**Skill / Experience Summary:**

* 10+ Years of IT with complete SDLC experience from understanding the requirements to code Deployment in Production.
* Hands on Experience with **Big Data Integration - Cloudera** **Hadoop – HDFS, PIG, Hive, Sqoop and Zookeper.**
* Excellent understanding / knowledge of Hadoop architecture and HDFS.
* Rich experience in Data Architecture, Development and Integration using ETL and BI solutions with **Informatica Power Center 8.x/9x and Talend Big Data Enterprise Platform 5.4, Tableau 8.0, SSIS, SSRS and SSAS 2012.**
* Hands on Experience in creating **PIG Latin and HQL Scripts and** NoSQL Database **HBase.**
* Hands on Experience in creating **Map Reduce** Code using Java.
* Worked on loading Big Data using parallelization techniques into relational databases from Hive and viceversa.
* Engage with business stakeholders to understand business needs for DW/BI and Big Data solutions.
* Define detailed business requirements for reporting, dashboards, analytics and ad-hoc reporting using Tableau 8.0.
* Created Various Informatica and Talend Jobs using various Transformations and Components
* Can work on Java to Develop and Enhance Map Reduce Code.
* Developed jobs to Read/Write Data from/to Hadoop and SAS.
* Worked on Projects and various POCs to integrate Talend Big Data Enterprise with SAS and Hadoop.
* Design, develop, debug, and optimize ETL Mappings and Jobs for Data Extraction and Data Loading for Reporting and Analysis.
* Lead User Acceptance Testing, Data Validations, created Test Scripts and Participate in Project Communications, Release Planning and Change Management.
* Experience of working with **Star Schemas and Snowflake Schemas** and XML files, Flat files and various RDBMS such as Oracle 11g, Netezza Appliance 7.0, SQL Server 2005 and Teradata TD12
* Created Mapplets and Joblets for Data Standardization using Power Center designer to improve Data Quality and Data profiling tasks using Informatica Developer.
* Experience working with FSShell, Unix Shell Scripting including sed and awk programming.
* Optimized performance and Debugging of ETL Code and well Versed in Query tuning and optimization including Partitioning.
* Involved in Unit testing and smoke testing of ETL Code.
* Create or update functional and technical and design documents.
* Interacted directly with the client to know the business requirement and executed the same in the projects.
* Exposure to other tools: Toad, PL/SQL Developer, Autosys, Defect logging tools – HPQC, Jira.

**EDUCATION: Bachelor of Technology**

|  |  |
| --- | --- |
| Technology Experience | Big Data, HUE, Oracle, Netezza, Teradata, SQL, PL/SQL, Extract – Transform - Load (ETL), Oracle 9i/10g, SQL Server 2005, Unix - Shell Scripting, Reporting, Dashboards, Business Intelligence |
| Tools | Informatica Power Center 8x and 9x, Talend Big Data Enterprise Platform 5.4, Informatica Developer, Micrstrategy 9.3 , Autosys, HPQC, Toad, PL/SQL Developer, Aginity Workbench |
| Domain Experience | Telecom, Banking, Retail and Healthcare |
| RDBMS | Oracle 9i/10g/11g, Netezza 7.0, TeradataTD12 |

Experience Profile – Key Projects

|  |  |
| --- | --- |
| Project: Visa Inc | Duration: Apr 2013 - Current |
| **Customer name:** Visa Inc (US)  Location : Foster City, CA | |
| **Project Description:**  Visa Inc is a leading provider of Credit and Debit cards to various financial banks, Financial institutions, Merchants and Governments. They also provide pre-paid products and have offices all around the globe.  This is a ETL and Big Data Development and Integration project carried out with an ETL tool Talend Big Data Platform 5.4 and Cloudera Hadoop 1.2.1 HDFS, Map Reduce, Pig, Hive, HBase, Sqoop and Zookeeper. Developed Map Reduce Code in Java. Loading Data (>100 TB) from Hadoop into Hive and HBase using PIG, Hive and HBase Scripts. The data in Hive is then loaded into RDBMS using Talend. I have done various POCs for different projects and developed code from scratch in Production. I have read and write data to Hive and SAS using various Talend Big Data Platform.  **Tools:** Informatica 9.1, Talend Big Data Enterprise Platform 5.4, Cloudera Hadoop 1.2.1  **Responsibility**: Big Data/ Hadoop Developer | |
| **Role Description:**   * Involved in developing the Pig, Hive, HBase scripts. * Hands on Experience in creating Mapper and Reducer Code using Java Classes. * Worked on data analysis in HDFS using Hive and PIG jobs * Involved in creating tables, partitioning, bucketing of table. * Extracted Hive query output to local file system in text/csv format. * Developed the sqoop scripts in order to make the interaction between Pig and Db2 Database * Acessing Query editor to fire quieries and browse the File System, Job browser and Metastore Manager using HUE. * Created HBase tables to store variable data formats coming from different sorces. * Developed various jobs in Talend using Compoents like tMap, tJoin, tAggregateRow, tFilterRow, File related components, Database related, XML related. * Developing, Designing and Migrating the Ab Intio graphs to Talend Jobs using various Talend ETL and Big Data components. * Created custom components in components, worked with SAS related components such as tSasInput, tSasOutput and Hadoop related such as tHiveinput, tHiveoutput, tHiveconnection, tSqoopImport, tSqoopExport etc * Creating and Designed various joblets and Subjobs and Masterjobs. * Developed jobs for reading and writing/ Integration with big data Hadoop, Pig, Hive and Sqoop. * Worked on various databases such as DB2, Oracle, MySql and XML and Big Files. * Performing unit/integrated testing of ETL process. * Developing Informatica mapping to implement techniques like slowly changing dimension (type1, type2), slowly growing targets, and Simple pass through. * Created various Transformations like Lookup, Joiner, Aggregate, Expression, Filter, Router, and Update Strategy to implement the business logic. * Creating tasks/Workflows to perform the ETL execution. * Perform data validation and eliminate issues as a source. * Build Informatica ETL processes to load data through the Staging layer in to EDW * Creating HLD(High level design), TDD(Technical Design docs)and Mapping document as per business requirement. * Created various Transformations like Lookup, Joiner, Aggregate, Expression, Filter, Router, and Update Strategy to implement the business logic. * Creating tasks/Workflows to perform the ETL execution. * Performing unit/integrated testing of ETL process. * Daily communication with Scrum Calls, Weekly Status Calls and SME/Business Analysts via email, chat, teleconferences. * Internal Quality Auditing of the deliverables before delivery. | |

|  |  |
| --- | --- |
| Project: Aetna ETL **Development and Migration** | Duration: Dec 2011 – Mar 2013 |
| **Customer name:** Aetna Health Care (US)  Location : Downers Grove, IL | |
| **Project Description:**  Aetna is one of the largest health care companies in US with more than 5 million customers all around the country. This project contains development of data warehouse, data marts and ODS in Informatica Power Center 9.1 and database being Oracle and Netezza. They are migrating the code from Oracle to Netezza and therefore involves migration as well and its unit testing and performance comparison compared with the legacy system. Almost migrated more than 200 mappings including the data marts and ODS. It involves creation of reporting and dashboard using Tableau as the reporting tool.  **Tools:** Informatica 9.1, SSIS/SSAS/SSRS, Tableau 8.0  **Software:** SQL, PL/SQL  **RDBMS** : Oracle 10g, Netezza, SQL Server 2012  **Responsibility**: Senior BI Developer | |
| **Role Description:**   * Requirements Analysis – Architecture and Identify sources of data and document business rules. * Developing, Designing and Migrating the Informatica mapping to implement techniques like slowly changing dimension (type1, type2), slowly growing targets, and Simple pass through. * Created and Designed various Transformations like Lookup, Joiner, Aggregate, Expression, Filter, Router, and Update Strategy to implement the business logic. * Created Mapplets for Data Standardization using Power Center designer to improve Data Quality and Data profiling tasks usingInformatica Developer. * Manage the production and distribution of Dashboards and Reports including Tableau reports. * Perform in-depth analysis and insight into key online metrics; monitor trends, identify patterns, and provide recommendations for making informed business decisions * Creating and Designed tasks/Workflows to perform the ETL execution. * Working in Oracle as Source system to extract data and Netezza as a Target System to load data into it. * Performing unit/integrated testing of ETL process. * Convert existing processing of data in to Informatica mappings and move all tables into Landing/Staging area. * Perform data Extraction to move all data to Enterprise data ware house (EDW) * Perform data validation and eliminate issues as a source. * Build Informatica ETL processes to load data through the Staging layer in to EDW * Support down-stream users and migrate the down-stream users from Oracle to Netezza * Creating HLD(High level design), TDD(Technical Design docs)and Mapping document as per business requirement. * Developing Informatica mapping to implement techniques like slowly changing dimension (type1, type2), slowly growing targets, and Simple pass through. * Created various Transformations like Lookup, Joiner, Aggregate, Expression, Filter, Router, and Update Strategy to implement the business logic. * Creating tasks/Workflows to perform the ETL execution. * Performing unit/integrated testing of ETL process. * Daily communication with Scrum Calls, Weekly Status Calls and SME/Business Analysts via email, chat, teleconferences. * Internal Quality Auditing of the deliverables before delivery. | |

|  |  |
| --- | --- |
| Project: EBAY DW BI **Development** | Duration: Mar 2011 – Dec 2011 |
| **Customer name:** Ebay Enterprise (US)  Location : King of Prussia, PA | |
| **Project Description:**  Ebay Enterprise formerly known as GSI commerce is a second largest e-commerce company in the world having various types of products available for buying and selling online. Ebay enterprise has several clients for which they design and create Data Warehouse for respective web stores.  This Data Transformation Program (DTP) requires transformation of various webstores such as Spanx, Sony, ToysrUS, Steinmart and other store’s data from its various Source Systems like Oracle and FlatFiles according to defined business rules, and loading them into Netezza as the target system (DW). This data is then used for their reporting purpose This ETL process is carried out with the help of Informatica Power Center and Talend Studio.  **Tools:** Informatica 8.6/ 9.1, Talend Studio, SSIS/SSRS/SSAS, SQL Server Management Studio  **Responsibility**: DW BI Lead | |
| **Role Description:**   * Lead BI implementation project efforts to understand overall analytical needs of the e-Commerce organization for E-Bay Supply chain management. * Developed various jobs using Compoents like tMap, tJoin, tAggregateRow, tFilterRow, File related components, Database related, XML related. * Worked on Java Code at backend and to create custom components. * Creating and Designed various joblets and Subjobs and Masterjobs. * Perform in-depth analysis and insight into key online metrics; monitor trends, identify patterns, and provide recommendations for making informed business decisions. * Developing Informatica mapping to implement techniques like slowly changing dimension (type1, type2), slowly growing targets, and Simple pass through. * Creation and maintenance of logical and physical data models. * Created various Transformations like Lookup, Joiner, Aggregate, Expression, Filter, Router, and Update Strategy to implement the business logic. * Convert existing processing of data in to Informatica mappings and move all tables into Landing/Staging area. * Perform data Extraction to move all data to Enterprise data ware house (EDW) * Perform data validation and eliminate issues as a source. * Build Informatica ETL processes to load data through the Staging layer in to EDW * Creating HLD(High level design), TDD(Technical Design docs)and Mapping document as per business requirement. * Developing Informatica mapping to implement techniques like slowly changing dimension (type1, type2), slowly growing targets, and Simple pass through. * Created various Transformations like Lookup, Joiner, Aggregate, Expression, Filter, Router, and Update Strategy to implement the business logic. * Creating tasks/Workflows to perform the ETL execution. * Performing unit/integrated testing of ETL process. * Daily communication with Scrum Calls, Weekly Status Calls and SME/Business Analysts via email, chat, teleconferences. * Internal Quality Auditing of the deliverables before delivery. | |
| Project: **Data Warehouse Development for BT Customers** | Duration: Nov 2008 to Feb 2011 |
| **Customer name:** BT  Location: India, UK | |
| **Project Description:**    BT is a British [multinational](http://en.wikipedia.org/wiki/Multinational_corporation) telecommunications services company headquartered in London, United Kingdom. It is one of the largest telecommunications services companies in the world and has operations in over 170 countries. Through its [BT Global Services](http://en.wikipedia.org/wiki/BT_Global_Services) division it is a major supplier of telecoms services to corporate and government customers worldwide. Its [BT Retail](http://en.wikipedia.org/wiki/BT_Retail) division is a leading supplier of telephony, [broadband](http://en.wikipedia.org/wiki/Broadband_Internet_access) and [subscription television](http://en.wikipedia.org/wiki/Subscription_television) services in the UK, with over 18 million customers.  This Assignment requires transformation of BT 21CN Customer data from its various Source Systems according to various business rules, and loading them in target system (DW). This data is then used by various businesses for reporting purpose This ETL process is carried out with the help of Informatica Power Center.  Tools: Informatica 8.6  Responsibility: ETL Developer | |
| **Role Description:**   * Reviewing requirements with business and both onsite/offshore management team. * Creating ETL design docs and Mapping document as per business requirement. * Experience of working with Teradata SQL and PL/SQL Procedures, Packages, Functions and Triggers. * Experience working with Unix Shell Scripting. * Optimized performance and Debugging of ETL and Teradata Code. * Experience working with TD Utilities such as FLoad, Mload, FExport, TPump etc. * Created various Transformations like Lookup, Joiner, Aggregate, Expression, Filter, Router, and Update Strategy. * Assigning the work from functional & technical perspectives. * Implemented techniques like slowly changing dimension (type1, type2, and type3) and Simple pass through mapping with ETL tool Informatica. * Creation of various tasks, sessions, worklet and workflows and their scheduling using Workflow Manager. * Interacting with the offshore and Onsite management. * Troubleshooting and Rescheduling in case of issues and problems with the Workflows * Managing, coordinating, planning and executing project. * Ensuring proper utilization of resources. * Handling team size of 4 Developers. | |
| Project: **Customer Data Warehouse Development** | Duration: Apr 2007 to Oct 2008 |
| **Customer name:** Qualcomm (US)  Location: Santa Clara, CA, US | |
| Project Description:  Qualcomm is world-leading provider of wireless technology and services. Qualcomm is a global company, a firm with many facets, with each business division changing the way and work through its own unique contributions. But no matter what each Qualcomm business does, all are united by a single, driving passion to continue to deliver the world’s most innovative wireless solutions  This Assignment requires transformation of Qualcomm Customer data from its various Source Systems according to various business rules, and loading them in target system (DW). This data is then used by various businesses for reporting purpose This ETL process is carried out with the help of Informatica Power Center.  Tools: Informatica Power Center 8.1  Software: SQL  **RDBMS**: Oracle 9i, SQL Server 2005  **Responsibility**: Data warehouse Developer | |
| **Role Description:**   * Reviewing requirements with business and both onsite/offshore management team. * Creating ETL design docs and Mapping document as per business requirement. * Experience of working with Oracle SQL and PL/SQL Procedures, Packages, Functions and Triggers. * Experience working with Unix Shell Scripting. * Optimized performance and Debugging of ETL and PL/SQL Code. * Created various Transformations like Lookup, Joiner, Aggregate, Expression, Filter, Router, and Update Strategy. * Assigning the work from functional & technical perspectives. * Implemented techniques like slowly changing dimension (type1, type2, and type3) and Simple pass through mapping with ETL tool Informatica. * Creation of various tasks, sessions, worklet and workflows and their scheduling using Workflow Manager. * Interacting with the offshore and Onsite management. * Troubleshooting and Rescheduling in case of issues and problems with the Workflows * Managing, coordinating, planning and executing project. * Ensuring proper utilization of resources. | |

|  |  |
| --- | --- |
| Project: **Data warehouse of BSS Customers** | Duration: Jun 2003 to Mar 2007 |
| **Customer name:** BSNL  Location: India | |
| **Project Description:**  Bharat Sanchar Nigam Limited (abbreviated BSNL) is an Indian [state-owned](http://en.wikipedia.org/wiki/State-owned) [telecommunications](http://en.wikipedia.org/wiki/Telecommunications) company headquartered in [New Delhi](http://en.wikipedia.org/wiki/New_Delhi), [India](http://en.wikipedia.org/wiki/India). It is the largest provider of [fixed telephony](http://en.wikipedia.org/wiki/Telephony) and fourth largest [mobile telephony](http://en.wikipedia.org/wiki/Mobile_telephony) provider in India, and is also a provider of [broadband](http://en.wikipedia.org/wiki/Broadband_Internet_access) services.  It is India's oldest and largest communication service provider (CSP). It had a customer base of 95 million as of June 2011.  The purpose of the project was to develop and maintain their GSM and CDMA Network BSS (Business Support System) Data warehouse using PL/SQL. This comprises of operations and maintenance of BT BSS customers 24 x 7 including their Billing, Mediation, Payment and Revenue systems.  Tools: IRB 3.0 ( Billing Application), Kenan Fx  Software: SQL, PL/SQL  **RDBMS**: Oracle 9i  **Responsibility**: ETL Developer | |
| **Role Description:**   * Creation and Maintenance of BSS data warehouse using PL/SQL. * Experience of working with SQL and PL/SQL Procedures, Packages, Functions and Triggers. * Designed and created database tables, views and stored procedures and Built and maintained SQL scripts, indexes, reports and queries for data analysis and extraction. * Created different Cursors and Reference Cursors in various Procedures and Packages. * Extensive experience in creating and executing SQL queries and Query tuning and optimization * Created test scripts for application module and performed Unit Testing of application. * Interacted with the end users to define and document business requirements as the foundation for database design and analysis. * Mentoring the Team in troubleshooting routine issues and problems and also with the new CRs and Developments. * Co-ordination with the Client in order to understand the business needs. | |

Training / Continuing Education Summary

|  |  |  |
| --- | --- | --- |
| **Program or Course** | **Coverage** | **Year** |
| **Oracle Certified Associate - Developer** | **Oracle 9i - SQL and PL/SQL** | **2010** |
| **Oracle Certified Associate - DBA** | **Oracle 9i - Fundamentals of Database Administration concepts** | **2010** |
| **Informatica Power Center** | **Informatica setup, mappings, case studies, Designer, Workflow manager and Monitor** | **2007** |
| **Netezza** | **Introduction to Netezza Architecture** | **2012** |
| **International Standard Training Quality Board (ISTQB)** | **All Testing Methodologies and concepts** | **2009** |
| **Business Objects** | **Introduction to reporting processes and concepts 6.0** | **2011** |
| **OBIEE** | **Introduction to OLAP concepts and reporting using OBIEE 10.1.3** |  |
| HP- Quality Center | HPQC overview and concepts | 2008 |
| Unix/ Shell Scripting | Unix commands and Shell script writing using sed and awk commands | 2006 |