

Sean Anderson

937-509-8797 / sean.anderson@datavirtue.com

Data Engineer / Architect / Developer

SUMMARY

Senior data engineer and software developer who is passionate about open source community driven advancements in cloud computing, software development, and dev-ops. Looking for a team that is energized by the same desires to learn, create, and contribute.

TECHNOLOGY

Big Data: Hadoop, Hortonworks, Yarn, Hive, Ambari, Tez, MapReduce

Azure: HDInsight, Application Insights, VMs, Resource Templates (ARM), SQL Database, Blob Storage, Data Lake Store (ADLS), Administration

Amazon: DynamoDB, API Gateway, S3, EC2, Administration

Databases: SQL Server 2005, SQL Server 2008, SQL Server 2012, SQL Server 2014, MySQL, MongoDB

Coding: TSQL, SQL, C#, Xamarin.Forms, Java, .NET, PowerShell, PHP

Tools: SQL Server Integration Services (SSIS), PowerBI, Excel(VBA), Visual Studio, TFS, Team Services, SQL Server Management Studio (SSMS), Fiddler, SOAP UI, MongoChef

Data Virtualization: Cisco Information Server

Application Servers: Internet Information Server (IIS 7, IIS 8)

Professional Experience

2017 to Present: Green Dot Corporation (Pasadena, CA & Blue Ash, Ohio)

[RushCard was acquired by Green Dot in Feb 2017]

Principal Systems Engineer

- Exploring Pivotal Cloud Foundry, Docker, and Terraform to gauge effectiveness for improving: software delivery, scalability, high availability, and vendor-neutral cloud adoption
- Leading a project to rearchitect the core message broker service (RabbitMQ) for RushCard
- Planning data center migrations and DR implementations for two subsidiaries
- Consulting with the enterprise monitoring group to develop a clean monitoring and alerting solution

2013 to 2017: RushCard (Blue Ash, Ohio)

Application Engineer / DBA / Systems Architect

- Data platform owner responsible for data engineering and for ensuring fault-tolerance, performance, and recoverability of a \$150MM 24/7 highly-transactional multi-terabyte database environment (10 TB, 16 Servers, 16 AG, 200 Cores)
- Led big data project using HDInsight (Hadoop) and Data Lake Store on Microsoft Azure to create a centralized, standardized repository for all vendor, partner, and production data (data lake)
- Directed database administrators to ensure system performance, backups, 24/7 support, and sound data engineering projects

- Created and revised star-schema data models and developed ETL/ELT for data warehouse using SSIS on SQL Server 2014
- Worked with application architects to develop data models in Entity Framework for new application features and acted as the gate keeper for any code changes (LINQ queries) affecting database access
- Successfully launched a data integration project for a new company acquisition utilizing SSIS to load text files daily from payment processors (MasterCard, FSV, FIS) for risk and customer service reporting
- Led a successful large-scale data conversion over to a completely new data model involving hundreds of tables in a 10TB production database--six people under my direction across a series of milestones for a total system redesign with hard cutover; project lasting 18 months
- Developed complex T-SQL for reports, analytics, data blending, cleansing, and applications utilizing advanced SQL such as windowing (analytical) functions along with proper use of table variables and indexed temp tables in adhoc queries, stored procedures, functions, C# CLR, and table value functions
- Developed many SSIS projects/packages for converting data, transferring files (SFTP), and loading tables with billions of rows utilizing ETL and ELT methods
- Developed applications/scripts using C#, PowerShell, and T-SQL to prepare and verify proper file formats and process batches of data through APIs--using the language appropriate to each situation
- Coached analysts, developers, and report writers on T-SQL features and query optimization (joins, set operations (union), correlated sub queries, analytic functions, grouping, aggregation, temp table and table variables, indexing, sargability)
- Used “downtime” to perform in-depth application security analysis, identifying and resolving numerous attack vectors and compliance violations. Conducted developer OWASP training.
- Identified root cause of application bugs and wrote/submitted C#/.NET code to resolve issues
- Documented complex systems and data flows using Visio for technical and non-technical audience

2010 to 2013: Southern State Community College (Hillsboro, Ohio)

Systems Analyst

- Successfully implemented an application integration and upgrade for a Microsoft Access-based application to SQL Server and automated the exchange of information between the ERP system
- Secured the ERP environment with Microsoft Forefront Threat Management Gateway (Firewall/IDS)
- Discovered and reported on security vulnerabilities by performing penetration testing on ERP web application and network (network and application security)
- Developed an Excel application (object oriented VBA), in cooperation with the head chemistry professor and Turning Point representatives, to fill a gap in Turning Point response system reporting metrics to track student progress throughout a semester with aggregated and visualized results for a flipped classroom initiative.

FORMAL EDUCATION

Associate of Applied Business: Management Information Systems
Minor in Applied Science: Electrical Engineering (Miami University extension)
Southern State Community College: Hillsboro, Ohio (sscc.edu) 3.87 GPA

CERTIFICATIONS

Microsoft Certified Solutions Associate:
<http://www.mycertprofile.com/Profile/7894728660>
SQL Server Data Warehouse and BI Development

Microsoft Certified Office Specialist:
Microsoft Excel 2010
Microsoft Access 2010

ACCOMPLISHMENT SUMMARIES

Large System Implementation (Data Conversion)

Converted 3TB of account servicing and transaction data to a completely new data model going from a monolithic database with hundreds of tables to 15 domain-based databases to support a full application stack rewrite. Worked with the business, key vendors, and the development teams to support a hard cut over on a strict deadline. I wrote most of the conversion ETL/ELT while leading a team of developers and analysts assisting me on other aspects of the project—both offshore and in-house.

Analytics API

I was engaged on several projects for the Added Value division of Kantar to support the Harley Davidson account.

I built an analytics API for survey data; making it possible for applications or BI tools to easily visualize metrics. Taking a process whereby data scientists would ingest and manipulate survey data in R, I built an API in Cisco Information Server that would publish that survey data as an API from a SQL Server database instead. Using their R scripts as the “template” for the API parameters, I worked with the data scientists to clarify requirements and verify my results via the API. Added Value intended to use this system to power metrics on the Harley Davidson (HD) dealership portal.

On another project for Harley Davidson, I built a visualization and reporting solution for end users by capturing email status messages (JSON) from the SendGrid web hook, and feeding those through AWS Lambda (.NET) to a SQL Server instance on EC2/S3; ultimately transforming the data in SQL Server and presenting it through PowerBI to end users. This let HD dealers access a history of email click-through statistics for easy comparison.

Azure Big Data Lake

Successfully led a project to begin consolidating all company data onto Hadoop using the Hortonworks distribution on Azure (HDInsight) and Data Lake Store. I built, configured, automated, and documented the HDInsight instances and storage provisions and wrote Hive scripts to prove out ETL methods. During the project, I discovered several issues and bugs and assisted Microsoft support in documenting and reproducing the issues for Hortonworks. Ultimately, I worked around the bugs and deficiencies in the Hive SQL implementation and could ingest data into efficiently partitioned and bucketed ORC files to reduce storage consumption, speed ingestion, and provide good performance when querying for new data models. While I worked in Hive, I had a data engineer focused on automating ETL directly from our production database using Azure Data Factory. The project proved successful, however our company was undergoing a round of re-capitalization which led to our acquisition—essentially killing this project and all others which were in flight at the time.

Major Financial Audit

Performed in the Lead Data Analyst role for a company audit. The audit was required to satisfy regulators and partner banks after a full system conversion. I worked closely with the company Controller to define requirements and create a complex account balance timeline for 2.5MM accounts. The balance timeline was used in the audit, contributing to the success of the project, and is still in use for an account-level reconciliation effort.

API Redesign

Re-engineered Green Dot API implementation that supports over a billion dollars in transactions. The system was rebuilt at one point to support a revised Green Dot API allowing Green Dot to contact our systems to approve and complete cash loads from Wal-Mart. The rebuild was not designed properly and had serious bugs that would prevent cardholder funds from being loaded and/or would cause serious reconciliation problems for the finance team. I identified the transactional issues and redesigned the system (web service and data flow) to prevent failed cash loads and reconciliation issues. Lacking a formal project management team, I recruited another developer to implement and deploy the changes--successfully resolving the issues.

Improved Large Scale Application Issues

Reduced downtime and outages related to application performance and stability issues for major customer facing and customer service applications and websites. A ground-floor member of a new Application Support group, I quickly learned a large, multi-faceted system and drove improvements to the software and database code. Writing code and submitting detailed suggestions to the application development team, working with them to develop, enhance, test, and verify changes and document data flows. Working alongside business users that would report problems, I identified solutions and followed through to see that the solutions were implemented, maintaining communication with the users to ensure they were satisfied. After these improvements and the introduction of extensive business logic alerting, a 66% reduction in down-time year over year was realized, saving up to a million dollars in call center expenses.

Professional Software Integration

Successfully integrated a third-party application to provide a real-time security badge printing and tracking solution for four college campuses. I took a single user, Microsoft Access based, solution suffering from access contention and data loss to a true multi-user, transactional solution on Microsoft SQL Server. Implementing the software upgrade, converting data, and writing custom software to support ERP data integration and reporting I eliminated contention and data loss while greatly enhancing customer service and creating a pleasant workflow for the users. I created documentation that enabled help desk technicians to support the software after my departure.