|  |  |  |  |
| --- | --- | --- | --- |
| **GUPTA GUNDLAPALLI** | | | |
| A highly accomplished IT consultant with 19+ years across multiple domains and roles including Apache Hadoop Data Analyst, AWS Cloud Consultant, ETL Developer, and Integrated BI/ETL Test Lead, Analyst, BI Manager, Data profiler and Project Lead, demonstrating significant successes, achievements and experience across a range of leading organizations. A respected figure operating within demanding environments; service driven with the ability to initiate change and optimizing opportunities for collaboration.  **EXPERIENCE**  2020 January – Present JP Morgan Chase.  **Vice President- Commercial Bank – Data Engineer**  **Client:** JP Morgan Chase, TX, Plano.   * Re-design /Modernize EDH & MDM on premise CDH platform to hybrid cloud Architecture. * Currently **Application owner** for EDH application, governing the next generation modernizing hybrid platform. * Successfully completed a **POC in evaluating Kubernetes deployment of** Spark +Hadoop deployment. * Successfully completed a **POC in evaluating AWS by adopting ECDP frame work**. * Enhancing the existing ingestion framework (Python Ingestion Frame work) suite to cloud architecture. * Developing/Design of Kubernetes deployment in GKS. * Building data censes and Data lineage for approx. 200 + source in Data lake on CDH. * Building Micro services to enable ingestion frame work.     2019 March– 2020 January Capital One  **Data Management - Data Warehouse Developer/Architect**  **Client:** Capital one VA  Modernized cloud based solution for IRIS (Interactive Research Information System) i.e. Desktop application used by agents (Phone, Chat, Back Office, etc) to service accounts . The Three main groups of data created as a result of interactions in IRIS: System Events, Workflows and Business Events. Bizevents data is modernized using pyspark using AWS EMR cloud infrastructure .     * Lead the application development team to deliver Cloud Modernization of IRIS Card Bizevents including Cloud and DevOps * Ensure a robust view of the overall business performance of Bizevents Data stored in warehouse * Perform data exploration/analysis identifying trends and insights * Experienced in handling large datasets using **Partitions, Spark** in Memory capabilities, **Broadcasts in Spark**, Effective & efficient Joins, Transformations and other during ingestion process itself. * Designed, developed and did maintenance of data integration programs in a Hadoop and RDBMS environment with both traditional and non-traditional source systems as well as RDBMS and NoSQL data stores for data access and analysis * Extensively used **AWS services like ECS, EC2, Lambda, EBS, S3, ELB, RDS, IAM, Autoscaling, CloudWatch, CloudTrail.** * Strong experience in writing applications using python using different libraries like **Pandas**, **NumPy**, SciPy, Matpotlib etc * Design **EMR cluster “Steps”** architecture design to orchestrate Data pipeline . * Implemented Trex (**Technical Recovery Exercise**) for Bizevents EC2 instances , EMR and S3 replication policies. * Implemented the Regional isolation of EAST and WEST successfully. * Using Jenkins successfully maintained IRIS bizevents API to Auto Rehydrate . * Using Different **APIs data is fetched to raw bucket (Data Lake)** **in AWS S3** and latter using Spark Data frames and Spark SQL data is transformed as required and stored into S3 bucket. * Key participant in all phases of software development life cycle with Analysis, Design, Development, Integration, Implementation, Debugging, and Testing of Software Applications in client server environment, Object Oriented Experience in using various IDEs **Eclipse, Atom, Jupiter and repositories Git**   2012– March 2019.  **Big Data Analyst, Ab initio ETL Designer , Spark Developer,Data Management, Business Intelligence (BI), Project Management, ETL Business Requirement Analyst,**  **Client:** SunTrust Richmond VA   * Developed ETL Generic “TEST HARNESS” Suite * Analyzed core requirements and key features of projects which involved three work streams: Data Ingestion, Data Distribution and Data Lake Analytical Processing with a primary object to build Enterprise Data Lake for Business Analytics, Trending Reports and Fraud Analytics * Good understanding of NoSQL databases and hands on work experience in writing applications on NoSQL databases like HBase, and **MongoDB** . * Implemented schema extraction for **Parquet and Avro file** Formats in Hive * Expertise with the tools in Hadoop Ecosystem including  **Hive, HDFS, Sqoop,, Spark**,. * Experienced in working with **Amazon Web Services (AWS) using EC2** for computing and **S3** as storage mechanism. * Used **Spark API over Cloudera Hadoop YARN** to perform analytics on data in **Hive partitioned and bucketed tables.**   **Environment:** Ab Initio GDE 3.2.5, Co-op 3.3 (on UNIX and on Windows), Hadoop Stack cdh5.9.0,Hbase 1.2.0  **Computational Frame work:** Spark 1.6.0, Apache Hadoop (HDFS 2.6.0 + Yarn)  **Diff File formats:** PARQUET DATA FORMAT (2.1.0), PARQUET(1.5.0),AVRO DATA FORMAT(1.7.6)  2009 – 2012 COGNIZANT TECHNOLOGIES  **ETL Developer**  **Client:** Capital One, Richmond, VA   * Instrumental in the development of a data warehouse with the aim of ensuring adherence of Basel II and FAS114 regulatory requirements for the Commercial Bank line of business * Tasked with automating the business processes in the Special Assets division of Capital One to meet the objective of building a data storage for automating the data capture for calculating the Loss Given Default (LGD)   **Projects completed:**  **Chevy chase conversion, ING conversion, CDI,HSBC conversion, Charger Project, Charger Finale, E-Time, Toro Recap, US Cards 3rd Level Support**   * Managed the insights of managers, directors and stakeholders to ensure all parties agreed with the current scope and roadmap of projects * Playing a cross functional role of ETL Designer, Lead Business analyst, Technical Lead * Implemented Capital One Specific ETL standards * Loaded 500 GB history data into Teradata Database using mload from text files   **Environment:** Ab Initio GDE 1.15, Co-op 2.15 (on UNIX and on Windows), Teradata 12.0, Microsoft SQL Server 2008 R2 (RTM), GNU/Linux, Windows XP, HP Service Manager, Web EME.  2007 – 2009 MPHASIS Internal Finance Corporation WBG  **BI Manager**   * PMO and Offshore subject matter expert (SME) support; utilized expertise to setup onshore and offshore delivery teams * Ensured standardization and supported the transition from a fully onsite model to a hybrid onshore by embedding processes, guidelines and training tools * Demonstrated influencing and communication skills by managing executive level discussions to establish an onsite-offshore execution framework * IT Process redesign and implementation (on architecture artefacts’, project engagements, design reviews and approvals) plans, allocation of available resources. * Managed multiple engagements at a given point of time * Supervised offshore vendor development activities for various implementation | **CONNECT**  202-468-6280  c0aaedf014aa40f0b5acc93d8512b409  Gupta.gundlapalli@gmail.com  [https://www.linkedin.com/in/radhakrishna-gupta-gundlapalli-2b743b6](https://urldefense.proofpoint.com/v2/url?u=https-3A__www.linkedin.com_in_radhakrishna-2Dgupta-2Dgundlapalli-2D2b743b6&d=DwMBaQ&c=pLULRYW__RtkwsQUPxJVDGboCTdgji3AcHNJU0BpTJE&r=5WY_Nit5r-IScEfdCKVA26yHhAx-aM3WLmlAP-TMOIU&m=z8WD6xzf8-p9lutLxgQoeYeSGTCS08aObWfa-OoFt4o&s=El31zMHv_5GdzqVmyFtOopYr0myZZ9T60JOoSCowa6k&e=" \t "_blank)  **EXPERTISE**  **Hadoop Stack:** cdh5.9.0, Hbase 1.2.0  Computational Frame work, Spark 1.6.0 (2.3.1 new Version), Apache Hadoop (HDFS 2.6.0 + Yarn)  **Diff File formats:** PARQUET DATA FORMAT (2.1.0), PARQUET(1.5.0),AVRO DATA FORMAT(1.7.6)  **Tools:** Ab initio 3.2.5 Informatica 9+, Talend 5.x (Projects, Visio, SharePoint, Office), Primavera, SQL-Assistant, Toad, BTEQ  **Languages:** SAS 9.4, SQL, PL/SQL, Shell Scripting, C++, Hive 1.1.0, Scala 2.10.5, pyspark 2.3.1  **Databases:** ORACLE, DB2, Mongo DB,TERADATA, MS SQL Server, DB2, MS-Access.  **PROJECT MANAGEMENT**   * Project rollout * Practitioner of leading agile methods * Excellent understanding of Scrum framework * Product roadmap development * Program release planning and management * Structured analysis and design * Budget management and forecasting * Stakeholder and Third-party vendor management * Contract management * Business case creation and benefits management * Scope management * Resource management and forecasting * Risk assessment * Regulatory - Governance/Compliance   **PEOPLE**   * Motivational Leadership * Leadership Development * People Management * Resource Management * Succession Planning   **TECHNICAL SKILLS**   * Cloudera Developer * HiveQL to SparkQL * Sqoop to import data into HDFS from RDBMS and vice-versa * Data models * Data flow diagrams * Data dictionary/glossaries * Source data analysis * Data modeling * Data quality analysis * QA testing * Target mapping * ETL and data cleansing design * Normalization * Data patterns * ETL data flows   **EDUCATION**  Master Computer Science, VTU University, India 2002.  **TRAINING & CERTIFICATIONS**  **AWS -Solution Architect Associate.**  **CKAD- Certified Kubernetes Application Developer Training – CNCD**  **Hadoop Developer CDH04**  Databricks Certified Developer - Training  Teradata VR12.5  Data Analyst Training  **ADDITIONAL EXPERIENCE**  2003 – 2007 Wipro Technologies  **ETL Developer**  2002 – 2003 Tata Consultancy Services  **SQL Developer** |