Databricks to Synapse Workspace Migration

Challenge

In response to a growing realization that the client's utilization of Databricks was falling short of optimal efficiency, the client made a strategic decision to transition to Azure Synapse. Upon making the decision to migrate, they recognized a need for expertise in assessing their current environment and setting up Azure Synapse to streamline their operations. Furthermore, they had the challenge of converting many existing Azure Databricks notebooks. In order to proceed with this project, the client turned to INSPYR Solutions for expert guidance.

Solution

Recognizing the need for expertise to migrate to a new platform, the client sought the assistance of INSPYR Solution's Data & Analytics Practice. INSPYR Solutions assigned one of its most senior Data & Analytics Practice leaders to the project. As an expert in this area, the consultant was able to analyze the project and performed the following steps to give the client a roadmap to success:

Step 1. Assess the current Databricks notebooks. This consisted of checking for dataflow tasks to determine if these can be added to the existing Azure Data Factory pipelines and implementing a roadmap for converting the Databricks notebooks to Synapse notebooks. The INSPYR Solutions team also completed a breakdown of pipelines to check if they could be combined to iterate through parameters/variables,

> By leveraging our team of experts, the project was delivered on time and at a significantly lower cost than had the client performed the work internally.

inspyrsolutions.com

as well as a breakdown of datasets to create unique dynamic datasets per source/landing zone.

Step 2. Prepare the different environments for the Azure Synapse workspace resources. We assisted in prepping the resource groups and necessary Azure Identity Management permissions required to set up the Synapse workspace with a dedicated SQL pool. All names were standardized with respect to industry standards and used templates to deploy services to ensure that moving from environments would be easier. We also got the network teams involved and helped them understand the cloud architecture better.

Step 3. Implement DevOps practices. The INSPYR Solutions team implemented practices to streamline moving code from development to testing, and then to production at a very efficient rate by using CICD within Azure DevOps. This involved automated execution of test cases (YAML) within the CICD process to avoid processes remaining in code review as a bottleneck.

Step 4. Understand the cloud cost model. The clients were made aware of the cost that would be incurred for the services that were deployed during the development phase and implemented a dashboard that monitors the usage stats in a timely manner for current and future project expansion. Optimization efforts were implemented to avoid costs in the migration process.

Step 5. Successful project implementation. The INSPYR Solutions team was deployed to implement the conversion of Databricks notebooks into Azure Synapse workspace notebooks. Files were dropped from REST API calls and landing zones into a data lake, which in turn was converted using Synapse notebooks to be staged in the dedicated SQL pool and consumed into the Azure Synapse data warehouse. The team also created templates for most of the data solutions for ease of replication and use within the organization.

Outcome

The result was a unified analytics platform that was scalable, intelligent, and cost-effective. By leveraging our team of experts, the project was delivered on time and at a significantly lower cost than had the client performed the work internally. By saving the client money and time, INSPYR Solutions provided an efficient solution tailored to the client's business needs.

Client Profile

The client is a physician-led healthcare organization that partners with hospitals, health systems, and healthcare facilities to offer clinical services spanning the women's and children's continuum of care. The client is an experienced clinical specialist trusted by patients, hospitals, and referring physicians to take great care of each patient, every day and in every way.

Technologies Supported/ Created

Azure Synapse, Databricks, Data Lake, Data Factory.

About INSPYR Solutions

Technology is our focus and quality is our commitment. As a leading expert in delivering flexible technology and talent solutions, we strategically align industry and technical expertise with our clients' business objectives and cultural needs. Our tailored offerings include a wide variety of professional services, project solutions, managed services, and talent resources, all bolstered by our strategic partnerships with cutting-edge technology services. By always striving for excellence and focusing on the human aspect of our business, we work seamlessly with our talent and clients to match the right solutions to the right opportunities. Learn more about us at www. inspyrsolutions.com.