

DP 200 - Implementing a Data Platform Solution

Lab 9 - Monitoring and Troubleshooting Data Storage and Processing

Estimated Time: 75 minutes

Pre-requisites: It is assumed that the case study for this lab has already been read. It is assumed that the content and lab for module 1 to module 7 has been completed.

Lab files: The files for this lab are located in the *Allfiles\Labfiles\Starter\DP-200.9* folder.

Lab overview

The students will be able to define a broad monitoring solution that can help them monitor issues that can occur in their data estate. The student will then experience common data storage issues and data processing issue that can occur in cloud data solution. Finally they will implement a disaster recovery approach for a Data Platform technology.

Lab objectives

After completing this lab, you will be able to:

1. Explain the monitoring capabilities that are available
2. Troubleshoot common data storage issues
3. Troubleshoot common data processing issues
4. Manage disaster recovery

Scenario

As the Senior Data Engineer at AdventureWorks you have been tasked with defining the standard operating procedures for monitoring that data estate within the organization. You will start by defining the monitoring tools that will be used to support the approach.

You will then explore some of the common data storage and data processing issues that can occur during the normal operation of your infrastructure.

There are concerns around the recovery of the Products database that is stored in the awcodbstudxx Cosmos DB. The IS department has asked you to provide high level steps that would be taken in the event that the products database has become unavailable through an accidental deletion or removal of the database.

At the end of this lab, you will have:

1. Explained the monitoring capabilities that are available
2. Troubleshoot common data storage issues
3. Troubleshoot common data processing issues
4. Managed disaster recovery

Exercise 0: Issue review

IMPORTANT: In this lab, you will refer to the issue(s) that you have encountered in any provisioning or configuration tasks in all of the labs in the document located at *\Labfiles\DP-200-Issues-Doc.docx*. Open this document up ready for the following exercises.

Exercise 1: Explain the monitoring capabilities that are available

Estimated Time: 15 minutes

Group exercise

The main task for this exercise are as follows:

1. Defining a corporate monitoring approach.
2. The instructor will discuss the findings with the group.

Task 1: Defining a corporate monitoring approach.

1. From the lab virtual machine, start **Microsoft Word**, and open up the file **DP-200-Lab09-Ex01.docx** from the **Allfiles\Labfiles\Starter\DP-200.** folder.
2. Spend **10 minutes** in a group discussing and identifying the monitoring tools that would be the most useful tool for you within your organization. Find two examples and outline your justification.

Task 2: Discuss the findings with the Instructor

1. The instructor will stop the group to discuss the findings.

Result: After you completed this exercise, you have created a Microsoft Word document that identifies the monitoring tools that would be the most useful tool for you within your organization.

Exercise 2: Troubleshoot common data storage issues

Estimated Time: 20 minutes

Group exercise

The main tasks for this exercise are as follows:

1. Assessing issues that are data storage issues.
2. The instructor will discuss the findings with the group.

Task 1: Assessing Data and Storage Security Hygiene.

1. In the **\Labfiles\DP-200-Issues-Doc.docx** document, share your findings and work with the group to identify which of the issues are data storage issues.
2. Work with the group to see if there are **common data storage issue** that the group experienced while setting up the various data platforms.
3. As a group select two data storage issues that you have identified, and log them into the document **DP-200-Lab09-Ex02.docx**.

Task 2: Discuss the findings with the Instructor

1. The instructor will stop the group to discuss the findings.

Result: After you completed this exercise, you have created a Microsoft Word document that lists two data storage issues.

Exercise 3: Troubleshoot common data processing issues

Estimated Time: 20 minutes

Group exercise

The main tasks for this exercise are as follows:

1. Assessing issues that are data processing issues.
2. The instructor will discuss the findings with the group.

Task 1: Assessing Data and Storage Security Hygiene.

1. Review with the Group the questions outlined in **DP-200-Lab09-Ex03.docx**
2. Work with the group to see if there are **common data processing** issues that the group experienced while setting up the various data platforms.
3. As a group select two data storage issues that you have identified, and log them into the document **DP-200-Lab09-Ex03.docx**.

Task 2: Discuss the findings with the Instructor

1. The instructor will stop the group to discuss the findings.

Result: After you completed this exercise, you have created a Microsoft Word document that lists two data processing issues.

Exercise 4: Manage disaster recovery

Estimated Time: 20 minutes

Group exercise

The main tasks for this exercise are as follows:

1. Manage Disaster Recovery.
2. The instructor will discuss the findings with the group.

Task 1: Manage Disaster Recovery

1. From the lab virtual machine, start **Microsoft Word**, and open up the file **DP-200-Lab09-Ex04.docx** from the **Allfiles\Labfiles\Starter\DP-200.9** folder.
2. As a group, spend **10 minutes** discussing and listing the data requirements and data structure that your group has identified within the case study document.

Task 2: Discuss the findings with the Instructor

1. The instructor will stop the group to discuss the findings.

Result: After you completed this exercise, you have created a Microsoft Word document that outlines the high level steps required to restore the products database.