**Migrating MongoDB databases to Azure CosmosDB**

**Project Topic: Azure Database Migration Service**

**Project Topic Description**: Typically is focused on the migration of database schema and data from one database format to a specific database technology in the cloud.

**Problem Description**: We have a few MongoDB instances on-premise or on-virtual machines in the cloud. We would like to take advantage of Azure CosmosDB cloud service for DR, replication, scalability, and high availability. There is a need here to migrate of MongoDB databases to Azure CosmosDB database service.

**Sample Data:**

1. Northwind csv files from <https://github.com/tmcnab/northwind-mongo/archive/master.zip>. Northwind sample database is from Microsoft Access product. There are 11 tables with 3308 rows.
2. mongodb\_cases.json from <https://community.jaspersoft.com/wiki/mongodb-city-cases-example>. There are 10,000 json documents in this dataset. It is about cases in a city.

**Hardware and Software Used:**

1. Windows 7 and MongoDB Enterprise Server 3.6.1.

**High-Level Steps**

1. Install MongoDB, clean the data and load sample data into databases
2. Backup the databases in MongoDB which are ready for loading to CosmosDB
3. Create a CosmosDB account on portal.azure.com
4. Determine restore paramters for batchsize and number of workers.
5. Restore the backup to CosmosDB using MongoDB API. Pre-create CosmosDB database and collections if the default throughput 1000 RU/s is too slow.

**Issues and Lessons learned**

1. Uploading JSON documents to CosmosDB is slow. For a simple document, need about 0.072s which is equivalent of 14 documents/s. For a million documents, the rough estimate is that we need about one day.
2. We could increase the throughput above 10,000 RU/s to spead up the upload, but we need to implement sharding. When the collection size is larger than 10GB, sharding is also required. As JSON documents are schema free and no required constraint for a partition key as required by sharding. Thus, the application needs to ensure that there are candidate partition keys in the document.

**YourTube URLs**

2 min Video: <https://youtu.be/9vDRBgPeT50> 15 min Video: https://youtu.be/KHFuEFOMIH4