

Arun Muruganandam

Big Data Engineer

Singapore

- Email me on Indeed: <http://www.indeed.com/r/Arun-Muruganandam/9fd24ee23fd70f37>

- Around 12 years of experience in Big Data platform and Software Implementation, Upgrade and Migration projects.
- Understanding the business requirements, needs and drawing the roadmap for big data platform.
- Cloudera Certified Administrator for Apache Hadoop, Monitoring, job performance and capacity planning of the Hadoop cluster.
- End-to-end installation configuration/administration of multi-node Hadoop cluster CDH 5 in cloud and data center environment.
- Very good exposure on Hadoop technology like HDFS, Yarn, MapReduce, Hive, HBase, Sqoop, Zookeeper, Spark, HUE, Flume, Zookeeper, Oozie, and Sentry.
- Design and Developed Predictive Model to identifying credit card defaulters using Python Spark.
- Expert in understanding the data nature, Designing, Develop, schedule and maintain ELT jobs for seamless ingestion of data into the Big data platform using Oracle ODI 12c.
- Design and Develop streaming processes to enable real-time data ingestion, processing and transforming using flume and Spark Streaming.
- Working on various uses cases including Analytics, processing, storing and Reporting of Voluminous, rapidly changing, Structured, Semi-Structured and Un-Structured data.
- Mining the Social web and generating data asset for Banking and Telecom Domain using python.
- Working on Hive Metastore backup, Partitioning and bucketing techniques in a hive to improve the performance.
- Involved in Kerberos security Implementation in AD Integration with Cloudera Manager
- Created users in Active Directory and map the roles in each group for the users in Apache Sentry
- Involved in LDAP implementation for different types of accesses in an AD for Hue and Hive
- Working knowledge of SQL, PL/SQL for stored procedures, ELT tool, user defined functions, and HTML.
- Working knowledge in Java Script, Shell Script and Perl Script.
- Follow-up all request for change which compliance to ITIL processes.
- Responsible for application deployment, backup, rollback and operation maintenance.
- Involved in creating documentation of High level/Low level design and Functional Spec
- Hands-on, a motivational team leader who maintains an authoritative presence and 24 X7 technical supports

Work Experience

Senior Big Data Engineer

DBS Bank - Singapore

August 2019 to Present

- Working on Hive Metastore backup, Partitioning and bucketing techniques in hive to improve the performance.
- Working closely with BAU team and understand the requirement clearly and create hive table on HDFS.

- Understanding the data nature, Designing, Develop, schedule and maintain ELT jobs for seamless ingestion of data into the Hadoop platform
- Design and Develop streaming processes to enable real time data ingestion, processing and transforming using flume and Nifi.
- Working on various uses cases including Analytics, processing, storing and Reporting of Voluminous, rapidly changing, Structured, Semi Structured and Un-Structured data.

Big Data Engineer

CIMB Bank - Malaysia

October 2017 to Present

Duration October 2017 to Till Now

Environment Oracle BDA 4.7, Cloudera CDH 5, HDFS, Map Reduce, Hive, Flume, Impala, Hue, Python, Oracle R, Jupyter Notes, R Studio, Perl Script

Responsibility:

- Understanding the business requirements, needs and drawing the road map for big data initiatives.
- Management of Hadoop cluster CDH 5.9, with all included services, including resolving any ongoing issues with operating the cluster in data center environment.
- End-to-end installation configuration/administration of multi node Hadoop cluster CDH 5.13 in Cloud environment
- Installation and Configuration of spark 2.x and run Python Spark computation securely on data in Hadoop cluster for data exploration, analysis, modeling, and visualization
- Design and Developed Predictive Model to identifying credit card defaulters using Python Spark.
- Design and develop data model and optimizing queries on Hive.
- Working on Hive Metastore backup, Partitioning and bucketing techniques in hive to improve the performance.
- Working closely with Data science team and understand the requirement clearly and create hive table on HDFS.
- Understanding the data nature, Designing, Develop, schedule and maintain ELT jobs for seamless ingestion of data into the Big data platform
- Design and Develop streaming processes to enable real time data ingestion, processing and transforming using flume and python spark streaming.
- Working on various uses cases including Analytics, processing, storing and Reporting of Voluminous, rapidly changing, Structured, Semi Structured and Un-Structured data.
- Mining the Social web and generating data asset for bank using python.
- NEWS streaming data has been extracted and analysis for insurance X-sell based on Incidents in nearby location.
- Property portal data has been extracted and analysis refine hidden preferred customers.
- Troubleshoot the R-Studio/Hive/Spark issues raised by Data science team and provide the solutions same
- Involved in Kerberos security Implementation in AD Integration with Cloudera Manager
- Created users in Active Directory and map the roles in each group for the users in Apache Sentry
- Involved in LDAP implementation for different types of accesses in AD for Hue, Hive
- Involved in creating documentation of High level/Low level design and Functional Spec
- Discussed issues with multiple Vendors on the meeting, update progress status and report to team head.

Big Data Engineer

Vodafone Telecommunication
February 2017 to September 2017

India

Duration Feb 2017 to Sep 2017

Environment Cloudera CDH 5, HDFS, Map Reduce, Hive, HBase, Sqoop, Spark SQL,
Zookeeper, Impala, Hue, and Oracle BI Publisher 11g

Project Description:

Historical data analysis is an essential feature for Vodafone India, a clear requirement to keep the data in archived form to free up storage from operational systems. It has also been important for them to perform advanced analytics that help improve business decisions. Vodafone India has been archival the old data.

Responsibility:

- Data has been archived and purged based on archival and purging policies.
- Archival Data is to bring huge volumes of data from Oracle DB into Hadoop environment.
- Archival data provide support for various uses cases including Analytics, processing, storing and Reporting of

Voluminous, rapidly changing, structured data.

- Understanding the data nature and designing the injection processes for HDFS using Sqoop.
- MapReduce and Spark used for data processing and transformation.
- Designing data model and optimizing queries on Hive and spark SQL.
- Designing HBase (No SQL DB) column schemas.
- Partitioning and Bucketing techniques in hive to improve the performance.
- Exploring the feasibilities of Oracle business publisher as the reporting tools.
- HUE UI interface will be used for dynamic query & batch retrieval of archived data.
- RESTful API interface between HBase and CRM screen for online retrieval of archival data.
- Using Kerberos and LDAP security authentications.
- Discussed issues with multi stakeholder on the status call and update daily progress status and weekly status report to management.
- Driving initiative to automate the recurring manual activities for monitoring and operations using Unix Scripting.
- Plane release schedules and align with project and program timelines
- Involve in customer experience transformation, Risk/Assumptions/Issues/Dependencies (RAID) status.

Software Consultant

StarHub Ltd - Singapore

July 2016 to November 2016

Environment Siebel IP15, Oracle DB 12c, Oracle BI Publisher 11g, SQL Developer

Project Description:

In Siebel e-communication application, Sales Order Management (SOM) focuses on the creation, modification and completion of Order and End to End solution. The key functional areas are Sales Order, Asset, Network, activities and Billing.

Responsibility:

- Involved in Siebel application upgrade from Siebel 8.0 to Siebel IP15 development, testing and deployment.

- Sound technical knowledge in Siebel application Usage/Navigation/Familiarity, Business Objects/Components, Applets, Views, LOVs, Positions, Responsibilities, Workflow Debugging/Troubleshooting and Siebel econfigurator.

- Involved with the Siebel application design team to analyze the technical feasibility of solutions provided and

Understanding the impact of the designs.

- Sound technical knowledge in SQL queries.

- Involved in testing that included review of the requirement documents, studying the functional requirement

Page3

document for developing test plans, test scenarios, test scripts and test cases.

Solution Architect

JAI Accounting Pte Ltd - Singapore

February 2016 to June 2016

Duration Feb 2016 to June 2016

Environment AWS, Cloudera CDH 5, HDFS, Map Reduce, Hive, HBase, Sqoop, Spark, Zookeeper, Impala, Hue

Project Description:

Extract data from social media and creating a data framework and deploying tools to capture, transform, analyze and store of structured and unstructured data

Responsibility:

- Installed and configured Cloudera Hadoop cluster on AWS environment
- Working on various uses cases including Analytics, processing, storing and Reporting of Voluminous, rapidly changing, Structured, Semi Structured and Un-Structured data.
- Mining the Social web and generating data asset for Banking and Telecom Domain using python.
- Loaded semi structured and unstructured data into Hadoop File System (HDFS).
- Design and Develop streaming processes to enable real time data ingestion, processing and transforming using flume and Spark streaming.
- Created hive table on HDFS and performance tuning on Hive.
- PPC Manager and optimization of client accounts utilizing Ad Words, Facebook, LinkedIn and Twitter
- Creation, management and optimization of AdWords search campaigns for 50 small to medium business clients in the Singapore. Extensive reporting, integration and analysis.
- Pay Per Click (PPC) campaign setup and optimization, online media buys and optimization of Google AdWords and SEM through ongoing testing of ad copy, landing pages, bidding strategies, ad extensions and more.

Software Consultant

Emerio GlobeSoft Pte Ltd - Singapore

February 2013 to February 2016

Environment Siebel 8.0, BSCS Billing, EAI, Oracle 11g, PL/SQL Developer, Dynatrace APM

Project Description:

In Siebel e-communication application, Sales Order Management (SOM) focuses on the creation, modification and

Completion of Order and End-to-End solution. The key functional areas are Sales Order, Asset, Network, activities and Billing

Responsibility:

- Involved with the design team to analyse the technical feasibility of solutions provided and understanding the impact of the designs, review the IA and approved.
- Provided reports to the business users by exporting or scanning the affected data from Siebel to provide user count reports, adhoc reports, documenting the production support procedures.
- Investigation, troubleshooting, and analytical work to resolve user's queries on SOM, Network and Billing Cases.
- Identify the problem root cause, raise and fix the defects.
- Execute workaround steps to facilitate business operations. e.g. perform the SOM related order recover
- To work with system engineers and database administrators to ensure the high availability of applications.
- Develop automation workaround and proactively propose solutions to constantly improving the systems and business operations through CIP (Continuous Improvement Program).
- Perform Billing Dispute and Conduct the data reconciliation among of Siebel and BSCS.
- Application monitoring via Dynatrace tool.
- Knowledge Transfer / Knowledge Sharing with team members and other relevant colleagues.
- Involved in 24/7 on-call support for production environment and UAT environment

Application Analyst

MDIS

May 2011 to February 2013

Environment SQL Server, C#, Visual Studio 2010, ASP.NET

Description: University Management System (UMS) application is a framework for student management of MDIS Business College integrated with UOB (UK), UOS (UK), OCU (USA) and SCU (Australia) Universities. Student Relationship Management (SRM) module handles the services provided for the students in the university. This Module primarily concerns with Configuration and Administration of student details.

Responsibility:

- Requirement gathering from user to design the user approval and entitlement to executive system.
- Involved in the development of Presentation Logic for GUI of ASP.NET pages.
- Used ADO.NET to interact with SQL Server Database.
- Used Asp.net validations Controls for side validations.
- Data migration projects to combine student information from one application to another application.
- Utilized Quality Center for directing, tracking and outlining needs during a project SDLC.
- Formed detailed UAT and QA test plans, managed system testing and preserving testing documentation.
- Produced complete data mappings and data conversion need documents.
- Carried out regular triage meetings to recognize and determine critical testing faults.

Project 7:

Siebel Developer

Tata Consultancy Services

February 2008 to April 2010

Client Nokia e-consumer Good Application, Finland

Duration February 2008 to April 2010.

Environment Siebel 7.8.2.4 Tools, Oracle 10 G, SQL Developer.

Project Description:

Nokia Xpress is a framework for Application Management (AM) of Nokia B2B applications. Retail Net is a Nokia

Business analysis application and it consists of three major modules FFM- Front End, FFM- Back End and Analytics.

Siebel EIM Developer:

- Involved in performing the EIM tasks and preparing the required mapping to migrate, Subject data to populate the Siebel base tables through EIM Interface Tables.
- Used SQL Loader Utility and Control Files to stage the data from Flat file, which is coming from the legacy System.
- Validated the data mapping sheet using interface table reference guide and tools.
- Involved in deleting the duplicate data from the Siebel database using EIM process.
- Changing the parameters for debugging the EIM process whenever we faced errors and hence running it again and again to get a comprehensive error log file.
- Executed EIM task by creating IFB file with the required parameters.
- Discussed issues with the team members on the status call.

Siebel Administrator:

- User Creation, Generate new DB, DB extraction, Handling Ad-hoc request, Server Restarting.
- Creating LOV's, Views, Responsibilities, Remote Administration, Assignment Manager, Workflow, synchronization and Monitoring EIM Batch jobs.
- Recommend resolution of performance issues at the database, architecture and application level.
- Address existing performance and stability bottlenecks.
- Prepared Backup and Recovery Strategies/Coordination with client vendors.
- Application server Sizing.
- Routine User Support and performance tuning/Offshore interaction and management.
- Application Administration and Understand the concept of Integration & data exchange etc
- Release Planning
- Server Monitoring, Component Monitoring, Task, GMB and ITSM Monitoring.
- Prepared a Procedure document for "Maintenance & Operation Guide" and "Deployment & Rollback

Task

Page5

Guidelines"

Education

Master of Technology in Technology

School of Computing - Thanjavur, Tamil Nadu

2005 to 2007

Bachelor of Technology in Technology

School of Electrical - Thanjavur, Tamil Nadu

2002 to 2005

Skills

DATABASE (8 years), SQL (7 years), ORACLE (7 years), SIEBEL (5 years), APACHE HADOOP HDFS (5 years)

Certifications/Licenses

Cloudera Certified Administrator for Apache Hadoop

May 2017 to Present

Additional Information

Technical Skills:

- Big Data Platform: Cloudera CDH 5 Cluster and Oracle BDA 4.7
- Hadoop Technologies: HDFS, Yarn, MapReduce, Hive, HBase, Sqoop, Zookeeper, Spark, HUE, Flume, Zookeeper.
- ELT Tool: Oracle ODI 12c
- Programming Languages: Java, Python, R, C, C++, and C#
- IDE: Eclipse, Jupyter, R studio and Spyder
- Scripting language: Java Script, Perl, UNIX shell scripts
- Database: Oracle, MS SQL Server, MYSQL and MS Access
- ReportiTool: OBIEE 12c, Crystal Report
- Siebel CRM Applications: Partner Manager, Partner Portal, e-Consumer good and e-Communications Application.