**Surendra Kumar** Email ID: surendrakumarbd@gmail.com

Big Data Consultant Phone: 6128897646

|  |  |
| --- | --- |
|  |  |
|  |  |

**Summary of experience:**

**Over 12 years of experience** implementing large scale IT applications predominantly in Consumer Product group, Manufacturing, Telecommunications, Entertainment Industries using BigData Technologies, Amazon Web Services,Data warehousing, Business Intelligence Applications as Solution Architect and Developer.

**Professional Experience:**

* 2+ years of hands on experience using **BIGDATA - Hadoop /Spark Ecosystems including HDFS, Map Reduce ,HIVE ,PIG, Talend Connectors, Spark SQL ,Spark Streaming**
* Experience with large scale **Hadoop/Spark environments** to build and support including design, configuration, installation, performance tuning and monitoring.
* Experience in AWS,EC2,S3 –Cloudera Distribution .
* Good amount of experience ingesting both structured and unstructured Data using **Sqoop**, **Flume** into HDFS from Legacy systems including Streaming Data.
* Involved in the development of **POC's** to migrate map reduce programs into Spark RDD’s.
* Hands-on experience in developing ETL solutions using **Talend** for Big Data
* Familiar with Python, Scala programming languages.
* In depth knowledge and Hands on experience using **NOSQL** databases like HBase and Cassandra.
* Hands-on experience in setting up and configuring **Apache Hadoop** and **Cloudera CDH** clusters on **Ubuntu, Red Hat** Linux distribution environments.
* Good Expertise working with different types of data including semi/un-structured data.
* Experienced in optimization techniques in sorting and phase of **Map reduce programs**, and implemented optimized joins that will join data from different data sources.
* Hands on Experience in performing analytics on structured data in Hive with Hive queries, Views, Partitioning, Bucketing and UDF’s using **HiveQL**.
* Experience in performance tuning of HIVE queries and Map Reduce programs for scalability and faster execution.
* Experience in Solr full text Search Platform. Familiar with Splunk and Elastic search.
* Hands on experience in writing Map Reduce jobs on Hadoop Ecosystem using **Pig Latin** and creating Pig scripts to carry out essential data operations and tasks.
* Experience in processing different file formats like **XML, Avro, Parquet, JSON** and sequence file formats
* Used **Apache Oozie** to create workflows by combining multiple jobs for Map Reduce, Hive, Pig, Sqoop into one logical unit of work.
* Experience in **reviewing and managing logs**.
* Good Knowledge on Cloud Computing with Amazon Web Services like EC2, S3 .
* Experience in using Cloudera Manager for monitoring cluster machines.
* Experienced with using different kind of **compression techniques** to save data and optimize data transfer over network using **Lzo, Snappy**, etc.
* Good Experience in **Agile Engineering** practices, **Scrum methodologies**, and **Test Driven Development** and **Waterfall methodologies**.
* Hands on experience in database design using PL/SQL to write Stored Procedures, Functions, Triggers and strong experience in writing complex queries, using Oracle, DB2 and MySQL.
* Ability to transform complex business requirements into technical specifications.
* Diverse Experience in working with variety of Database like **Oracle**, **MySQL**, **Teradata**, **DB2.**
* Hands-on Experience in Object Oriented Analysis, Design **(OOAD)** and development of software using **UML Methodology.**
* Good Experience in creating Business Intelligence solutions (ETL using SAP Data Services, Business Objects, SAP BW, Tableau).
* Good working experience on different OS like UNIX/Linux, Apple Mac OS-XWindows.
* Experience with Implementation of Java **Web services (JAX-WS and JAX-RS).**
* Excellent project documentation skills which includes creation of design documents which constitute UML diagrams.
* Exhibited strong written and oral communication skills. Rapidly learn and adapt quickly to emerging new technologies and paradigms.

**Technical skills**

|  |  |
| --- | --- |
| **Big Data Technologies** | Hadoop, HDFS, YARN Map Reduce, Hive, Pig, Sqoop, Flume, HBase, Cassandra, Spark. |
| **Business Intelligence Solutions** | WEBI, Crystal, SAP BW, HANA,Tableau. |
| **ETL** | Talend ,SAP Data Services (BODS) |
| **RDBMS** | Oracle, My Sql, and DB2 |
| **Programming Languages** | SCALA, PYTHON, JAVA, SAP ABAP |
| **Operating Systems** | Windows , Linux, Unix, Solaris |
| **Other** | SAP ERP,CRM,SEM,BW,HANA |

**Work Experience**

**Novartis Jan 2016 – Till Date**

**Hadoop Solution Architect/Developer**

**Responsibilities:**

* Involved in Design/Developing Hive Warehouse using Hive/Impala/SPARK SQL
* Manage HDFS/Raw Ingestion layer/Warehouse.
* Build data Pipelines/ELT using shell scripting and Python Scripting
* Work with multiple data sources (SAS,Files)
* Worked with multiple files (SAS, text, sequence, Parquet) serialization/De-serialization.
* Ingest data into S3,EC2 Cluster.
* Create Hive Tables and load data.
* Create proto type on SPARK Core/SQL /Streaming

**NTT DATA August 2013 – Dec 2015**

**Hadoop Solution Architect/Developer**

**Responsibilities:**

* Involved in complete SDLC life cycle of big data project that includes requirement analysis, design, coding, testing and production
* Extensively Used **Sqoop** to import/export data between RDBMS and hive tables, incremental imports and created Sqoop jobs for last saved value.
* Created **Talend** workflows to configure various jobs using various **Bigdata Connectors like Hive ,HBASE** consumers to provide clean data abstraction between producers and consumers.
* Created Map Reduce programs in order to analyze data and used **Pig Latin** to transform data.
* Installed and configured Hive and wrote **Hive Generic UDF** to successfully implement business requirements.
* Involved in creating hive tables, loading data into tables and writing **Hive** queries.
* Experienced with using different kind of **compression techniques** to save disk usage and optimize data transfer over network using **Lzo, Snappy** etc in **Hive** tables.
* Implemented custom interceptors for **flume** to filter data and defined channel selectors to multiplex the data into different sinks.
* Implemented POC to migrate map reduce jobs into **Spark RDD** transformations**.**
* Developed Scripts and Auto sys Jobs to schedule a bundle (group of coordinators), which consists of various Hadoop Programs using **Oozie**.
* Used Solr for Full Text Search.
* Hands on experience with Accessing and perform CRUD operations against **HBase** data using T**alend**.

***Environment:*** *Hadoop, Spark, HDFS, Map Reduce, HBase, Talend, Hive, Flume, Sqoop, PIG, CDH*

**Warner Bros, January 2013 – August 2013**

**Role: Hadoop Developer**

* Installed and configured Hadoop **Map Reduce**, **HDFS** and developed multiple Map Reduce jobs in Java for data cleansing and preprocessing.
* Responsible to manage data coming from different sources.
* Involved in gathering the business requirements from the Business Partners and Subject Matter Experts.
* Proactively monitored systems and services, architecture design and implementation of Hadoop deployment, configuration management, backup and disaster recovery systems and procedures.
* Supported **Map Reduce** Programs those are running on the cluster.
* Involved in HDFS maintenance and loading of structured and unstructured data.
* Written **Pig Latin** scripts for data transformation and handling unstructured data.
* Imported data using **Sqoop** to load data from **Oracle** to HDFS on regular basis.
* Creation of **HBase** tables to store variable data formats of data coming from different applications.
* Created **Hive** queries for data analysis to meet the business requirements.
* Loading log data into HDFS using **Kafka, Flume** and performing ETL Integration.

**Environments**: Hadoop, HDFS, Map Reduce, Hbase, Hive, Pig, Sqoop, Kafka, Flume, Linux, Oracle 10g.

**Prior Experience:**

|  |  |
| --- | --- |
| **Company** | **Role** |
| General Mills -Jun 08 to Jan 2013 | Business IntelligenceDeveloper/ DataArchitect (SAP BW,BODS ETL, Business Objects ,HANA,Oracle) |
| ConAgra Foods -Jan 07 to May 08 | Business Intelligence Developer (SAP BW,BIA,BODS ETL, Business Objects) |
| SAP GDC Sept 06 to Dec 06 | SAP BI Consultant |
| IntelliGroup - JUN 04 to SEP 06 | Data Warehouse Consultant |

**General Mills: Jun 2008-Jan 2013**

**Business Intelligence Data Architect/Consultant.**

**HANA/Data Services (BODS) Implementation:**

* Gather/Document user requirements.
* Design Review Data acquisition Models.
* Create Open Hubs from BW to extract data from BW
* Create HANA Landing Physical Tables
* Create Attribute, Analytical, Calculation Views.
* Create BODS Jobs to accommodate delta Data from Open Hubs.
* Created Roles and Security in HANA.

**BO Explorer Migration to HANA:**

* Worked with Business SME’S to understand Requirements.
* Reviewed existing Business Objects Explorer Models on BWA.
* Configured necessary Security for Project Implementation and User Setup.
* Designed Necessary HANA Information Models (Attribute, Analytical, Calculation Views).
* Created Physical tables and Hana Views.
* Setup Security, Application Package, Connection and Jobs in Data Services.
* Created New Information Space, Security and Connection setup Business Objects Explorer

**Environment:** SAP BOBI 4.0 SP4/ SAP BOBI 4.1 SP1, Dashboard Design 4.0, Design Studio 1.1, Lumira 1.1, SAP BW, HANA database

**Trade Reporting:**

* Interacted with **end-users** and **business analysts** to identify and develop **business requirements** and transform it into **technical requirements**
* Analyzed multiple approaches to meet requirements by giving development & testing estimation and ultimately responsible for delivering the solution
* Built real-time & non-real time dashboards for multiple business units like Trade Marketing, Trade Finance using SAP Dashboard Designer 4.0/Xcelsius
* Used different data connections like **BICS, BIWS, QAAWS, Flash variables** etc. effectively to meet the user needs and supported them with detail level drill-down Web Intelligence reports for further analysis
* Built **BEx queries** with **Restricted Key Figures, Calculated Key Figures, customer exit variables** and **text variables** which was utilized in **Xcelsius dashboards**
* Created attribute, analytical views on SAP HANA database and built IDT universes to satisfy business requirements
* Developed complex WebI reports with Revenue & Expenses data for various Cost Centers & Profit Centers using features like **Input Controls, Element Linking, opendoc hyperlinking**
* Developed **WebI reports** on **HANA analytical & calculation views** using **IDT**
* Designed and developed mobile enabled WebI Reports, Xcelsius and Design Studio dashboards which are accessed by the SAP BI Mobile app
* Built dashboards on Design Studio. Created custom defined CSS classes to satisfy the formatting requirements
* Enhanced user experience by using html 5 for mobile enabled webi reports & dashboard
* Worked on Performance Enhancements for Xcelsius Projects built by other teams, by looking into various aspects like Architecture, Design, Server Scalability and Environment Load Handling
* Created BI workspaces for different user groups like finance ,marketing .
* Provided extensive training for end users on WebI reporting, Mobile reporting, Dashboard and Explorer use
* Performed cross team SIT and UAT operations to give out the best output available

**Environment:** SAP BOBI 4.0 SP4/ SAP BOBI 4.1 SP1, Dashboard Design 4.0, Design Studio 1.1, Lumira 1.1, SAP BW, HANA database

**Talent Management:**

* Involved in gathering, analyzing, and documenting business requirements and data specifications for Business Objects Universes and Reports
* Understanding Functional and creating Technical documents for development of reports and dashboards
* Worked on Designing & Developing complex interactive Dashboard “HR Org View” on top of Webi reports built using BIWS. Dashboard has multiple levels of Org level hierarchy Drill Down & Drill UP. Both implemented using custom function in Webi and Excel..
* Created Dashboards for different departments based on User needs, using compatible Data Visualizations, Charts, Dynamic Selections & Visibility, Tabs, Alerts, User Selection Components, Multiple Drill Downs etc. in Xcelsius 2008 Sp4
* Worked on Data Modeling and Design as per the requirements for building the dashboards in MS Excel using Excel Functions like Index Match, Vlookup, choose, nested if conditions etc.,
* Trouble shooting dashboard issues using fiddler analysis & snapshots; improved performance using best practices as recommended SAP
* Design and Develop Complex Webi reports using **Variables, Drill Downs, Sections, breaks, and Merged Dimensions, Context Operators etc.,**
* Dynamic Reports Using additional 4.0 Features such as Content Linking , Element Linking, Folding etc
* Created **IDT universes** on multi providers to build info spaces for developing explorer views
* Created query definitions using the BEx Query Designer. Integrating key figures, characteristics, hierarchies, variables, properties, exceptions and conditions in the query definition Using Open Docs to connect SAP Dashboard with relevant Webi Reports and passing Parameters as required based on User Selection
* End user training on self-served WebI reporting and Xcelsius dashboard usage
* Performing Unit testing, supporting UAT and technical document preparation for post-production help

**Environment:** SAP BI BO 4.0, WebI, Rich Client, Explorer, LCM, Dashboard Designer, BI Launch Pad, SAP BW

**NADB Implementation:**

* Designed application by integrating product knowledge with users requirements.
* Created complex reports on WebI Rich client and WebI thin client using graphs, pie-charts and multiple data providers on daily, weekly, monthly and quarterly basis.
* Developed dashboards using Crystal Xcelsius using QaaWS ,Live Office and Spread sheet in a form of chart, dial, dynamic dashboards etc., refresh QaaWs every time when end user run the report, Use Triggers in M/D.
* Documented the design.
* Created templates for standardizing reports using company logo, header and footer settings, color coding, fonts and maintained consistency in viewing cross-tabs, blocks in report.
* Created user defined objects to manipulate the data by using database specific functions to fulfill the reporting requirements.
* Synchronized reports using data from multiple data providers in **Web Intelligence by merged dimensions functionality.**
* Extensively used **calculations, variables, filters, conditions, breaks, sorting, drill down, slice and dice, master/detail, cross tab, charts and alerts for creating Business Objects reports.**
* Created Users/Groups/Sub-Groups & assigned profiles there by restricting un-authenticated access to resources.
* Implemented row & object level security for different user group requirements.
* Developed issue log and prioritized the issues.
* Involved in resolving problems like ‘Chasm Trap’ and ‘Fan Trap’.
* Checked the integrity of the Universe schema for any errors in objects, joins, conditions, loops and cardinalities of the Universe.
* Enhanced reports which included standardizing format of report content, redefining query filters and prompts, linking summary report to a set of detail reports.
* Used @Functions such as @Prompt- to restrict return values and to enter prompt, @Select- to reuse the select statement of another object, @Variable- to restrict the currently logged user etc.
* Distributed documents inside and outside of the business using CMC and administration for scheduling the reports (like REP, PDF, EXCEL, TEXT and HTML format).

**Environment:** Windows XP, SAP Business Objects XI 3.0/3.1,BODS, Oracle 10g, Crystal Xcelsius 2008**.**

**PVR Implementation:**

* Performed Enterprise Administration for SAP Business Objects environment.
* Migrated all universes/reports from BO XI R2 to BO XI R3.1 using Import Wizard
* Migrated whole environment from XI R2 to XI R3.1 using Import Wizard
* Migrated XI R 3.1 Environment to XI R4.0 using upgrade Management tool.
* Used BI Launch Pad in XI R4.0 instead of Info View.
* Used UDT in XI R4.0 to create and modify Universes.
* Partnered in creation of BO XIR 3 users; user groups, folders, universe connections, securities, server and server groups in development environment.
* Created users, user groups, folders, universe connections, securities, server and server groups in development environment in 4.0.
* Assigned privileges to respective reports and universe to users and user groups.
* Converted Desktop Intelligence reports to Web Intelligence reports using report conversion tool
* Configured reports on user profiles to burst at different destinations like FTP Server in PDF, Excel formats.
* Configured and fixed all issues encountered by universes/reports during migration.
* Modified variables and formulas from partially converted rich client reports to Web Intelligence XIR3 for exact same results.
* Created Ad-hoc reports using Business Objects functionalities like Cross Tab, Master Detail, Slice & Dice and Drill Down, etc.
* Imported / Exported the Universes from / to repository to make them accessible to end-users.
* Defined Aliases and Contexts to resolve loops in the universes.
* Addressed Chasm trap and Fan Trap using Multiple SQL statement features of Universe Designer tool.
* Created and worked on Row level security, Object level security using Designer and various predefined security settings depending upon users and group needs using CMC.
* Created several users and groups inside the CMC and implemented the access levels as per the best practices.
* Responsible for documentation of various stages in development including design, debug, maintenance etc.
* Involved in trouble shooting of various reporting errors.
* Created Business Objects reports, ad hoc queries with constant interaction with the end users.
* Created new user with appropriate user groups and granting them access level.
* Dealt with the user issues on new environment with their access level as well as with their reports.

**Environment: Windows 2003/7, SAP Business Objects XI R2/3.1/4.0, SAP BW Oracle 10g/11g, Teradata.**