Microsoft Azure - Starter Kits for Partners

Hands on Lab

Advanced Analytics Scenario

Create an Azure SQL Data Warehouse and Visualize with Power BI

Last Update: Aug 2016





**MICROSOFT MAKES NO WARRANTIES, EXPRESS, IMPLIED OR STATUTORY, AS TO THE INFORMATION IN THIS DOCUMENT.**

The information contained in this document represents the current view of Microsoft Corporation on the issues discussed as of the date of publication. Because Microsoft must respond to changing market conditions, it should not be interpreted to be a commitment on the part of Microsoft, and Microsoft cannot guarantee the accuracy of any information presented after the date of publication.

Complying with all applicable copyright laws is the responsibility of the user. Without limiting the rights under copyright, no part of this document may be reproduced, stored in or introduced into a retrieval system, or transmitted in any form or by any means (electronic, mechanical, photocopying, recording, or otherwise), or for any purpose, without the express written permission of Microsoft Corporation.

Microsoft may have patents, patent applications, trademarks, copyrights, or other intellectual property rights covering subject matter in this document. Except as expressly provided in any written license agreement from Microsoft, the furnishing of this document does not give you any license to these patents, trademarks, copyrights, or other intellectual property.

The descriptions of other companies’ products in this document, if any, are provided only as a convenience to you. Any such references should not be considered an endorsement or support by Microsoft. Microsoft cannot guarantee their accuracy, and the products may change over time. Also, the descriptions are intended as brief highlights to aid understanding, rather than as thorough coverage. For authoritative descriptions of these products, please consult their respective manufacturers.

© 2016 Microsoft Corporation. All rights reserved. Any use or distribution of these materials without express authorization of Microsoft Corp. is strictly prohibited.

Microsoft and Windows are either registered trademarks of Microsoft Corporation in the United States and/or other countries.

The names of actual companies and products mentioned herein may be the trademarks of their respective owners.

Contents

[Overview 4](#_Toc459726295)

[Objectives 4](#_Toc459726296)

[Prerequisites 4](#_Toc459726297)

[Exercises 5](#_Toc459726298)

[Infrastructure Provisioning 5](#_Toc459726299)

[Exercise 1: Creating an Azure SQL Data Warehouse 5](#_Toc459726300)

[Exercise 2: Loading data into Azure SQL Data Warehouse using Polybase 6](#_Toc459726301)

[Exercise 3: Azure SQL Data Warehouse integrated with Power BI 6](#_Toc459726302)

## Overview

In this lab, we walk you through building Power BI Reports and Dashboard using SQL Data Warehouse (SQL DW) for a publicly available dataset -- the NYC Taxi Trips dataset which was stored in Azure Blob Storage.

**Estimated time** to complete this lab: **180 minutes**.

**Audience**: IT Pro, Architect, Application Owners and Developers

### Objectives

In this hands-on lab, you will learn how to:

1. Create an Azure SQL Data Warehouse
2. Load data from a cloud data store using Polybase
3. Visualize the data in Power BI

Each task is separated into ‘Parts’ in this demo script.

The following messaging points will be delivered as part of the tasks above:

1. Ease of deployment
2. Cloud-based data warehouse service
3. Region availability
4. Elastic-scale capability including pause
5. Cost friendly
6. Control over compute separate to storage
7. Data encryption
8. Compliance and auditing
9. Integrated experience
10. Common tooling
11. Rapid and rich visualizations

### Prerequisites

The following is required to complete this hands-on lab:

* A Microsoft Azure subscription - [sign up for a free trial](http://aka.ms/WATK-FreeTrial)
* [Azure PowerShell 0.7.4  or higher](https://azure.microsoft.com/en-us/documentation/articles/powershell-install-configure/)
* Visual Studio 2015

## Exercises

## Infrastructure Provisioning

### Exercise 1: Creating an Azure SQL Data Warehouse

1. ***Pre-requisites***
2. 1. Azure Subscription with $50 of credit
3. ***Scenario***

Today businesses are collecting more information than ever before to drive insight based on data. To succeed we require three things – scaled data stores, tools to deliver insight and an agile approach to get the business answers quickly.

Let’s start by addressing agility and scale. The Azure SQL Data Warehouse is the industry’s first enterprise class cloud-based data warehouse with dynamic scale. What used to take days or weeks to achieve, we can literally do in minutes.

1. ***Steps***

Creating a data warehouse couldn’t be simpler.

We simply specify the data warehouse details (server, resource group, database, region, etc.).

If we look at the process, we don’t need to select the size of your datawarehouse on disk – Azure grows it dynamically as you need it. This is what is called ‘elastic scale’ and Microsoft are the first to bring this capability to the cloud data warehouse.

We are building an MPP data warehouse and so query speed is super important. By simply moving the slider to the right we add more CPUs into the mix and get the query speed I need. That’s a whole lot simpler than building a network of machines.

The default is 400 DWU (Data Warehousing Unit) and is a good starting point for creating a data warehouse. More information on DWU and concurrency can be found at the following URL:

<https://azure.microsoft.com/en-gb/documentation/articles/sql-data-warehouse-develop-concurrency/>

We can also programmatically alter the amount of compute using PowerShell, T-SQL or REST. More information can be found at the following URL:

<https://azure.microsoft.com/en-gb/documentation/articles/sql-data-warehouse-manage-compute-overview/>

**Step by step:** <https://azure.microsoft.com/en-us/documentation/articles/sql-data-warehouse-get-started-provision/>

### Exercise 2: Loading data into Azure SQL Data Warehouse using Polybase

1. ***Pre-requisites***
2. Visual Studio 2015 Community Edition has been provided for this part of the lab.
3. We have provided a template script to perform all of the stages of this lab. The script can be found within the virtual machine at the following location:
4. **Sample Data folder**
5. ***Scenario***

So we have created our data warehouse and now we want to load data into it. We can do this through the traditional ways of ETL and tooling such as SQL Server Integration Services or third-party tooling. However today we are going to use Polybase; a Big Data technology created by Microsoft in partnership with Gray Systems Laboratories in Cambridge which allows querying of data located in Hadoop, Azure HDInsight or Azure Blob storage.

Our source data is located in Azure Blob Storage in the West US region and we will query and load this data using Polybase. The data is based on the NYC Taxi public dataset which was originally obtained from <http://chriswhong.com/open-data/foil_nyc_taxi/>.

**Step by step:**

<https://azure.microsoft.com/en-us/documentation/articles/sql-data-warehouse-load-polybase-guide/>

<https://azure.microsoft.com/en-us/documentation/articles/sql-data-warehouse-get-started-load-with-polybase/>

### Exercise 3: Azure SQL Data Warehouse integrated with Power BI

1. ***Pre-requisites***

A Power BI subscription is required (free or premium)

1. We will use the **Microsoft Edge** browser for this lab part.
2. 1. **Browser** (Separate private Mode Browser tabs)
3. Home Page <http://portal.azure.com>
4. Username: **<your live account\azure login>**
5. Password: **<your password>**
6. Home Page http://www.powerbi.com
7. Username: [??????@**yourcompany.com**](mailto:??????@yourcompany.com)
8. Password: ??????
9. ***Scenario***

Storing data at scale is one thing – but what we really want to do is get insight from that data. With PowerBI we can gain insight into any data to make informed business decisions efficiently and safely.

**Step by step:**

<https://azure.microsoft.com/en-us/documentation/articles/sql-data-warehouse-integrate-power-bi/>

* 1. **Appendix A: Structure**
  2. **DB Structure**
  3. Custom schema (NYC Taxi dataset)
  4. **West US Data Centre**:
* Resource Group = rgmyclouddw
* SQL DW Server = dwwestus.database.windows.net
* Database = myclouddw