Sean Anderson

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

Senior Data Engineer / DBA / 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 inspired by the same desires to learn, create, and contribute.

TECHNOLOGY

Big Data: Hadoop, Hortonworks, Yarn, Hive, Ambari, Tez, MapReduce
Cloud: Azure: HDInsight, Application Insights, VM provisioning, Resource Templates (ARM), SQL Database, Blob Storage, Data Lake Store (ADLS), Cloud Administration, AWS: DynamoDB, API Gateway, S3, EC2
Databases: SQL Server 2005, SQL Server 2008, SQL Server 2012, SQL Server 2014, MySQL, MongoDB
Coding: TSQL, SQL, C# .NET, Xamarin.Forms, Java, Python, PowerShell, PHP, VBA, RabbitMQ
Distributed: RabbitMQ message broker
Tools: SQL Server Integration Services (SSIS), PowerBI, Excel XLS(Macros/VBA), Visual Studio, TFS, Team Services, SQL Server Management Studio (SSMS), SQL Server Data Tools (SSIS/SSRS), Fiddler, SOAP UI, MongoChef, vSphere (Web/Desktop), Power CLI VMware automation
Application Servers: Internet Information Server (IIS 7, IIS 8)

Professional Experience

2013 to Present: RushCard / Green Dot Corporation (Blue Ash, Ohio)

Software Engineer / DBA / Principal Systems Engineer

- Hands-on data platform owner responsible for data engineering and for ensuring fault-tolerance, performance, and recoverability of a $150MM 24/7 high volume, transactional, multi-terabyte database environment (10 TB, 16 Servers, 16 AG, 200 Cores)

- Led cloud-based big data (data lake) project using HDInsight (Hadoop) and Data Lake Store on Microsoft Azure to create a centralized, standardized repository for all vendor, partner, and production data to prepare for the use of analytics in applications and reporting

- Created and revised star-schema data models and developed ETL/ELT for data warehouse using SSIS on SQL Server 2008 - 2014

- Principal data modeler, worked with developers to create optimal data models in Entity Framework (C# objects mapped to SQL tables)

- Successfully launched a data integration project for a new acquisition to ensure data sources were available to business ops teams

- Built a robust FTP/SFTP application in SSIS to handle all vendor file uploads and downloads in production

- Led a successful large-scale data conversion 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, proper use of table variables and indexed temp tables, stored procedures, functions, C# CLR, and table value functions

- Developed many SSIS projects/packages for converting data and loading tables with billions of rows utilizing ETL and ELT methods

- Developed applications and scripts using C#, PowerShell, and T-SQL to cleanse and verify proper file formats and process batches of data against SOAP/REST 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 in C# and SQL. Conducted developer OWASP training.

- Successfully spearheaded and completed a project to improve our production distributed message broker system (RabbitMQ) for high availability and zero downtime maintenance; reducing resource utilization by 50% while improving reliability and performance (automated deployment and configuration)

- Identified root cause of application bugs and committed code to resolve issues (C# web and console applications)

- Documented complex systems and data flows using Visio for technical and non-technical audiences

- Handled a variety of data analysis efforts with executives and department heads to mitigate issues with partners and regulatory agencies

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

- Lead network security engineer for student data—managed complex 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 Technology representatives, to fill a gap in Turning Point response system reporting metrics to track student progress throughout a semester with aggregated and visualized results.

FORMAL EDUCATION

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

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 (HD), I built a visualization and reporting solution for end users by capturing email status messages (JSON) from the SendGrid web hook API, and feeding those through AWS Lambda (Node.js) 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 perform ETL. 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 Finance 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 at MetaBank.

API Redesign
Re-engineered Green Dot API integration 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 point-of-sale stations. The previous 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 issues and provided work-arounds and operation support to customer service while I redesigned the web service data flow. Lacking a formal project management team, I recruited another developer to help implement, test, 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 high-volume customer facing and internal customer service web applications. As 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. I would write code or submit detailed documentation 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 then follow through to see that the solutions were working to their satisfaction. 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.