Siva Kumar Kandivalasa

New Jersey, USA • (908) 210-3790

skandivalasa24@gmail.com• linkedin.com/in/siva-kandivalasa-80143777

SUMMARY

Highly proficient, versatile and resolution-focused Data Architect/Data Engineer/Cloud Architect offering expertise in data warehousing architecture, data engineering, data modeling, and project management seeking a challenging role to contribute my expertise and skills to develop scalable and reliable data warehouses.

- **16+** Years of Professional IT experience in Data warehousing, Data Lakes, ETL, ELT, Data modeling, Data Analysis, Business Intelligence(BI), Cloud, Database, and Project management.
- Seasoned expertise in Data lake, Data warehouse and Lakehouse architectures
- Proficient in **Snowflake, Amazon Web Services(AWS) and IDMC** with extensive experience in designing, building, and optimizing data warehousing solutions using cloud-based technologies
- Expertise in Snowflake - Architecture, designing data lake architectures, data modeling, implementing ELT solutions using Snowflake SQL, complex stored procedures and standard DWH and ETL concepts
- Having good abilities in development and support for the successful execution of projects using industry-leading BI tools and technologies.
- Designed and Implemented data warehousing projects in **Banking, Insurance, Telecom, Life Sciences, Retail, and Health Care** domains at various client locations across the Globe.
- Deep understanding of relational as well as NoSQL data stores, methods, and approaches (star and snowflake, dimensional modeling)
- Experienced in integrating with external systems such as **Snowflake, AWS, SAP (BW/HANA/ABAP/ IDoc/ BAPI), Structured and Semi-structured data, Salesforce, WebServices using power exchange**
- Certified **Snowflake Advanced Architect,** well-versed in advanced concepts including warehouse, query and cost optimizations, data engineering, shares, replications, policies, masking etc.
- **3x AWS Certified** and exposure to various AWS services including **IAM, S3, EC2, EMR, Auto Scaling, Elastic Load Balancing(ELB), VPC, Cloud Formation, CloudFront, DynamoDB, RedShift, Lambda, and CI/CD**.
- Experience in using Serverless services including **S3, Lambda, DynamoDB, API Gateway, StepFunctions, messaging services SQS, and SNS**
- Proficiency in **Python scripting**, including the use of control flow, loops, functions, and data structures, as well as libraries such as Pandas, CSV, NumPy, and Matplotlib.
- Knowledge of Hadoop technologies including HDFS, MapReduce, Spark, and HIVE
- Expertise in designing and developing OLTP and OLAP data models to optimize performance and support business intelligence processes

MANAGEMENT SKILLS

- Successfully managed complex projects with global implementation, and rapidly evolving requirements.
- Collaborated and communicated effectively with cross-functional teams including business users, data analysts, and other stakeholders to comprehend their requirements and provide technical solutions.
- Extensive experience in driving projects under challenging circumstances, including tight deadlines, high-pressure situations, and ambiguous conditions
- Experience in project planning as well as project estimations and resource allocation
- Adept in providing analytical support to key business applications/solutions
- Acted as liaison between business users and the development team in translating requirements, resolving issues ,and handling escalations
- Possess hands-on experience in operating with the Onshore-Offshore model, skillfully led teams ranging from 5 to 10 members, and delivered projects ranging from 1 to 3 Million

TECHNICAL SKILLS

Snowflake & AWS	Snowsight, SnowPipe, SnowSQL,S3, EC2, RDS, RedShift, Lambda, Cloud Front, Step Functions, Cloud Formation, ECS, EMR
BigData&CI/CD Tools	Databricks, HDFS, MapReduce, Spark, Hive, GitHub, BitBucket
ETL/BI Tools	Dbt, Informatica PC, IICS, ICRT, SSIS, PowerExchange,IDQ,Tableau
Databases	Snowflake, Teradata 12.0/14 ,Netezza, SAP HANAOracle, DB2, MS-SQL Server(TSQL), MongoDB
Data Modeling Tools	Erwin 4.2/4.0, MS Visio
Programming Languages	SQL, Python Scripting, Pandas, UNIX Scripting, PL/SQL, and Core JAVA
Other Tools	Azure DevOps (ADO), Jira, SAP GUI, HP QC, Aqua Studio, Force.com, Teradata SQL Assistant, WinSQL, Toad, SQL Developer, Putty, PVCS, Aginity Workbench for Netezza, SAP HANA Studio, PAC2000 and Service-Now
Scheduling Tools	IBM Tivoli (TWS), Tidal, Stone Branch and Autosys

EDUCATION & CERTIFICATIONS

- B.Tech (Electronics&Comm) , J.N.T.U, Hyderabad, India; Year 2000-2004;

- SnowPro **Advanced** Architect & Core Certified, Snowflake
- AWS **3x** Certified (Certified Developer, Solution Architect-Associate and Cloud Practitioner)
- **Python** Data Structures - University of Michigan
- Informatica Power Center Certified Developer, Informatica Corporation.
- Oracle Database 11g: SQL, Oracle Corporation
- Teradata 12 Certified Professional, Teradata Corporation
- Informatica Cloud 101 and 201

Certification Record: https://www.credly.com/users/siva-kumar-kandivalasa/badges
https://www.credential.net/profile/skandivalasa/wallet

WORK EXPERIENCE

Client: **Novo Nordisk**
Role: **Snowflake Data Architect\Data Engineer, Lead** **Feb'24 to till date**
Employer: **Brillio**

Project Name: **Enterprise Data Warehouse (EDW) - HCP Marketing Systems(Campaigns)**
Summary:

- Accountable for developing and implementing new applications to analyze HCP marketing data for Adobe Experience Management (AEP), Adobe Campaign Management (ACM), NBA media, Marketing Mix, and various downstream systems. The Enterprise Data Warehouse (EDW) functions as a central repository, aggregating data from both third-party and direct vendors containing HCP campaign information. The EDW enhances this data to underpin omnichannel dashboards and support any downstream systems that utilize the data.
- Coordinate with stakeholders, line of business (LoB) teams, and both downstream and upstream owners to gather requirements.
- Responsible for developing project plans, establishing clear timelines, estimating efforts, identifying system dependencies, and collaborating with Scrum Masters
- As a project lead, accountable for delivering all project outcomes and ensuring client satisfaction
- Designed and implemented ETL/ELT pipelines to transfer vendor files from AWS S3 external buckets to Unix systems, from Unix to AWS S3 internal buckets, and ingested data into Snowflake tables
- Created and implemented data lake architecture on AWS S3 to consolidate structured and semi-structured data
- Created Snowflake stages(internal/external), File Formats, Sequences and Pipes to copy and build ETL pipelines
- Created snowflake datashares to share data across multiple teams within the organization
- Managed access to database objects via RBAC at various levels and ensured minimum grants provided to users
- Applied time travel on core tables for data retention and recovery
- Ingesting data using snow pipe using AWS SQS for small workloads
- Processed and managed a variety of datasets, including IQVIA, LAAD, and Komodo
- Implemented quality control checks on data files, master control files, and metadata to ensure high standards of data quality
- Responsible for leading issue triaging, root cause analysis, solution design and implementation until issue is resolved
- Interact with on-site and off-shore technical resources to elicit business requirements, lead technical design, provide guidance on best design practices
- Responsible for sprint planning, capacity management, and ADO tracking
- Led and mentored a team of 6, overseeing their development and optimizing team performance

Environment: Snowflake, SnowSQL, AWS S3, AWS SQS, AWS SNS, Python Scripting, Postgre SQL, Unix, Informatica, Azure DavOps and Service Now

Client: **Tapestry, Inc.** **Oct'21 to Feb'24**
Role: **Snowflake Architect/Data Engineer**

Project Name: **S4 HANA to Snowflake Migration/E2EF Performance Management Program**
Summary:

- Interacted with Business teams to understand their requirements, gathered all data elements, and defined timelines
- Designed and created data model and database objects in Snowflake and HANA S4 for new, and enhancements
- Integrated Snowflake with Informatica & DbT, created projects and models, custom SQLs, cloned objects, and developed ETL pipelines
- Created data ingestion process in Snowflake to copy multiple files(bulk loading) and load in Snowflake stage tables.
- Created Snowflake stages(internal/external), File Formats, Sequences and Pipes to copy and build ETL pipelines
- Built ETL pipelines to process files from AWS S3 to Snowflake stages based on AWS SQS notifications and automated
- Demonstrated expertise in data modeling, database design, and schema management for different types of databases, such as relational, NoSQL, and columnar databases, ensuring efficient data storage and retrieval in alignment with the project requirements
- Developed a centralized data lake for the client to store and analyze customer behavior data, enabling personalized marketing strategies
- Reduced load time by **40**% by optimizing snowflake database queries using pruning, restructuring database tables, and minimizing data spillage to remote and local disks
- Have used Snowflake utilities, SnowSQL, and SnowPipe as and when needed

- Implemented Snowflake query and search optimization methods for query optimization, time travel, zero-copy cloning, optimized virtual warehouses, created resource monitors and defined credit thresholds, defined roles (RBAC), and managed users
- Implemented a robust data validation framework, employing automated checks and validations, to maintain data integrity and quality throughout the project lifecycle.
- Performed rigorous data validation to ensure the accuracy and integrity of the migrated data. Compare data between the SAP HANA S/4 system and Snowflake to identify any discrepancies or inconsistencies.
- Created data reconciliation framework in ETL to validate weekly/monthly/quarterly/yearly counts and quantities between Snowflake and HANA S4 tables
- Implemented multiple sub-projects including Gift-card Escheatment, OMS Reconciliation for orders, Demand Sales, Traffic, DC Inventory Checks, and Store end of day comments in HANA and deployed to Snowflake
- Created Inbound and Outbound processes and designed end-to-end ETL solutions including MFT jobs creation, file generation, cleansing, loading, auditing, validation, scheduling, and archiving
- Created Unix shell scripts for XML source file validation, count checks, list file generations, and archiving
- Implemented error reporting and audit process for ETL loads

Environment: Snowflake, Informatica, Dbt, SnowSQL, SnowPipe, AWS-S3, AWS-SQS, Tableau, SAP HANA S4, Teamcenter, PythonScripting, DataFrame, Unix Shell Scripting, MFT(Axway SecureTransport), Jira & ServiceNow

Client: **Wells Fargo** **Oct'19 to Oct'21**
Role: **Cloud Data Architect**

Project Name: **Capital Markets Platform Management**
Summary:
- Involved in migration of a legacy on-premise application to AWS cloud
- As a key member of the AWS Solution Architecture team, provided inputs in designing, building, and modernizing applications, software, and services on the AWS platform.
- Collaborated with cross-functional teams to design and implement cloud-based solutions using AWS. Conducted system analysis and provided recommendations to improve performance and scalability
- Managed S3 buckets, setup policies, granted least access to applications and services via I AM roles
- Configured AWS CloudTrail to log IAM events for auditing and compliance
- Developed strategies to build serverless models versus traditional models

Environment: AWS, S3, Autoscaling, EC2, Route 53, RDS, Load Balancer, Cloud Watch, Cloud Trail, Informatica, Oracle, Putty, WinScp, SQL Developer, UNIX scripting, and Tivoli Workload Scheduler (TWS)

Client: **Tapestry (Coach)** **Feb'19 to Oct'19**
Domain: **Retail**

Project Name: **Tapestry S4 Implementation Phase 3 (Legacy to S4 HANA 1709 Data Migration)**
Summary:
The objective of this project is to migrate the data which is needed to carry on business in the new SAP system from the legacy systems. Migrated data will include transactional data such as inventory, open sales orders, open purchase orders, and the related materials (articles), customers, vendors, sites, etc. This migration is an iterative and interconnected approach where we analyze the data, extract and transform the data, validate the data, load it into target system(s) and reconcile the data loaded in the target system back to the source system.

- Understanding and analyzing functional specification documents.
- Developing ETL jobs for extraction, transformation, pre-load, load and post load processes.
- Extracted data from SAP R/3 system to Netezza stage using Power Exchange SAP ALE Integration.
- Load process involves loading data into SAP IDOCs or BAPI
- Generating validation scripts for reconciliation.
- Collaborated with Business and SAP functional teams to implement the conversions.

Environment: Informatica Power Center 10.2, Netezza, Oracle, Putty, WinScp, SQL Developer, SAP GUI and UNIX shell scripting

Client: **Wells Fargo** **May'18 to Feb'19**
Domain: **Banking**

Project Name: **AMCT Reporting Streamlining Project**
Summary:
The objective of this project is to centralize AMCT reporting for CIO's areas which allows them to project a better and more accurate status of all application outrages, upgrades, and compliance to standards. AMCT (Application management calendaring tool) is an internal application tool provides application managers and their planning/support teams a consolidated view of activities that impact their applications and the underlying assets – and the ability to plan and manage these activities in a single tool. However data entering AMCT is very manual and error prone. The data being given to the CIO's and their reports from the infrastructure areas are extremely difficult to decipher, and each one of these areas send reports in different formats.

To overcome these difficulties and avoid manual process, ETL process has been used to extract data from various input files, transforms them to a normalized format and load them to a database for further processing.

- Involved in analysis, requirement gathering and designing activities.
- Collaborated with various business and technical teams to gather requirements and created design documents.

- Analyzed source data received from various source systems (WAM, iTRACKER, PTS, WSUS, Past Due Patches, QVRu etc.) using join and enterprise discovery profiling methods to identify primary keys and integrity checks.
- Created DQ Rules and score cards using profile results.
- Created DQ mappings for data cleansing and validations, using different transformations and Integrated with power center mappings.
- Created/Modified file validation, file creation (indirect) and file archival scripts using UNIX shell scripting.
- Participated in creating data model and performance tuning of ETL process.

Environment: Informatica Power Center 10.2, IDQ 10.2, MicroStrategy 10.3.0, Oracle, Putty, WinScp, SQL Developer, UNIX scripting, and Tivoli Workload Scheduler (TWS)

Client: **Campbell Soups Company, NJ** **Apr'17 to May'18**
Domain: **Retail**

Project Name: **Cornerstone Americas'**
Summary:
Project Cornerstone Americas aims to deliver processes to collect necessary plant floor data to drive process optimization, reduce process variability, minimize product variation and provide data to R&D to perform continuous quality evaluation of manufactured products.
Project aims to deploy a standard set of business processes, tools and applications to facilitate collection of key in-process data at various data points across the manufacturing operations

As part of ETL implementation, extracted data from key data sources, MES(Manufacturing execution system) which helps to collect the real time plant quality related data from plant equipment as the process order gets executed, QLM which helps to collect master data related to product specifications and product characteristics, SAP ECC helps to collect process orders, MINI MINT(feeds to Mexvision 1.0) which collects production run details and actual production quantities at various operations of a process order and loading into mexvision data warehouse.

A framework in ETL is created to use common code for all plants which extracts data from all plants and stores into same data warehouse tables. Plants data is scheduled to collect for every one hr. and process orders data is scheduled to collect for every 15 mins into DWH.

- Contributed in the development of system requirements and design specifications.
- Created ETL specification documents using functional documents.
- Participated in the design and development of Dimensional modeling.
- Participated in creating database scripts using Erwin forward and reverse engineering techniques.
- Extracted data from SAP R/3(ECC) using ABAP integration with Informatica Power Exchange
- Created ETL framework to reuse the code across all plants.
- Created DQ Rules, data/custom profiles and score cards using profile results.
- Created DQ mappings for data cleansing and validations, using different transformations.
- Integrated DQ mappings with power center mappings.
- Created Informatica objects and Scheduled in IBM TWS.
- Created and modified UNIX shell scripts to run the jobs using framework.
- Created database Views, MV'S and Stored Procedures to refresh the MV's, as and where needed.
- Participated in Database objects, ETL and Tivoli Jobs migration to QA and Production environments.
- Interacted with various business teams, users and provided support in all the phases of project.

Environment: Informatica Power Center 10.x, 9.x, Informatica Power Exchange for SAP NetWeaver, IDQ 10.x, MicroStrategy 10.3.0, Unix Shell Scripting, SQL Server, Oracle, SAP R/3, Putty, WinScp, SQL Developer and Tivoli Workload Scheduler(TWS)

Client: **The Bank of Tokyo-Mitsubishi UFG, NJ** **May'15 to Mar'17**
Domain: **Banking**

Project Name: **OFSAA Leones 42**
Summary:
The main objective of the project OFSAA Leones (42) is to 'Establish a New Bank single environment Profitability Reporting for MUFG Americas'. Currently MUFG UB (Union Bank) and Head Quarters of the Americas ("HQA") have two independent OFSA 4.539 systems for Management and Profitability Reporting. Both Banks are on systems that are more than 10 years old. Both HQA and UB have profitability systems Oracle Financial Services Applications (OFSA) that require upgrades to improve management reporting. In order to support MUFG Union Bank and HQA's ability to analyze its business on an organizational, customer and product level, the combined organization is implementing OFSAA (version 6.0.3). The new OFSAA system will be maintained by MUFG Union Bank and located in the Chandler, AZ data center.
HQA will provide data feeds from EDW and Oracle GL to MUFG Union Bank via Informatica ETL/ Connect Direct for the single OFSAA instance.

Current OFSA ETL process will be modified to be use with the new OFSAA environment (6.1). HQA ETL processes for the current OFSA processes will be received from the HQA EDW to the MUFG Enterprise Data Hub (EDH) within the MUFG UB entity. Current MUFG UB OFSA files will be modified and sent to the EDH via file gateway. The EDH will bring the data into separate staging areas and then incorporate the new OFSAA logic required to feed the OFSAA staging, instrument and look up tables.

- Designed ETL mappings/sessions/workflows for the monthly feeds to extract data from Core Bank DWH and generate files to provide to EDH.
- Involved in Logical / Physical data modeling of the Banks HQA DWH.
- Interacted and worked closely with HQA Business users in data validations and reconciliation of OFSA 4.5 and OFSAA 6.5 systems.
- Involved in SIT, UAT , Pre and Post PROD deployment activities.
- Scheduled and Monitored workflows on the TIDAL server.
- Worked on performance tuning of the existing ETL framework (OFSA 4.5)

Environment: Informatica Power Center 9.1.0 Unix Shell Scripting, DB2, Oracle, TOAD, Putty and Tidal.

Client: **Campbell Soups Company, NJ** **Jan'14 to May'15**
Domain: **Retail**

Project Name: **Financial Reporting and Business Intelligence (FIRE)**
Summary:
Campbell's corporate and business unit financial analysts and leadership have been using BPC for both planning and reporting activities, with limited standardization and visualization capabilities with respect to reporting and analysis. It is the purpose of this project to deploy a solution that will improve financial reporting and analysis capabilities by deploying a robust set of scorecards and dashboards, in addition to managed and ad-hoc reporting. This will be enabled by decoupling the reporting system from the planning system, incorporating a purpose-built data mart that will be optimized for reporting and analysis requirements, in addition to enhancing the quality of reporting output with existing enterprise visualization and mobility tools.

As part of ETL, has to extract Sales transactional data residing in SAP BPC System and Master and Hierarchies data from Oracle and Flat files sources(PELEUS) using Informatica Power center and transformed them into Stage and Data warehouse systems with applied business logic.

- Contributed in the development of system requirements and design specifications.
- Created and modified table scripts using Erwin forward and reverse engineering.
- Created database Partitions and Sub-Partitions to improve the performance and ease the maintenance.
- Created indexes on partitioned and non-partitioned tables.
- Created Views, Materialized Views (MV) and Stored Procedures to refresh the MV's, as and where needed.
- Worked with Informatica Data Quality 9.6.1(IDQ) Toolkit for analyzing, standardizing, cleansing, matching, conversion, exception handling, reporting and monitoring the data
- Experience in implementing Data quality rules using IDQ.
- Created scorecards to review data quality.
- Created simple to complex mappings using Informatica Power Center components.
- Created and Scheduled Job Steams and Jobs in Tivoli (TWS).

Environment: Informatica Power Center 9.5.1, Informatica Developer\Analyst 9.5.1, SAP, Informatica Power Exchange for Web Services, XML, Oracle 11.2.0, UNIX Shell Scripting, Oracle SQL Developer and Putty.

Organization: Cognizant Technology Solutions
Client: Merck Co **Oct'09 to Jan'14**
Role: Solution Architect/ETL Architect/Tech Lead

Project Name: Merck Service Rendered for ET
Summary: Merck Services Rendered for ET is development and enhancement project. It is an umbrella project in which multiple projects execute in parallel. It provides extract, transform and load (ETL) services to 26 USHH project teams across 5 business areas. Each project is independent from the other projects and requires strict timelines. An Umbrella structure of Merck Services Rendered for ET project provides centralized control over the ETL project management. Each individual project goes through entire Software development life cycle which involves phases like Requirement Gathering, Analysis, Design, Development, Verification and Validation of code and Post Implementation Support. Technologies used in the project are Informatica 9/8.6.1. , Oracle 9i/10g, UNIX Shell scripting, Cognos 8 and Spotfire.

Sub Projects:
- Emerging Markets (EMEA) – Turkey, Russia & South Africa
- Activities and X-force Dashboard (EMEA)
- Sales & Activities (EMEA)
- GENESYS
- UNIVADIS
- GHH Divisional Data warehouse (DDW)
- MCC Divisional Data warehouse (DDW)

- Managed requirements and design phase towards identifying mutually agreed solutions with technical teams and business partners
- Responsible for providing requirements, functional knowledge and design documents to ETL team for development
- Responsible for converting requirements into comprehensive detail design specifications (Integration, databases, data flows, transformations, interfaces etc.) for solutions
- Responsible of creating project plan and working with customer on regular basis to show the progress at each stage
- Responsible for all ETL deliverables in SDLC methodology such as integrated functional specification, source to target mapping sheet, design, coding, unit testing and deployments
- Integrated Informatica with Salesforce Cloud in extracting Customer flags and RFM data
- Integrated Informatica with Web Services using Power Exchange in extracting data

- Other source systems majorly used are Teradata, Oracle, SQL Server, XML and Flat Files
- Extensively used Teradata utilities MULTILOAD, FASTLOAD, TPUMP, TPT and BTEQ Scripts to load data into Data warehouse
- Created UNIX shell scripts to run the Informatica workflows, Dynamic parameter files creation, indirect file creation, file watcher and email notification etc
- Responsible for driving the team to meet the objective of the project by proactively taking decisions

Environment: Informatica Power Center 9.1, Informatica Power Exchange 9.1 for SAP NetWeaver, Teradata, Unix Shell Scripting, Sources (SAP BW, Share Point List, MS SQL Server, Flat files & Oracle) , Teradata SQL Assistant , Salesforce, Force.com, Putty, Autosys & Soap UI

Organization: **Fujitsu Consulting Services**
Client: **Principal Financial Group** **Oct ’06 to Jan’ 09**
Role: **ETL Developer**

Project Name: Treasury Consulting Project and GE Securities
Job Duties:

- Designed ETL jobs and scheduled using third party scheduler's
- Implemented performance tuning techniques to increase the session performance
- Responsible for Unit testing and creating migration plans
- Coordinated with the various teams including database, unix and stake holders involved in project

Environment: Informatica Power Center 8.1/7.1, DB2 UDB, Oracle, SQL, PL/SQL & UNIX Shell Scripting.