PARUL SHARMA

Datawarehouse & Business Intelligence

Cupertino, CA 95014 | 669.238.5613

parul0609@gmail.com | linkedin.com/in/parul0609 | github.com/parul0609

CAREER SUMMARY

Technical Manager/ Datawarehouse & ETL Architect

- With over 12 years of experience specialising in Enterprise Data Architecture, Datawarehouse/Business Intelligence, Data Modernization, Master Data Management, ETL & Data Visualization/ Reporting Architecture for OLTP and OLAP systems.

- Design and Developed end-to-end datawarehouse solutions for 45+ data models supporting 300 TB of critical financial KPIs data.

- Extensive experience in performance tuning of large data-intensive applications by leveraging multi-core server capability for parallel computing, database partitioning and SQL optimization.

- Experience of migrating 30+ applications, 50 PB of data to Cloud Based Datawarehouse Solutions.

- Proficient in working with RDBMS, NoSql DBs, SQL, ETL ,Big Data tools.

- Well recognized for thought leadership, excellent communication and managing complex project deliverables while collaborating closely with stakeholders across multiple business units and geographies.

SKILLS

- Datawarehouse/BI: MSBI Stack (SSIS, SSAS, SSRS),PowerBI, SQL
- RDBMS: MS SQL Server, MySQL, PostgreSQL
- ETL: SSIS, Pentaho, Datastage, Informatica
- Cloud Skills:
    - Google Cloud Platform (BigQuery, BigTable, Cloud Sql, Cloud Functions, Dataproc, Cloud Shell)
    - Microsoft Azure (Azure Data Factory, SQL database)
- Big Data Tool: Hadoop, Hive ,HBase, Spark, Oozie
- Languages: Python , C#
- Data Modelling: Microsoft Visio
- Version Control Tools: TFS, SVN, GIT, GitHub
- Agile: Scrum, Kanban

- Project Management Tools: Jira, Confluence, Sharepoint, Google Sheets, Excel

PROFESSIONAL EXPERIENCE

<u>Datametica</u>   2022 - Present

Technical Lead

Leading development for a large-scale data modernization to cloud platform that is expected to save more than 25% engineering efforts globally and save significant operational cost for a prestigious client.

- Large scale modernization and optimization of 50 PB data, 70K+ tables and 30+ project spaces by upgrading and automating client's data grid from Hadoop cluster to DataProc and Big Query.
- Closely collaborating with multiple teams including client application owners , offshore, downstream teams to have seamless integration of existing processes with the new upgraded platform.
- Developing metadata driven ingestion framework and workflows using Python,Spark and Oozie to support various forms of files (ORC, AVRO , Parquet) and extracting partitioned and non partitioned data from Hive.
- Working closely with Google to evaluate solutions to efficiently implement workflows using the right cloud services.

<u>TechM</u> (client: Thomson Reuters)   2017 – 2020

Technical Lead Manager

Led development of a multi-terabyte financial datawarehouse for a quant product.

- 300+ clients, 290 TB data storage, top product in quant space.
- Developed 45+ ETL/Data pipelines as deployed solutions and/or on cloud.
- Designed and managed dashboards for 10+ data modules in Power BI based on financial KPIs generated from the datawarehouse to be accessed by internal stakeholders to make key decisions on product offering.
- Achieved 18x performance improvement for an ETL pipeline (8 billion+ rows in one table) by tuning SQL queries, parallel computing and database partitioning.
- Worked with heterogeneous data sources including text, XML,TIBCO/MSMQ queues, databases etc.
- Scrum Master for planning sprints, stories and conducting daily stand ups.
- Received 9/10 client satisfaction rating consistently.

Thomson Reuters   2008 – 2017

Associate Development Manager/Technical Lead/ SSE

Owned Datawarehouse Architecture, ETL, Data Analysis & Reporting framework.

- Developed proprietary in-house reusable software infrastructure to download data, execute, monitor and deliver pluggable ETL feeds output to clients.
- Introduced and developed ingestion framework in new ETL tool Pentaho.
- Retired and migrated 7 legacy ingestion pipelines from C++ to SSIS and SQL.
- Successful transition of work from contractors that saved ~USD 3.3 million.
- Partnered with Microsoft, HP, Ab Initio, Informatica, Datastage for POCs to identify best software and hardware solutions to improve performance and scalability of ETL/Data pipelines.

Collabera   2007 – 2008

Database Engineer

Worked for one of the largest realtor companies.

- Successfully completed a migration project involving requirement gathering, schema analysis, development and optimization of complex stored procedures and ETL.
- Worked on multiple enhancements improving coding practices like changing dynamic queries to static, making use of the latest SQL server features like CTE and Ranking functions.
- Efficiently managed several maintenance tickets involving analysis, preparing approach, developing solutions and working with QA to scope out all possible test scenarios.

EDUCATION

Master of Computer Applications (MCA)
GJU University, India
Bachelor of Information Science (BIS) (*Computer Science*)
MDU University, India

ACHIEVEMENTS

- Multiple Individual Performance Awards
- Rewards & Recognition Team Award
- Member of ITrain Programme and Center of Excellence(COE) teams
- Tech Unconference Annual event Co-Speaker