

# SHIVARAM CHALLA

**BIG DATA DEVELOPER/DW + ETL ARCHITECT/SR. DATA ANALYST**

<http://challa.net>  
[shivaramchalla@gmail.com](mailto:shivaramchalla@gmail.com)  
(201) 362-3785

## PROFILE

I specialize in developing data centric solutions to fulfill organizations information needs. Experienced in designing, implementing and maintaining system interfaces with synchronous and Asynchronous communication, data extraction, data profiling, data integration, data quality, Metadata analysis; while using the most suitable tool and/or technology for the job, including Hadoop platform (both Cloudera and HortonWorks), Direct Database access, Web services (WSDL), Message Queues, CSV/text files, web scrapping, programming API, FTP, XML, XSL, Custom programming using C#, ColdFusion (CFML), VBScript and Batch programming.

I produce outstanding results leveraging on my strong technical, analytical and problem solving skills to analyze, design, develop and support highly available, highly performing systems, while leveraging on user requirements analysis, database design skills, strong programming background, excellent oral and written communication and adroit understanding of business.

## EDUCATION

Computer Engineering, Master of Science – Graduated 2005  
[Fairleigh Dickinson University, NJ, USA](#)

## PROFESSIONAL ABSTRACT

- Over **13 years** of compound IT experience. Over **8 years** of experience Data Warehouse Architect. Over **9 years** of experience designing, developing & implementing ETL/ELT solutions. **3 years** of experience working with Hadoop (both Cloudera and Horton works) with HBase, Hive, HDFS, Sqoop, MapReduce programs (using MRJob in python), Ambari, Cloudera Hue, HBase Master and more.
- Always looking to learn and understand new technologies and their relevance to the business
- Experience with Agile/SCRUM and RAD/SCRUM development environment
- Experience supporting mission critical and several other applications
- Experience working in large OLAP environment with over 1.2PB data in Cloudera & Horton Works Hadoop platform and/Action Paracel
- Experience working in large environment with Mission critical OLTP database applications with over 1000 transactions/minute, 60+ SQL Servers and 150 developers
- Experience building and maintaining MDM systems at multiple organizations
- Experience in designing, implementing, maintaining, and documenting complex OLTP applications, and data warehouse environments.
- Extensive experience in design techniques, ETL processes, relational database theory, N-Tier development
- Excellent report creation skills using Microsoft SQL Server Reporting Services (SSRS), Crystal report and custom ASP.net/PHP based reporting solutions.
- Progressive management experience, team leadership and understands the value and importance of delegating work and responsibilities
- Excellent communication, presentation and interpersonal skills and ability to prioritize and coordinate work across different teams
- Experience creating and maintaining system interfaces between Work management system, Customer Information System, Finance and Accounting system and HR system

## Skills/Tools:

Hadoop, HDFS, HBase, Hive, Ambari, HBase Master, MapReduce(MR), Zookeeper, YARN, Sqoop, Netezza, Aginity WorkBench, SAS DataFlux, Oracle Advanced Queue, IBM Web sphere MQ, MS SQL Server Integration Services, SQL Server 2012/2008/2005/ 2000 (32 bit & 64 bit), C#, Python 2.7, Bash, ASP.net, ASP, PHP, Java, M/R jobs, JavaScript, JQuery, ColdFusion, WSDL, XML, XSL, Database Application's supported include Issue/Bug tracking system, Order processing system, Order Fulfillment system, Customer Order tracking system, Back Office billing system, Vendor Management System, Production release Ticketing system, Version control system (Subversion, source safe), Legal Document management system, Client Contact and Preference management system, Web Reporting System, Windows 2003 & 2000 servers (32 bit & 64 bit), Transact SQL (T-SQL), Integration Services (SSIS), Reporting Services (SSRS), Oracle 10g, PL/SQL, Install, Upgrade OS, Configuration, Database Design, Partitioning Strategies, Performance monitoring & tuning VLDB, Index fine tuning, covering indexes, Normalization, De-Normalization, Automated Maintenance, Failover Clustering, Replication, Log shipping, Backup, Restore, Disaster Recovery, Storage capacity planning, ETL, Data Import & Export, Database Engine Tuning Advisor, Microsoft Security Baseline analyser, SQL Server Profiler, Database Mail, SQL Mail, IIS6.0, IIS 7.0, Secure FTP (SSH), SQL

- Experience creating and maintaining interfaces using Web services, Message Queues (MQ), flat files, direct data access
- Experience creating and maintaining Reporting infrastructure in Microsoft SQL Server Reporting Services, Crystal Reports, Oracle Reports, ASP.net and Classic ASP.
- Expert in analyzing and implementing normalization/de-normalization needs.
- Extensive experience with PL/SQL and T-SQL in creating Tables, Views, Indexes, Stored Procedures, user defined functions, Triggers, Cursors, Roles, Logins, Users, relational database models, self-sufficient data dictionaries and data integrity with constraints.
- Hands on experience and knowledge of high availability SQL Server solutions, including Log shipping, Replication and SQL Server clustering.
- Created and used the contingency plans for disaster recovery for the databases and operating systems.
- Extensive experience with Monitoring, performance tuning and analyzing database performance and allocate server resources to achieve optimum DB performance.
- Managed Data migration from Oracle, Teradata to SQL Server 2005 and Vice versa
- Developed PL/SQL packages, shell scripts to execute the packages and PL/SQL blocks to debug the package in production environment and migrated bulk data.
- Improved performance by using Explain Plan, creating appropriate indexes, queries optimization, utilizing table spaces and partitioning schemes.
- Excellent SQL code optimization skills using explain plans/execution plans.
- Strong understanding of SQL server Internals via Dynamic Management Views (DMV).
- Extensive experience in developing, deploying and monitoring SSIS packages, DTS Packages and DTS to SSIS conversion.

Server Security, Roles, Grant/Revoke permissions, RDBMS concepts, Maintain Mission critical OLTP Applications, Provide training / guidance, Analysis Services (SSAS), Star schema, snow flake schema, SSMS, BIDS, MS VISIO, Collabnet Subversion, MS Visual SourceSafe, QlikView, command shell, Windows power shell, Teradata Transact SQL, Teradata SQL Assistant 7.2, MySQL, MS Office Suite, Expert in Excel & Access, Swish Max, CSS, HTML4.0, XHTML1.0, JScript, VBScript, Shell scripting, XML/XSL, Visual Basic/VB.net, C#.net, Windows Vista/XP/2000, Red hat/Ubuntu Linux UNIX Apache web server, Active Perl5.10.0Build1002, SDLC, RAD-SCRUM,SOX Compliance, 24/7 On call duty.

## EXPERIENCE

### Senior Database Application Developer

April 2013 - Present

Merkle Inc. (7001 Columbia Gateway Drive, Columbia, MD 21046)

**Business Environment:** Merkle is the nation's largest and fastest-growing performance marketing agency. Merkle helps brands transform their marketing organization using data, analytics, and technology to create meaningful, personalized customer experiences that deliver competitive advantage.

### Certification, Recognition and Awards at Merkle:

**Certified Merkle Digital Professional:** I have attended 10 classes and passed 11 exams to be certified MDP

**Merkle One Team Award** – for Hadoop Migration work from Cloudera to Horton works

**Merkle attribute award - Passionate** – for helping another team for an oracle database related project

**Merkle attribute award - Sense of Urgency** – for diligently working on getting a client's corrupt reference base back.

### Projects and Responsibilities:

I was part of the team that was responsible for developing and maintaining an **Enterprise wide reference base (Master Data Management)** that hosts standardized, consistent and most current data for individual customers for each client's reference base. This reference base becomes the central source of information based on which the marketing campaign attributions are done. Various other platforms use this as the source data.

**Hadoop:** Hadoop platform was used to create a Connected Recognition (CR) platform, which include Customer Data Integration (**CDI**) and Digital Data Integration (**DDI**). The data was stored in HBase tables, Hive tables and HDFS files. I used Ambari to check the system stability and data node availability. Used HBase Master to monitor proper data distribution to avoid hot spots in HBase tables.

**Connected Recognition (CR) - Job automation:** I have implemented the automation of daily data ingestion and weekly report generation job for several clients. This includes developing several scripts to move files between Hadoop's HDFS and SAN, created several **Hive scripts** and run the Map-Reduce job to process the data, creating jobs in the enterprise scheduling software (Tidal), and export the report data into SQL server 2013 for billing. This automation framework was eventually applied to all the other clients by using the codebase I developed. This has and will save a lot of time for all client onboarding's. Involved Bash, Hive, Scoop, Java MR jobs and Python.

**HBase table Performance optimization:** The rows in HBase tables are sorted lexicographically using the row key. This allows for fast access (acts like a clustered index) by using the start row-key and stop row-key on each region which may

or may not be serviced by separate region server. When too many records within a range of start keys (like HASH keys starting with letter S) are clustered together (this phenomenon is called hot-spotting), I split the region and put these two regions on separate region servers. Hot-spotting usually leads to job failures as the uneven write load on a single region puts undue pressure on respective region server, causing region server failure. I used HBase Master to detect the regions server with too many regions and regions that are highly clustered together.

**Automated** Hadoop MapReduce job execution and Hive scripts and query executions using **Python 2.7**. On a separate project, I used/tested **Luigi** to run multiple MapReduce jobs and Hive scripts (Luigi is a python module used to build complex pipeline for batch and jobs).

**Automated** the billing process by using **Sqoop** to transfer the MapReduce job output from HDFS to SQL server 2012

**Legacy Key migration:** When migrating Clients from legacy platform to Hadoop based Customer Data Integration (CDI) platform, we lose the ability to attribute a sale to a given marketing campaign. This project allows us to tie the legacy IDs to be mapped to the Current Platform, giving us the ability to attribute a sale to a marketing campaign. \* **Skills Used:** Involved writing several Hive scripts and chaining them using Bash scripts. I created **External Hive tables** on top of HBase tables.

**Migration:** successfully migrated **MasterSource** reference base from legacy KL2 to CR-CDI platform. This is by far the biggest reference base of all the clients we manage on CDI platform.

**HBase Snapshots:** Successfully managed (created, Restored and dropped) HBase table snapshots in several instances to restore a corrupted reference base for individual clients

**Connected Recognition - CDI - POC implementation:** Proof of Concept database is a small scale database hosted on SQL Server 2012 database on a smaller scale of data (not exceeding 10 million records). Once the client team reviews the output data and approves the match rules, then the actual data is ingested on to the Hadoop Platform.

**Tier 3 support:** For connected recognition (CR) and Knowledge Link (KL), I am part of the team that provides **24/7** hours support on a rotating schedule basis. I was called several times late at night and I took I resolved the issues and brought the solution/process back online. Several times, I had to initiate SWAT calls, bringing everyone necessary and **took the lead** to successfully resolve the failures. \* **Skills used:** SQL expertise, System integration knowledge, Knowledge on Hadoop stack (HBase, Hive, HDFS file manipulation, Sqoop, HBase Master, HBase hotspot detection, etc.), Cisco TIDAL – Enterprise scheduling software and leadership and managerial skills.

**Data Analysis:** Perform data analysis by extracting data from HBase, Hive, HDFS files, SQL server 2008/2012 and Netezza

**Knowledge link 3 platform:** I have developed and implemented several functionalities and improvements to the platform. Some efforts include feature development and implementation, bug fixes, etc. Involved Bash scripting

**CR - Digital Data Integration:** I worked on modifying the T-SQL code embedded in the java code to process input data.

**Paracel Upgrade feasibility study:** Paracel (Actian Matrix) database is a parallel relational database system using a shared-nothing architecture and columnar orientation. I did the upgrade feasibility study; this involved, setting the environment on a newly built test cluster, import sample data, test connectivity from local and remote access. Ensure the SQL parsing engine would support the existing code base.

**Hadoop environment conversion:** I am part of the team that migrated the Hadoop environment from Cloudera to Horton Works. Played a key role in testing the successful migration of Sqoop, HDFS and HBase usability. The conversion team received recognition and award in a company-wide quarterly meeting for successful migration.

**On-boarded several clients onto CDI platform and KL35 platform:** On-boarded T-Mobile, Sanofi, PayPal credit, ESPN, MetLife, Abercrombie, Cedar fair, Decker's, Medical mutual, Regions bank, Sun power and more

**Worked on KL35 to CDI migration effort:** worked on migrating several clients like Oriental trading co, AAA, Dell, Carters, Chase home equity.

**On-boarded clients to Standardized DDI solution:** Geico, 1-800-flowers and Toyota

Following skills and tools were used to deliver a relevant solution in various scenarios. Bash scripting, Hadoop (HDFS file manipulation, written several Hive scripts and ad-hoc HiveQL queries, created, dropped and queried HBase tables and Created, dropped and restored snapshots of HBase tables), SQL Server 2012, T-SQL, Oracle PL/SQL, Hue, Ambari, Korn shell scripting, Python 2.6, Java 7

**Data Warehouse/ETL Architect and Admin**

**June 2009 - April 2013**

**Lockheed Martin IS & GS Civil, Energy services, Manassas, VA**

**Business Environment:** Energy Services division of Lockheed Martin provides IT services to several Electric utilities. And Northern Virginia Electric Cooperative is one of them. NOVEC provides electricity to about 150,000 customers in Northern Virginia with highest customer satisfaction in USA for 2012 (by J.D. Power)

**Projects and Responsibilities:** EDW/BI Project: As the technical lead on the project, I provided technical lead to Design, Develop, Implement, Test and Support EDW/BI solution over multiple phases and be the subject matter expert on the team. The responsibilities included working with the end-users for Requirements Analysis, Technical specification development, Data Integration, Gap Analysis, Data Mapping, Data Profiling, Data Cleansing, validation and the overall solution architecture. I developed ETL (Extract, Transform & Load) Architecture, self-contained **Master Data Management (MDM) system** within the oracle database, Performance Monitoring system for analysis and tuning, Role based security model for providing modular access to the Enterprise Data Warehouse. I provided the hardware specification for the Database server and ETL Application server; for Production and test environments.

As the primary **application admin** for Logica ARM WMIS (Work Management Information System), I installed, upgraded and maintained the Server infrastructure, which includes Oracle 11g database, Oracle Application Server, Oracle Advanced Queues, IBM WebSphere MQ and ARM IMFPlus Web Services. Maintain, upgrade and troubleshoot the interfaces between WMIS, GIS (Geographic Information System), Lawson (Financial & HR), Accounting (Power plant), Outage management and Crew Dispatch system (Clevest Mobile Field Force - MFF). I reviewed, installed, upgraded and patched the database, application and coordinated the fat client deployment on to over 200 PCs.

**SQL Server 2005/2008 DBA:** I have managed 14 SQL Servers, ensured the stability of the servers, planning for disaster recovery, monitoring for performance issues and tuning the database for optimal performance. With the server team, I planned, provided the hardware specifications and deployed the SQL Server software on Windows 2008 enterprise server for Mobile Field Force management software. **Oracle 9i/10g/11g:** As part of a team of three DBA's, I managed the various versions of Oracle databases supporting the enterprise applications such as Enterprise Data Warehouse, Work Management Information System, Geographic Information System, power plan (Accounting software)

**Consolidated the Reporting infrastructure** providing the end users one access point to all the organizations reports. The Reporting platforms consolidated include Crystal reports, Oracle Reports, ASP.net/ASP and Cold fusion based custom reports. I also provided access to some of the reports from within the client application.

**Consulted as the subject matter expert for databases and System Interfaces and system integration** for several projects along with the server team. The projects involved understanding the business challenges and working across business groups and technology teams to ensure alignment between business solution definition and system architecture for the organization.

**Skills and tools used** to deliver a relevant solution in various scenarios. Knowledge and Experience in System Integration, data integration, Data warehouse and ETL design and architecture, dimensional modeling, OLAP, OLTP, RDBMS, Normal forms, Requirements analysis, source system profiling, SDLC and DBA experience. SAS DataFlux Data Management server and studio, Toad for Oracle, PL/SQL Microsoft SQL Server 2005/2008, T-SQL, crystal reports, Oracle Reports, ASP.net, IBM Web sphere MQ, IMF Plus MQ and Oracle MQ based system interfaces. XML, XSLT, XHTML 1.1, JavaScript, MS Visual studio .net for web and application development using C# programming language, working knowledge of IBM DB2, Source System specific and ODBC Data connections, MS Visio and Office suite.

## **DBA & SME in DB Design, T-SQL, SSIS & SSRS**

**November 2007 – May 2009**

**Sprint Nextel, Reston, VA**

**Responsibilities:** Managing the Device Support Data Team's consolidated Data Warehouse Application, which is part of the Sprint Nextel's decision support system. My role was to establish an enterprise level database environment with 98% uptime and data availability at an agreed upon SLA with 20,000 users from coast to coast. This includes upgrading and maintaining 2 MS SQL Server 2000 to SQL2005 (64-bit & 32-bit), Overhaul of the then existing security infrastructure, automation of several data import and exports(SSIS), set up web reporting server (SSRS), setup SFTP server and optimally maintaining data that is close to .75 TB. Being the publicly traded, Fortune 500 Company that Sprint is, I've developed data access policies and procedures that strictly adhere to SOX compliance. Expertise in SQL coding and SSIS design skills were of great use here.

**Skills Used:** Database design and Architecture concepts. Understanding of MS SQL 2005 system objects/information schema, T-SQL expertise, Business acumen, customer communication, Understanding of SQL Server Internals, SQL Server Reporting Services (SSRS), provided training to developers in creating and deploying reports in SSRS, SSIS, PL/SQL, SQL for Teradata (TSQL), VBScript, Batch job (Windows Shell scripting). Efficient SQL Server configuration, Optimized database design, efficient partitioning strategies and effective maintenance, Successful Disaster Recovery, Understanding of SQL Server 2005's Role based security model, knowledge of SQL Server Internals via DMV's.

**Metadata project:** Purpose was to create a holistic view of the how data entered, processed and converted to output format in our system. Involved in every step of Software Development Life Cycle (SDLC). Took a novel approach for creating a self-sufficient system of obtaining and maintaining custom metadata for all the user objects as Extended Properties. Created nightly agent job that executes Procedures to refresh tables that holds inter object dependency data, like job to procedure dependency data. **Skills Used:** Database design and Architecture concepts. Understanding of MS SQL 2005 system objects/information schema.

- Picked up the Total Equipment Protection (TEP/ESRP) analysis project midway and consolidated and optimized the process. Added new code to address changing business needs. Automated this process to self-sufficiently update the data. This effort received accolades from the executive management. **Skills Used:** TSQL expertise, Business acumen, customer communication and worked with tight deadlines.
- Automated several data imports and exports and converted Several Data Transformation Services (DTS) packages to SQL Server Integration Services (SSIS) packages. Data sources include Oracle, Teradata, MS SQL server, MS Access, Spreadsheets, Flat files (CSV, TSV) and pulling and pushing multiple data files from and to FTP & SFTP sites. Also, used VBScript and Batch jobs for some of the automation. **Skills Used:** SSIS, T-SQL, PL/SQL, SQL for Teradata (TSQL), VBScript, Batch job (Windows Shell scripting).
- Optimized a data load from Oracle that used to take over 13 hours (using ODBC connection), to a load that takes less than One hour using Integration Services.
- **Successful Disaster Recovery:** Followed contingency plan created earlier for situations like this. Moved databases to a different server, reinstalled the Windows server 2003 OS and SQL Server 2005 64-bit enterprise edition, and recovered the databases without any data loss.
- Created a web report that shows the system wide data load status, tests all the linked servers and lists the failed jobs. This enables the team to quickly trouble shoot if there are any data load issues. **Skills Used:** Understanding of SQL Server Internals, SQL Server Reporting Services (SSRS) and T SQL.
- Migrated Crystal reports to Reporting services and trained the team in creating and deploying reports in SSRS.
- Created several procedures using Dynamic Management Views (DMV), to automate various administrative tasks. Following are a few of them (published some of these at SQLServerCentral.com). **Skills Used:** SQL Server 2005, knowledge of SQL Server Internals via DMV's, T-SQL.
  - o Index fragmentation: Automatically picks up the most fragmented index and re-indexes or Re-organizes the index. This takes less time and touches only the necessary indexes.
  - o Missing indexes: Provides a list of missing indexes in the system.
  - o Unused indexes: The unused indexes can be dropped, thereby reducing the data load times and meeting the Service Level Agreement (SLA).
- Automatically Deny SQL Connect to users during a certain time to meet a business need. Before this is triggered, all the users' sessions are to be killed forcibly with some exceptions. This is implemented using sys.sysprocesses and cursors.

**Database Administrator, T-SQL, SSIS & SSRS** **July 2005 – October 2007**

**InPhonic/Simplicity, Reston, VA**

**Responsibilities** – Manage enterprise database environment of 60 SQL Server systems, hosting multiple Tera - bytes of databases. The largest individual database size exceeds 400GB. Database systems included 5 active/passive SQL Server clusters on both SQL Server 2000 and SQL Server 2005. Internet storefront, Back Office, billing and fulfillment systems database needs provided entirely through Microsoft SQL Server.

**Tools used:** SQL server Management Studio (SSMS), SQL Server Integration services (SSIS), SQL Server Reporting Services (SSRS), SQL Server Replication monitor, SQL Enterprise Manager, SQL Query Analyzer, SQL Profiler, Red-Gate SQL Compare, LiteSpeed for SQL Server, MS Visual Studio.net 2003; Transact-SQL, PL SQL, Visual Basic.net, VBScript, Jscript, ASP.Net, ActiveX Scripting, DTS, SSIS

**Application and SQL Developer & Dev DBA** **Mar 2005 – Jul 2005**

Office of Treasure Technology, Trenton, NJ

**Windows Web Server admin, SQL DBA/Dev and web Developer** **Jan 2004 - Feb 2005**

Hamburger One LLC, Teaneck, NJ

**SQL server and .Net Developer** **Nov 2003 - May 2004**

SRG Inc., Iselin, NJ

**Database Development/Maintenance staff (Intern)** **Jun 2001 - Dec 2002**

SSI, Dharmapuri, TN, India

**\*\*\* References are available on Request \*\*\***