| Name: Subramonian Kumaresapillai  subbunear@gmail.com  206 245 6565  LinkedIn: https://www.linkedin.com/in/ksubram/ | |
| --- | --- |
| **Security Clearance:** | * DHS CBP Suitability Clearance 11/4/2019 * Top Secret clearance |
| **Education:** | * Masters of Business Administration (MBA) Seattle Pacific University, WA, USA Graduated on 2010. * MS (Computer Science) Madurai Kamaraj University Graduated on May 1996 * BS (Computer Science) Manonmaniam Sundaranar University Graduated on 1993 |
| **Certications and Special Training** | * AWS Certified Solutions Architect – Associate * Apache Hadoop Architect/Developer Training by Cloudera, CA, USA * ITIL v3 |
| Skills, Expertise and Knowledge Qualifications | |
| **Summary:**   * Over 21+ Years of extensive IT experience as a Big Data/Data Sciene/Cloud Architect. * Hands on experience in Data Analytics & Data Science. * Expertise in providing architectural directions using various AWS services and solution methodologies. * Experience in Agile/Scrum and Waterfall methodologies. * Domain experience in Federal, Aerospace, Retail, Hospitality,Telecom and Government sectors. * Expertise in Big Data, Data Lake, No-SQL and Open Source technologies. * Ability to work on multiple complex projects simultaneously in a team-oriented environment and tight deadlines. * Interpersonal skills to effectively work with clients and partners at all levels. * Outstanding leadership ability to direct, mentor and lead cross-functional teams.   **Skills:**   * Apache Hadoop (HDFS, Yarn, Mesos, Scala , Hive, Storm, Spark(2.1.1, 2.4.3), Accumulo, NiFi, Cloudbreak, Sparkmagic, Livy, Jupyter ) * Trifacta, DataRobot, Dataiku, H2O (**Data Wrangling & AI, Modeling tools** ) * NoSQL Databases (Cassandra, HBase,MongoDB) * AWS Services: S3, EC2, EMR,VPC, IAM, RedShift, DynamoDB, RDS, Athena,Glue * Websphere MQ, BEA Tuxedo, JMS, Condor, ETL, R. QlikSense * MS Visio, MS Project, Eclipse, TOAD, XMLSpy, Oxygen * Horton works, MAPR, Cloudera * SQL, XML, JSON, Perl, Python(2.7, 3.x), Anaconda * Java/J2EE, iBatis, Web Services * Oracle, DB2, SQL Server * Apache, Tomcat, Weblogic, Websphere, JBoss * Git, GitHub, JIRA, Confluence, Jenkins, Cruise control, Quality center, Clear Quest, SharePoint, Version one, Micro services, ,Pipeline, Nagios, Monit, Postfix, Dovecot. * ARIS, Agile/Scrum, Lean * HP-UX, Red hat Linux, Mac OS | |
| Professional Experience | |
| ManTech International  Senior Big Data Architect November 2019 – Present   * Responsible for the transformation of the current RDBMS Operational store to UADI Datalake as part of the Business Intelligence Support Services (BISS) to supporting TASPD for analytics. * Developed Spark/Scala framework to process larger volume of prep data to support TCNRM models (Spark/Scala,DataRobot, HBase, Impala, Hive and Cloudera) * Responsible for the implementation of UADI AWS Cloud Architectural Initiatives, technology roadmap and Proof of concepts. | |
| Synchronoss Technologies  Senior Big Data Architect Feb ‘19 – Nov ‘19   * Responsible for re-architecting data analytics product (SIP – Synchronoss Insights Platform) * Defined the Technology stack to support development for IOT. * Performed Query optimization, Spark performance tuning and troubleshooting (Spark 2.4.3, Scala, Python, MAPR, Hive,AWS) * Responsible for capturing and processing large volume of data to provide insights and make valuable decisions. * Designed the product architecture to be configurable for multi-client support. | |
| Hilton Worldwide McLean Virginia  Senior Bigdata Architect(Contractor) January ‘18 – February ‘19   * Create and maintain complex data pipeline architecture. * Designed and developed solutions for multiple initiatives to support the migration and transformation from multiple data sources using NiFi, Kafka and Sqoop. * Build re-usable data pipelines to ingest, standardize, and shape data from various zones in data lake. * Architected and developed data lake solutions using Scala, Spark, Hive, HBase and HDFS. * Data Analytics using Horton Works Data Platform 2.6 (Spark 2, HDFS, Hive, HBase, Ranger, Kerberos, Ambari etc.) * Designed the Data model for HBase using Spark processing. * Implemented data level security-using Ranger for Hive, HBase and HDFS. * Worked on architectural assessment, solution outline, Roadmaps (future state architecture) and vendor product assessment. * Build analytic tools to deliver actionable insights into customer acquisition, revenue management, digital and marketing areas for operational efficiency and critical metrics. * Design and build APIs on established enterprise Architecture patterns for data sharing from various sources. * Fussy Match analysis, Name matching and Loqate implemented using match IT Spark/HBase. * Implemented Business data Zone using AWS S3 & Redshift. * Responsible for translating the customer needs, and technology opportunities into detailed technical offering, solutions and proposal. | |
| Booz Allen Hamilton/Take2, Herndon Virginia  Senior Consultant (Contract) Apr’17 – Dec ‘17  Client: Centers for Medicare & Medicaid Services(CMS)   * Architected and developed solutions for Data extraction, Ingestion, processing, storage, querying and analyzing structured, semi-structured and unstructured data from multiple data sources (JSON, CSV and Text formats). * Worked to solve technical problems for the partners. * Designed and developed solution for a large-scale data lake on HortonWorks multi-node cluster with full suite of Hadoop ecosystem - YARN, HDFS and Hive. * Developed Spark applications using Spark (1.6.3) in Scala. * Responsible for Micro Services Architecting and building distributed data solutions using Mesos DC/OS, Dockers, ECS etc. * Implemented Spark/EMR/S3 based data processing solutions for multiple CMS projects. * Implemented AWS (VPC, EC2, S3, RDS, IAM, Splunk, DCS etc.) in multiple projects. * Responsible for troubleshooting various CMS AWS environments (Dev. Impl & Production). * Implemented AWS Monitoring using Nagios and Monit. | |
| CGI Federal, Fairfax Virginia  Lead Big Data Architect Mar’16 – Apr’17  Client : Environment Protection Agency(EPA)   * Worked on analyzing EPA Horton Works cluster using different Big data analytics tools such as Kafka, NiFi, Sqoop, Jupyter, Docker, Python, Livy, Storm, Spark and Hive. * Implemented Elastic search for the EPA’s Enterprise portal (Development and Infrastructure). * Leverage data science tools and technologies to implement and enhance the computational performance of clients' scientific/machine learning models. * Migrated a Chesapeake Bay Scenario Builder model step performed using SQL stored procedures to Apache Spark using pandas data frames, resulting in the ability to run the model hundreds to thousands of times faster depending on the number of inputs and AWS EMR cluster size. * Worked on advanced monitoring and remote sensing data to predict likelihood of Harmful Algal Bloom Events using SQL Server, Apache Storm and Microsoft Azure. * Developed prototypes to illustrate approaches to suite technology and business problems. * Responsible for analyzing, designing and developing end-to-end technical solutions for the customers based on their specifications and business needs. * Implemented Nifi streaming to receive real time data from the Twitter and store the stream data to S3. * Implemented AWS technologies and Services (VPC, EC2, S3, RDS, IAM, Athena, EMR etc.) for multiple projects. * Implemented Qilk visualization. * Built CGI’s Datalake platform using Horton Works clusters in the cloud. * Developed Mesos Spark applications using AWS, Dockers, Spark Magic, and Livy for the EPA. | |
| Booz Allen Hamilton, McLean, Virginia  Senior Lead Data Scientist Apr’15 – Feb’16  Clients : Department of Labor & National Geospatial Agency   * Worked on NGA’s data analytics using Apache Storm, Kafka and Accumulo. * Provided thought leadership towards architectural and technical approach to the Office of Information Technology (OIST) at Department of Labor (DOL) for Employee Training Administration (ETA) initative. * Conducted a study of ETA’s existing data architecture and analytical approaches. * Designed and implemented Data Lake platform POC that aligns with ETA’s modernization objectives, with the focus on providing reusable and consolidated data models to simplify and streamline analytics across the organization. * Developed a strategy for Full load and incremental load using Sqoop. * Mainly worked on Hive queries to categorize data of different claims. * Installed HortonWorks Data Platform on AWS and physical machines for the POC’s. * Involved in Design and Architecting of Big Data solutions using Hadoop Eco System. | |
| The Boeing Company, Seattle, Washington  Architect/Technical Lead Jul’05 – Apr’15  Projects  1) 787 –Maintenance Performance Toolbox  2) Toolbox - Cloud/Data Analytics/Big Data  3) Airplane Health Management Integration  4) Toolbox - 2016   * Worked closely with the customers, finalized requirements and created functional specifications. * As a Subject matter expert, lead a 15-member technical team in 787 CAS IT following the agile methodology. * Involved in Sprint planning and represented the team in Scrum of Scrums, Retrospective and Demos. * Developed proof of concept for 787 digital initiatives and worked with the development team to implement these solutions. * Worked on Business Cases and proposals to support new initiatives. * Performed machine learning research using various algorithms (KNN, decision trees, naïve Bayes, and regression) for the airline data for predictive analysis. * Developed strategies and architectural designs for the Hadoop based Data Consolidated Platform (DCP) that allowed cross-functional teams to enforce data governance, enabled data exploration and reduced operating cost. * Architected a proof of concept to establish a Data Lake platform using HortonWorks to process large amounts of data. * Architected and developed a highly scalable search mechanism using Elastic search. * Improved the scalability of the application with data partitioning, redundancy and database restructuring. * Leveraged AWS cloud services like IAM, EC2, VPC, S3, RDS, SQS and EBS for installing, configuring and troubleshooting. * Developed a migration strategy to move the existing environment to AWS Cloud, which reduced the turnaround time for application development & deployment and also reduced the operational cost. * Developed the technology roadmap, planning and execution against the roadmap. * Used Chef and implemented easy application deployment to the AWS & on premise that reduced operational cost & deployment time. * Reviewed existing architecture and implemented Microservices based solutions. | |
| SolutionsIQ, Bellevue, Washington  Technical Architect May’04-Jul’05  Client: Washington State Gov., Olympia, WA  AOC Justice Information System | |
| Hewlett Packard (HP), Worldwide IPG, Palo Alto, CA  Technical Consultant May’00-Apr’04   1. **HP Shopping, HP Product Registration**   **Team Leader / Consultant**   1. **Foundation Architecture Infrastructure (FAI)**   **Technical Lead / Consultant**   1. **Real Time Emailer (RTE)**   **Technical Lead/Consultant**   1. **CLAR (Common Login Application & Registration - Workflow Framework Wizard)**   **Consultant**   1. **HPD Load Site**   **Consultant**  **Responsibilities:**   * Development, Maintainence for core HP systems used in HP Registration/Printing division | |
| Consultant - Infogain, CA Nov’99 – Apr 2000  Software Engineer - NetERP India Private Ltd, Bangalore, India Jun 96 – Oct 2000 | |