

Contact

ameya00@yahoo.com

www.linkedin.com/in/katreameya
(LinkedIn)

Top Skills

Big Data Analytics

Amazon Web Services (AWS)

Cloud Computing

Languages

English (Full Professional)

Hindi (Native or Bilingual)

Marathi (Native or Bilingual)

Certifications

Teradata Certified Implementation
Specialist V2R5

Teradata Certified Application
Developer V2R5

Teradata Certified Design Architect
V2R5

Teradata Certified Administrator
V2R5

G-NIIT (Systems Management/e-
Commerce)

Honors-Awards

Top Performer *Extra Mile* - 2013

Top Performer *Extra Mile* - 2012

Teradata Excellence Award - 2007

Ameya Katre

☁️# Principal Data Engineer
San Francisco Bay Area

Summary

- Teradata Certified Master (V2R5)
- Diploma from NIIT (G-NIIT) in Systems Management/e-Commerce Technology

Design and Development of technical specifications using Hadoop technology

Apache PIG scripts to process the HDFS data

Hive tables to store the processed results in a tabular format

Hadoop scripts which take the input from HDFS and load the data into Hive

Load Hive external output processed data using Sqoop

Configuring, maintaining and monitoring Apache Kafka and Zookeeper for critical data flow pipelines for various applications

Using Logstash (ELK Stack) for data parsing and ingestion

Using Apache Spark for algorithmic data analysis and prediction

Data Visualization and BI implementation using Tableau, Splunk and Business Objects

Project Management & Delivery Management experience

Data warehousing solutions using Teradata methodologies and

Tools which involves Analysis, Design, Development, Testing and

Reporting using various Teradata TTU and also other third-party ETL tools

Team Leadership which includes managing onsite-offshore global delivery model ensuring the ETL/BI project work is on track as per the schedule

Data Warehouse Architecture and Data Modeling (E-R Modeling/ Dimensional Modeling)

Experience in consolidation of the project metrics, ensuring QA standards and reviewing technical documents for correctness & accuracy

Successfully leading an Onsite-Offshore Delivery Model

Excellent leadership qualities, communication skills & team player

Specialties:

Hadoop / HDFS
Apache Pig / Hive / Oozie
Apache Kafka / Zookeeper
Amazon Web Services (Redshift, EC2, S3, EMR, Kinesis)
Google Cloud Platform (BigQuery, Compute Engine, Dataproc,
Dataflow, Kubernetes Engine, VPC)
ELK (Logstash)
Tableau 9.2
Teradata 12.0/13.10
Linux/Unix
Data Analysis/Data Profiling
Data Modeling (ERWin)
Query Optimization
Complex Ad-hoc Reporting
ETL Tools (Wherescape RED, Informatica, Datastage, Abinitio)
Teradata Viewpoint
Teradata Tools and Utilities / BTEQ
IBM Tivoli Workload Scheduler

Experience

Dish Network

Principal Data Engineer / Scientist

March 2015 - Present

San Francisco Bay Area

Google Cloud Platform - BigQuery, Compute Engine, Cloud Storage, ML
Engine, Cloud Functions, Dataflow, etc.

Amazon Web Services - Redshift, RDS, S3, EC2, EMR, Kinesis Firehose,
Lambda, AWS Glue, AWS Athena, Redshift Spectrum, CloudTrail,
CloudWatch, SNS, SQS, etc.

Apache Kafka

Python / R / Shell scripting

NGINX

ElasticSearch/Logstash/Kibana (ELK stack)

Tableau

LAMP stack

Apache Spark

Splunk

Hadoop/HDFS

Apache Pig
Apache Hive
Sqoop
Couchbase
Machine Learning

- Lead the creation of the solution architectures.
- Involved in Design and Development of technical specifications using open source technologies.
- Create Real-time data streaming application pipelines using ELK stack for log processing.
- Work extensively on Amazon Web Services like EC2, EMR, S3, Redshift, Kinesis, Lambda, SNS, SQS, etc.
- Build real-time event-driven pipelines through AWS Lambda using Apache Spark / AWS Glue to store final processed data in parquet format in S3 and index them in Glue Catalog/Hive Metastore.
- Involved in recommending and prioritizing various project implementation.
- Use GCP BigQuery for query caching and faster retrieval.
- Developed different strategic vision and roadmap for the organization.
- Work closely with the engineering team to create data visualization using Tableau and Splunk to identify interesting user behavior and trend patterns for actionable insights.
- Create external tables in Hive and S3 (Redshift Spectrum).
- Working on creating custom training models using algorithms to develop various solutions using Collaborative Filtering and Regression Analysis in Apache Spark.
- Involved in moving all data files generated from various sources to HDFS for further batch processing.
- Create Apache PIG scripts to process the HDFS data and Hive tables to store the processed results in a tabular format.
- Utilize understanding of industry information management trends to innovate and provide new project/product ideas within the investment lifecycle.

eBay Inc
Sr. Data Engineer
September 2014 - March 2015 (7 months)

- Responsible for customer satisfaction and loyalty (NPS) programs.
- Building customer specific surveys to get feedback on newly launched products and/or key features for buyers and sellers.

- Focused on using both qualitative and quantitative research methods to gather customer feedback used to make eBay a better place for buyers and sellers.
- Designed and Developed data warehouse data model and inbound/outbound interfaces to Teradata.
- Lead the creation of the solution architectures.
- Involved in Design and Development of technical specifications using Hadoop technology.
- Create Apache PIG scripts to process the HDFS data and Hive tables to store the processed results in a tabular format.
- Create external tables in Hive.
- Production support for Hadoop scripts which take the input from HDFS and load the data into Hive.
- Developed MapReduce application using Hadoop, MapReduce programming and Hbase.
- Used Sqoop to export the Hive external output processed data into data-warehouse.
- Participated and recommended prioritization through the project implementation.
- Developed different strategic vision and roadmap for the organization.
- Utilize understanding of industry information management trends to innovate and provide new project/product ideas within the investment lifecycle.
- Code Reviews and ensuring Teradata best practices are being followed.
- Justify the technical design and approach to technical architects.
- Analysis and design of the data stored by information systems, concentrating on entities, their attributes, and their relationships.

WellPoint

Teradata ETL Architect

June 2014 - September 2014 (4 months)

Business Requirements

Data Warehouse Architecture

Teradata ETL (using Informatica 9.1 and BTEQ and AIX)

TekLink International Inc

Teradata ETL Architect

May 2013 - September 2014 (1 year 5 months)

Extensive experience with Data Analytics tools for Reporting and Data Visualization using Tableau 8.0 for massive amounts of data.

Experience with distributed analytic processing tools using Hadoop, HDFS, Hive, Pig etc.

Working on multiple projects in parallel which include development, testing and production support.

Business Requirements

Data Warehouse Architecture

Data Modeling

Teradata ETL

Wherescape RED

Database Management

Altria

Teradata ETL Architect

June 2008 - June 2014 (6 years 1 month)

- Responsible for Project Management and Delivery Management which includes –
 - o creating ETL roadmap and planning for Teradata Practice Team
 - o managing customer expectations, issue resolution and escalation management
 - o preparing monthly SLA reports and weekly reports for client
 - o Onsite-Offshore co-ordination
 - o supporting new project implementation including development, execution, testing and data consolidation of ETL scripts
 - o automation of EDW batch processes
 - o reviewing change requests and preparing impact assessments
 - o ETL failure analysis
 - o performing break-fixes and working on new enhancements
 - o tracking service request tickets (internal and business related)
 - o reviewing, updating and maintaining schedule books
 - o implementing Teradata ETL process with end-to-end solutions including regular audits and deviation analysis
 - o providing best practices for the client based-on and aligned-with the client's business objectives
 - o assisting clients with development and management of standards-based enterprise strategic technical Data Warehouse
- Team Leadership which includes managing onsite-offshore global delivery model ensuring the ETL/BI project work is on track as per the schedule.
- Perform requirement gathering and analysis to convert business requirements into functional and technical specifications.

- Data Warehouse Architecture and Database Management which includes DBA activities such as –
 - o maintaining the Teradata Data Warehouse
 - o Performance Tuning on the database
 - o Query Optimization
 - o User access management
 - o Data Analysis and Correction, etc.
- Perform Data Modeling which includes map-gap analysis, blueprinting, “proof-of-concept” models, data profiling, data validation and translating business requirements into LDM and PDM with optimization considerations.
- Perform Business Reporting which includes –
 - o ad-hoc queries and reporting
 - o creating, maintaining and optimizing pre-defined business report for tactical and strategic decision making

TekLink International Inc

Teradata - Offshore Practice Lead

September 2011 - May 2013 (1 year 9 months)

Deputed overseas to establish and grow the platform/infrastructure for the Teradata Offshore Practice Team (Development and Application Support Projects)

Project Management/Delivery Management

Team Leadership

Data Warehouse Architecture

Data Modeling

Process and Performance Improvements

Mentoring/Training

Teradata Corporation

Sr. Technical Consultant

July 2006 - September 2011 (5 years 3 months)

Client : Altria Group Inc.

Industry : Retail and Manufacturing

Project Type : Data Warehouse Design, Development, Implementation and Testing.

Environment : Teradata 12.0 on Windows Server 2003 / Unix & Business Objects XI

Role : Data Analyst, ETL Technical Lead, Reporting, Database Designing.

Responsibilities and Contribution:

- Analysis and Business Requirement gathering.
- System Design, Mapping of ETL Architecture.
- Ad-hoc extracts using complex queries for data analysis and on-demand business reporting and analysis.
- Data Profiling and Data Analysis
- Core ELT Development on Teradata (TPT/BTEQ)
- Business Reporting using Teradata scripts.
- Batch scripts in DOS
- Migration to UNIX environment using Shell scripts.
- Handling Onsite-Offshore Model.
- Experience in consolidation of the project metrics, ensuring QA standards and reviewing of the technical documents for completeness, correctness and accuracy.
- Preparing and maintaining project documents (Development documents, Allocation documents, Incident Reporting documents, Test documents)

NCR Corporation

EDW ETL Developer (Intern)

July 2005 - April 2006 (10 months)

Teradata India Pvt. Ltd. was a sub-division of NCR Corporation till Sep. 2007.

• Phase I

- o In collaboration with other team members, analyze raw data to ensure the data will meet the business need.
- o Developed complex SQL and Korn Shell scripts to load data into the data warehouse.
- o Analyzed the data mapping and developed test scenarios to verify the data migration from different sources.
- o Perform problem analysis and problem solution in a technical environment.
- o Coded test SQLs and analyzed results.
- o Interface with all levels and across organizational lines.

• Phase II

- o Participate in the project life cycle, from analysis to production implementation, with emphasis on data analysis, unit testing, system testing, and auditable documentation of results.

- o Reviewed and improved the design of Extract and Load specifications
- o Developed Korn Shell scripts based on Load and Extract Specifications
- o Accomplished data movement process that load data from existing source system into Teradata by the development of Korn Shell scripts, using Teradata SQL and utilities such as BTEQ, FASTLOAD, FASTEXPORT, MULTILoad, Teradata SQL Assistant and Teradata Administrator.
- o Use UNIX commands and utilities to monitor server side activities and performance.
- o Execute systems and data quality testing to ensure all development deliverables are production ready.

Education

University of Mumbai

B.E., Computer Science

Teradata

Teradata Certified Master V2R5, Teradata Certification · (2006 - 2010)

NIIT

G-NIIT, eCommerce / Systems Management · (2001 - 2010)