Kafka, Sqoop, Spark, Hive, Zeppelin), Data ingestion framework, Data zonings, Data processing frameworks

**Service Delivery**

**Client: National Professional Service organization**

* Implemented HDP cluster for advance analytics
* Installed and configured HDP and NiFi Cluster separately
* Installed and configured Ranger, Knox and KMS for secured cluster
* Installed and configured a Disaster Recovery solution using Falcon with two remote clusters
* Upgraded HDP 2.3 cluster into HDP 2.5 cluster

**Technical Environment:**

Hortonworks (NiFi, Sqoop, Kafka, Sqoop, Spark, Hive, Zeppelin, Falcon, Flume, Ranger, Knox, Oozie, Ambari views, YARN queue, KDC), Scala, shell scripting, Data ingestion framework, Data zonings, Data processing frameworks, Data consumption layer

**ERICCSON, India/Canada Mar 2013 – Feb 2016**

**Senior Solution Architect**

Ericsson is a leading Global product and service-based company. This position involves responsibility for the End to End Lead Solution Architect of proposed client ICT BSS / OSS and Billing within Solution Design, Development and Delivery, at a Global level (North America, South America, Middle East, Africa and South East Asian regions). My role sits within the Ericsson Research & Development and Delivery Organization for Global Telecom Accounts and covers both Solution Delivery and product development activities. I bridge the gap between business challenges and the delivery of BSS/OSS and Convergent-based full stack technical solutions.

**Client: Global Tier 1 Telecommunication Company (Singtel, Singapore)**

* Worked as a Customer Solution Architect to install and configure 80 nodes CRS- (CDH 5.x) cluster to maintain 400 TB of production data
* Conducted cluster sizing, tuning and performance benchmarking on a CRS data warehouse platform to achieve desired performance and storage metrics
* Integrated almost 16 source data systems to Ingested Machine data, Network data, operational data and partner’s data
* Used various ingestion technologies, WebHDFS, messaging system, mediation servers, NFS, Sqoop, flume
* Used Pig, Hive, Impala and HBase for data storage and analysis
* Integrated with downstream systems for data consumption
* Worked closely with Business leaders, Directors, Data scientist, EAs, System Administrators, developers and key business leaders to establish various models, SLAs and acceptable performance metrics for Data as a service offering

**Technical Environment:**

Ericsson CRS, Ericsson charging system, EMM,(Cloudera Manager, HDFS schema, YARN Queue, ACL, AD integration, Kerberos setup, All service, HBase, Impala, Sqoop, Flume, WebHDFS, Oozie, SPARK), Linux, Data zonings, Data ingestion framework, Data processing framework

**Senior Solution Architect**

**Client: Global Tier 1 Telecommunication Company (Globacomm, Nigeria)**

* Led the installation, configuration and benchmarking of Hadoop on 200+ nodes CDH 5.x cluster to maintain 1 petabyte data for an OLAP system
* Conducted cluster sizing, tuning and performance benchmarking on a CDH platform to achieve desired performance and storage metrics
* Developed a customer application to identify customer influence and loyalty data for enhancing customer experience
* Developed the data ingestion framework using Sqoop, Pig, Oozie for ELT and then Hive for Data analysis
* Performed data modelling for Cassandra NoSQL database
* Worked closely with Business leaders, Directors, Data scientist, EA’s, System Administrators, developers, and key business leaders to establish various models, SLAs and acceptable performance metrics for the Data as a service offering
* Developed several scripts and Hive programs in support of data saturation testing

**Technical Environment:**

Cloudera (Cloudera Manager, HDFS schema, YARN Queue, ACL, AD integration, Kerberos setup, All service, HBase, Impala, Sqoop, Flume, WebHDFS, Oozie, SPARK), Postgres database, Linux, Data zonings, Data ingestion framework, Data processing framework

**Client: Global Telecommunications Company (Rogers)**

* Worked as a Customer Solution Architect to migrate existing Oracle-based EDW into a Hadoop based solution
* Installed, configured and benchmarked a 120 node Hadoop cluster to handle 500 Terabytes of data
* Integrated this Hadoop cluster with enterprise northbound and southbound systems to establish end to end data flow
* Developed a data ingestion framework to ingest data into the data lake from Network systems, switching systems, CRM, Billing system, external data source using Sqoop, flume, NFS and processed using Pig, Hive and stores into HBase
* Exported the processed data into Oracle data-mart systems
* Worked closely with client development staff to produce ad-hoc queries, and data analysis on the newly created cross-platform dataset using Apache Hive and Pig
* Worked as a Hadoop cluster administrator and provided operational support for more than 6+ months
* Developed several scripts and Hive/Pig programs in support of data saturation testing

**Technical Environment:**

Cloudera (Cloudera Manager, HDFS schema, YARN Queue, ACL, AD integration, Kerberos setup, All service, HBase, Impala, Sqoop, Flume, WebHDFS, Oozie, SPARK), Postgres database, Linux, Data zonings, Data ingestion framework, Data processing framework

**GLOBAL TELECOMMUNICATION PRODUCT DEVELOPED COMPANY (Ericson/Huawei)**

**Lead Solution Architect**

* Worked in Research and Development unit where I had developed Next-Generation IoT full stack Big Data solution to integrate Enterprise level OLTP and OLAP systems.
* Worked in mission-critical Data based applications, where I have developed codes for Pig, Hive and Spark for converged solutions
* Designed and developed various structured and unstructured Data activity requirements like data ingestion, real-time data streaming, Data management, Data governance and Metadata management
* Designed and developed Datastax Cassandra Enterprise Big data analytics and search solution using Spark & Solr. Developed SOAP-based web services and performed complex application integration
* Performed numerous Linux and Unix server installations, product and service configurations in large data centres
* Migrated a Global Tier-1 NOC into GNOC and successfully transformed all 10 million subscribers flawlessly into the GNOC
* Worked as a Subject Matter Expert in Telecom BSS / OSS (Business support system / Operation support system)
* Expert in real-time charging and Billing system development
* Strong knowledge in telecom e-Tom, SID information models
* Worked as a national managed service manager where I established the GNOC process design, tools design, identified roles and responsibilities, recruited resources, identified and maintained the KPI and SLA for the customer

**Telco Vendor**

**Managed services May 2012 – Feb 2013**

* Establishing gnoc for telco clients
* Managed a people of 60
* Day to day operational maintenance for gnoc
* Platform reengineering
* Perform system integration for 3rd party

**Technical environment:**

OSS/BSS

**Ericcson**

**Lead solution architect Jan 2007 – April 2012**

* Architecting OSS/BSS System
* Performed data modeling for data warehousing systems
* Developed data ingestion framework
* Performed real time system integration
* Transforming business requirement into technology design
* Performed high level and low level design and automated data ingestion
* Developed data marts for different business users
* Requirement gathering, business analysis, system analysis and implementation solutions
* Performed end to end testing and delivered to clients

**Technical environment:**

Oracle data guard, D/W (Oracle, Hbase), Charging system, billing system, Linux, Shell script, SQL

**Preceding Assignments**

Jan’06–Jan’07 with Reliance Communication, Mumbai, Maharashtra as Team Leader

Jan’2003-Jan’2006 with BSNL Trichy, Tamil Nadu as Telecom Officer

Jun’2001-Jan’2003 with Alpha Software System Pvt. Ltd., Chennai, Tamil Nadu as Network Engineer

    

