We built a case for migrating SSIS packages to Data Factory, but a popular question, that is on everyone’s mind is, does Data Factory replaces SSIS?

To answer this question, Let’s first understand how the data scenario has evolved.

Traditional Data warehousing projects require creating ETL Packages that extract data from operational data sources, transform the data with SSIS and load the data into a Datawarehouse for reporting.

With the advent of Big Data, an organizations’ data is not limited to its operational databases. Today an organization generates a large volume of data that comes from a variety of data sources like Mobile phones, emails, Sensors, web pages, social media, web server logs, e-commerce apps and clickstream data at a very rapid pace.

The Data that is generated is vital to every app and experience that developers build today. With an increasing amount of data, organizations do not want to be tied down by increasing infrastructural costs that comes with using SSIS on-premise. Developers are realizing the need to start moving their workloads to the cloud to take advantage of its massive scale and flexibility.

So, you generate a large amount of data, but you may not have any idea about how to leverage the data to generate insights or answer your business questions, so rather than transforming the data and loading it into a data warehouse (like you did earlier), there is an increasing need to just dump the data into a data lake, so that it may be used in the future. And as the cloud drastically decreased the storage costs, you can afford to store all the data in a data lake.

Next, you may run big data jobs that process the data with big data tools like HDInsights and Data Lake Analytics to get the data into shape according to the Schema defined in a cloud data warehouse.

And to automate the entire process in the cloud, you need Data Factory.

And that’s the scenario of data today.

Coming to the question, whether Data Factory replaces SSIS, you can continue to use SSIS for the power of transformations, as many of them are not available in Data Factory, and also for the SQL Server investments you already made, And use Data Factory for handling big data workloads so you can inexpensively deal with Big Data in the cloud.

Think of Data Factory as a complementary service to SSIS. You can mix-n-match various SSIS and Data Factory components to get the best of both ETL and ELT worlds.