International Journal of Science and Research (IJSR)
ISSN (Online): 2319-7064
Index Copernicus Value (2016): 79.57 | Impact Factor (2017): 7.296

A Review of Incremental Data Load on Premises to Azure

Tejaswee Poya[1], Rahul Mishra[2]

[1]M.Tech Scholar Computer Science, GD Rungta College of Engineering & Technology, Bhilai (C.G)
Tejasweepoya[at]gmail.com
[2]Assistant Professor Computer Science, GD Rungta College of Engineering & Technology, Bhilai (C.G)
rahul.mishra[at]rungtacolleges.com

Abstract: ***This project explains to overcome with the disadvantages of legacy system to load data from on premises to azure. On legacy system Cloud based application should be offline while loading data into cloud which causes application unavailability, low customer satisfaction and impacting business revenue. This project helps in understanding the way of getting the advantages of the Cloud without having to migrate the entire database by just maintaining a copy of the data that needs to be accessible to internet-based users in Windows Azure SQL Database. The project allows SQL Server Management Studio (SSMS) to interface with the database engine and use Business Intelligence Development Studio (BIDS) to create a SQL Server Integration Services (SSIS) Package. SSIS will provide the mechanism necessary to upload the data from the SQL Server instance to SQL Azure and synchronize the uploaded data with existing data.***

Keywords: Azure Cloud, SSIS Package, SSMS, SQL

1.Introduction

Existing System

At present during the monthly deployments/releases, all the required SQL code which includes bug fixes and other additions to be manually executed and then bring the application offline or down for the required data to be pushed through jobs to Azure SQL database and then once all the new data is pushed to Azure, perform a manual data check and bring the application online.

Drawbacks of Existing System

In existing system once the application is offline; users are impacted and will be able to access the application till the jobs and the required data is pushed to Azure; so due to these drawbacks user are very dissatisfied and the interest towards using the application goes down.

New System and Enhancements

The goal of our solution is to use SSIS to copy data from on-premises SQL Server Database to a SQL Azure Database and to keep the data synchronized. In our SSIS package, we'll issue T-SQL MERGE statements against the SQL Azure Database to synchronize the data between the staging tables and target tables. We could use SSIS to stage the data in memory and then synchronize it with the data in the SQL Azure Database.

2.Literature Survey

Business will provide monthly data on excel sheet and there will be both old data (previously provided data) and new data (Current month data). On this monthly data ETL process will be applied to load data on- premises SQL database. Then business rule will be applied and cloud application will be offline to truncate and load Azure SQL database.

Data load on-premises to Azure database application enhancement will maintaining multiple set of same table on local SQL database and Cloud database to increase customer satisfaction, efficiency and to reduce data load times.

Business will provide monthly data on excel sheet and there will be both old data (previously provided data) and new data (Current month data). Excel data will be loaded into first set of table in on-premises SQL database ; will take a copy of Azure SQL database table into on-premises SQL database and will be maintaining in second set of table . Both first and second set of table will be compared and the difference (new data or current month data) will be stored in third table.

Business rule will be applied only on current month data and will be loaded into secondary table of cloud database. Batch job will load the data from secondary table to primary Azure SQL database table.

This will help us to load the monthly data into cloud database without bringing application server offline with minimal load time.

3.Problem Definition

Data load on-premises to Azure system load monthly data from business provided excel sheet to Cloud Azure Database for application use.

On existing system below problems are identified:

- Business is providing monthly data excel sheet which contains previously loaded data and current month transaction data. Application will load monthly data into on-premises SQL databases and apply business rule on entire data (previous loaded data and current month data); which is unnecessary and time consuming process.

- To load data on cloud; application need to be offline and entire truncate and load process will take minimum 3-4 hours. During this period no cloud based application will be available for end user use which is impacting company revenue and giving low customer satisfaction for the product.
- While loading data into Azure database if any issue occur it will take 2-3 hours to revert the changes.

4.Methodology

Functional Requirements

1. Truncate the staging table of SQL Azure.
2. Load the existing data from Azure main tables to on premises staging tables.
3. Apply the bug fixes or requirements code on the data available in on premise main tables.
4. Compare the data present in on premise staging tables and the main tables.
5. Push the incremented data, post comparison to Azure staging tables.
6. Compare the data present in the Azure staging tables and the main tables.
7. Push the incremented data, post comparison to Azure main tables.
8. Validate the changes made in backend are reflecting as per expectation in the application.

Architectural diagram

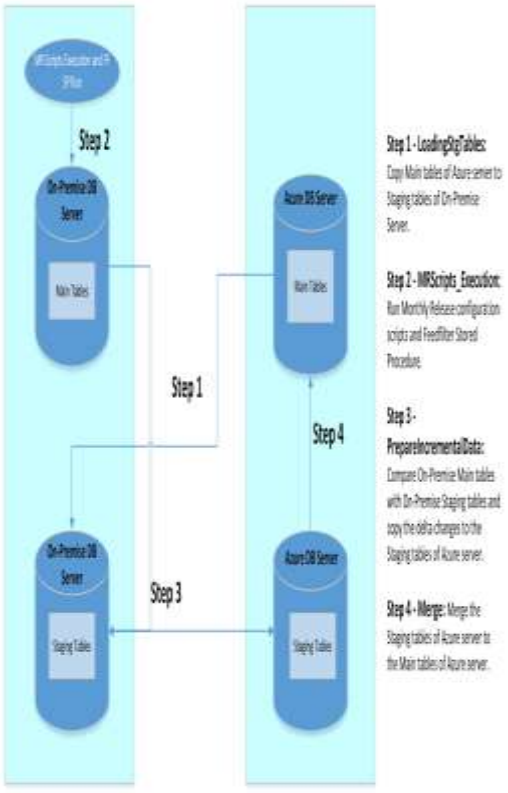

Figure 1: Architectural Diagram of Incremental Data Load on premises to Azure

5.Results

We do not need to bring the application offline while loading data on cloud and there is no need to truncate and load unnecessary data again and again.

6.Conclusion

After completion of data load from on-premises to azure application; to load monthly data on SQL azure database we no need to bring internet based application offline.

Following are the outcome of this application:

1. **Application availability:-** Cloud based application will be available to user.

2. **Minimum load time:** - During monthly data load; only new data will be loaded in SQL azure database, which will decrease the load time.

7.Future Scope

During data load if any existing records are accessed by business to generate the reports and the same records are modified in the backend and if that time customer access those data then below popup message will display to customer-"Value has been updated kindly refresh to get the updated and accurate report".

Benefits:

- In proposed system we do not need to bring the application offline while loading data on cloud.
- Customer satisfaction increases.

Limitation:

- In existing system we need to bring the application offline while loading data on cloud.

References

[1] http://blogs.msdn.com/b/zkap/archive/2012/04/01/moving-an-on-premise-database-from-sql-server-to-sql-azure-using-ssis.aspx
[2] https://www.simple-talk.com/cloud/cloud-data/uploading-data-to-windows-azure-sql-database/
[3] http://www.codeproject.com/Articles/430628/Migrating-data-from-on-premise-SQL-Server-to-SQL-A