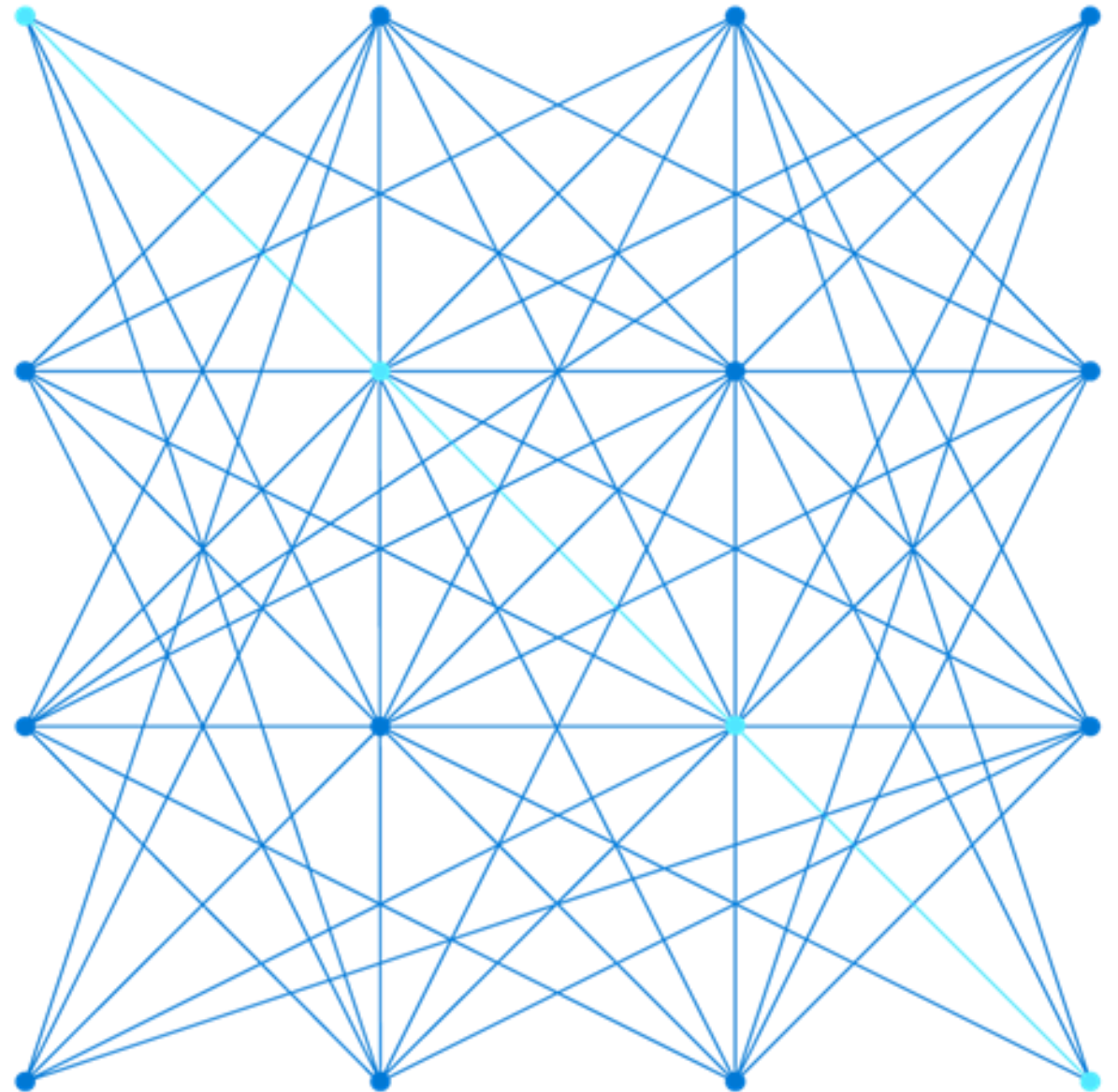


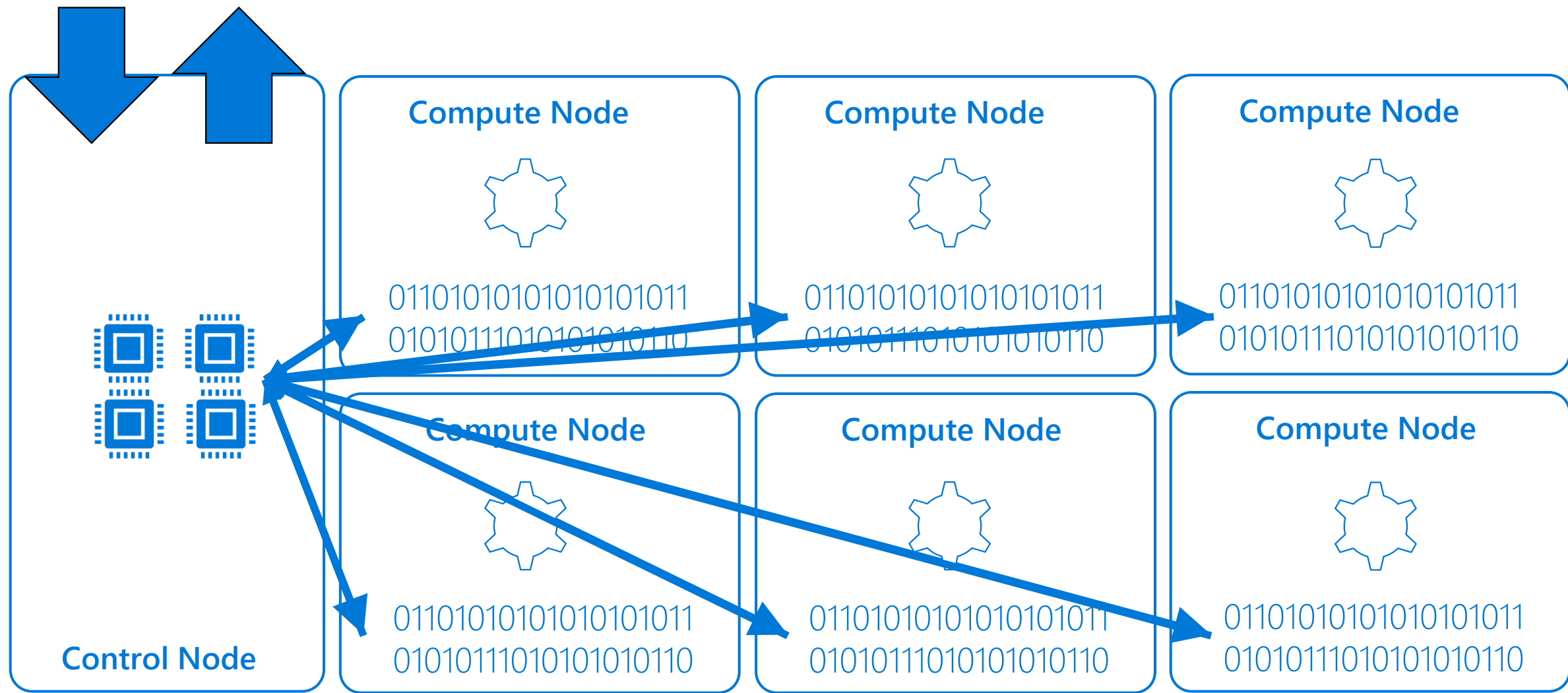
DP-203T00: Ingest and load data into the data warehouse



Lesson 01: Use data loading best practices in Azure Synapse Analytics



Dedicated SQL Pool architecture revision



Understand data load design goals

- Where is my data coming from?
- Is the data nett new? or do you receive changes from existing datasets?
- How often is the data being refreshed, added to or replaced?
- What formats are the data coming in?
- Is the data ingestible as-is? or are transformations and cleansing tasks required?
- Which takes priority, loading or querying/analysis?

Manage singleton updates



Control Node

Compute Node



01101010101010101011
01010111010101010110

Compute Node



01101010101010101011
01010111010101010110

Compute Node



01101010101010101011
01010111010101010110

Compute Node



01101010101010101011
01010111010101010110

Compute Node



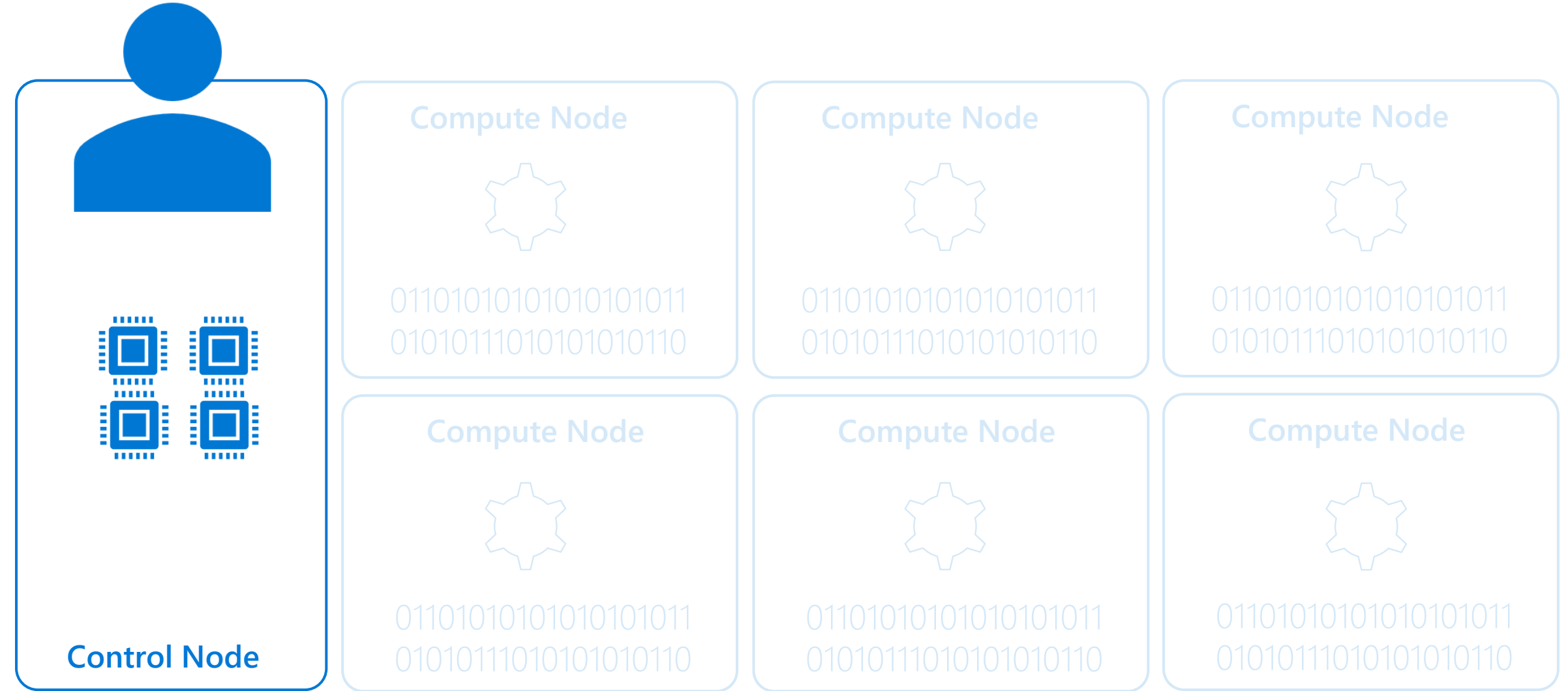
01101010101010101011
01010111010101010110

Compute Node

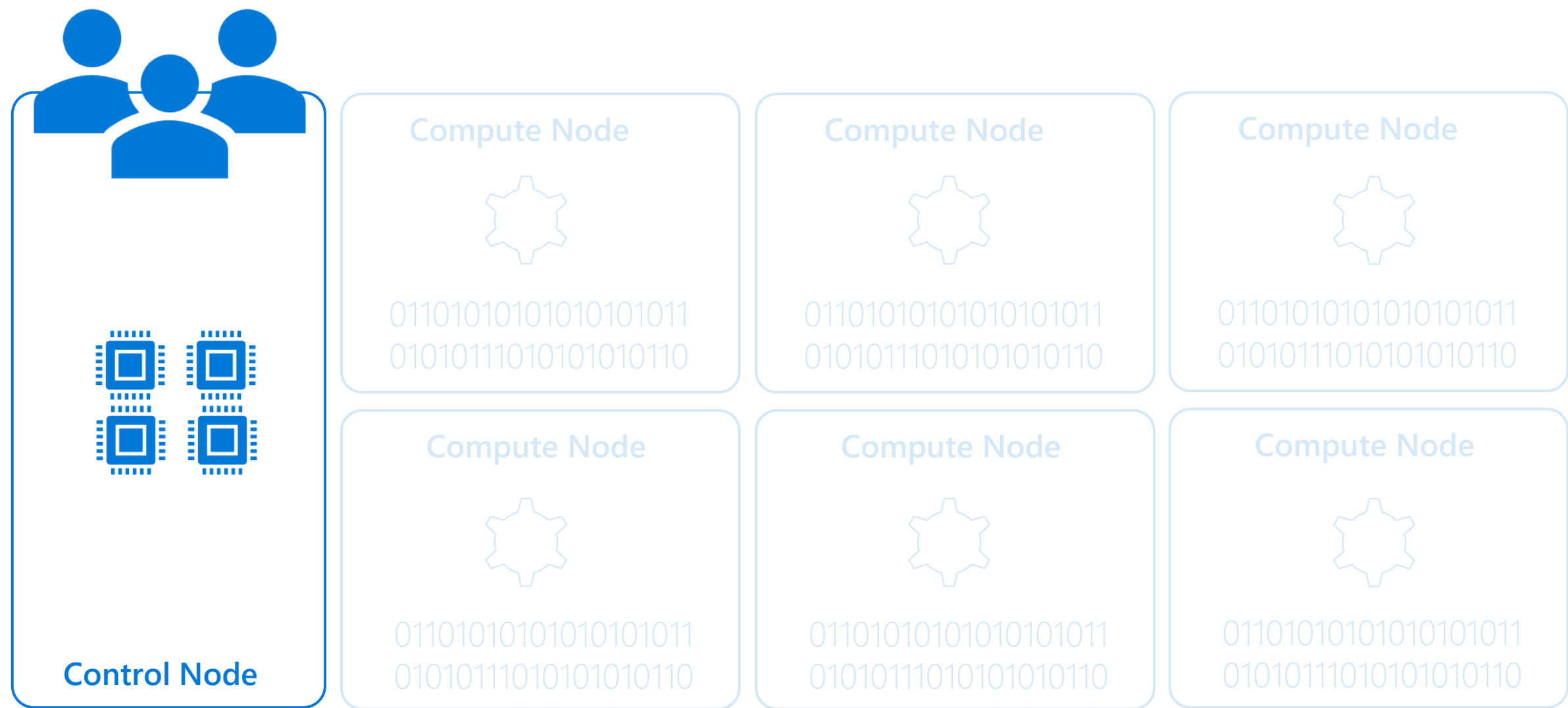


01101010101010101011
01010111010101010110

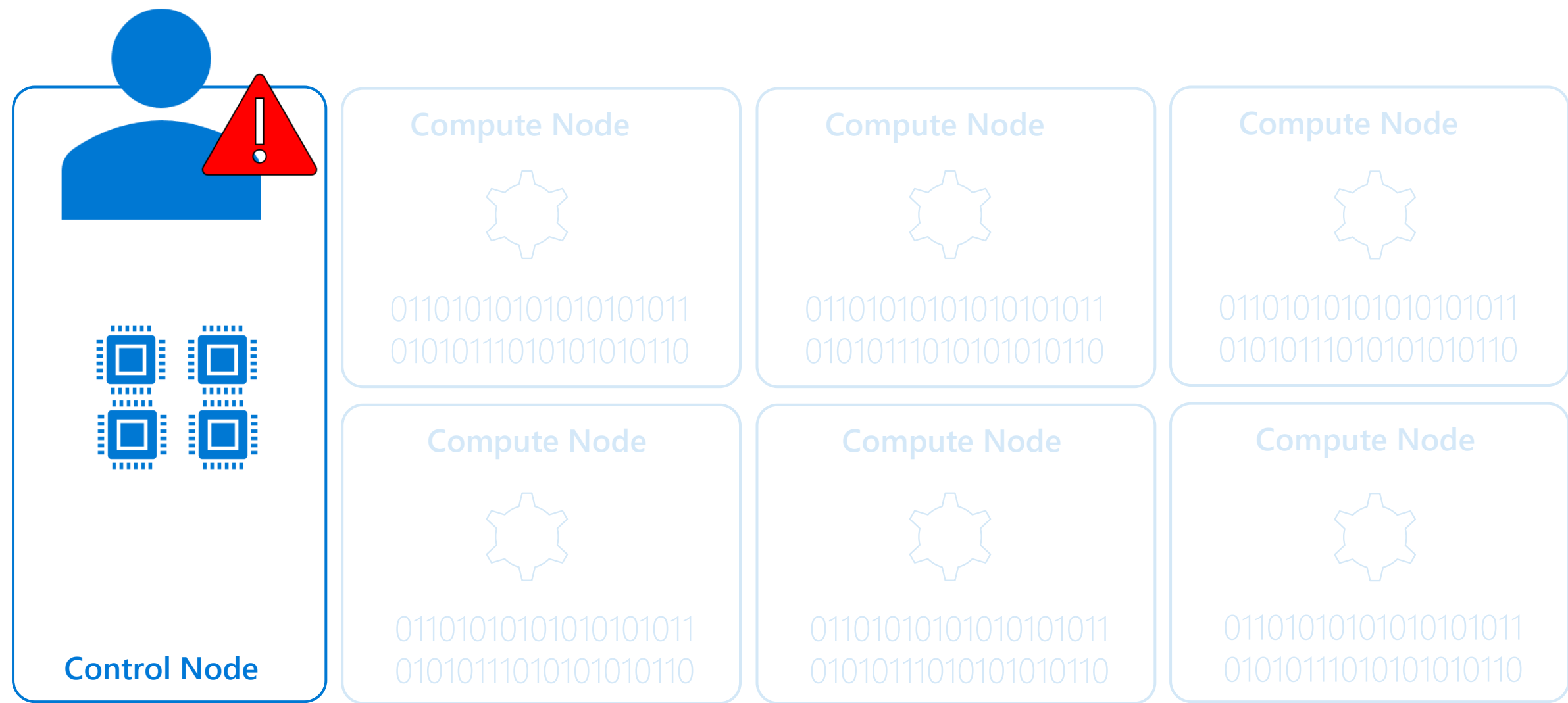
The diagram illustrates a control node and six compute nodes. On the left, a large rounded rectangle represents the **Control Node**. It contains a blue icon of a person at the top, four blue square icons with internal patterns in the middle, and the text **Control Node** at the bottom. To the right of the control node are six smaller rounded rectangles arranged in a 2x3 grid, each representing a **Compute Node**. Each compute node contains a blue gear icon, the text **Compute Node**, and two lines of binary code: 01101010101010101011 and 01010111010101010110.



Manage concurrent access to Azure Synapse Analytics

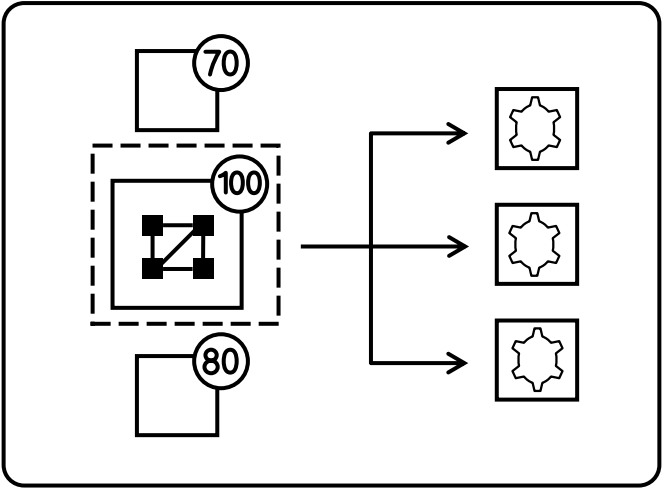


Implement Workload Management

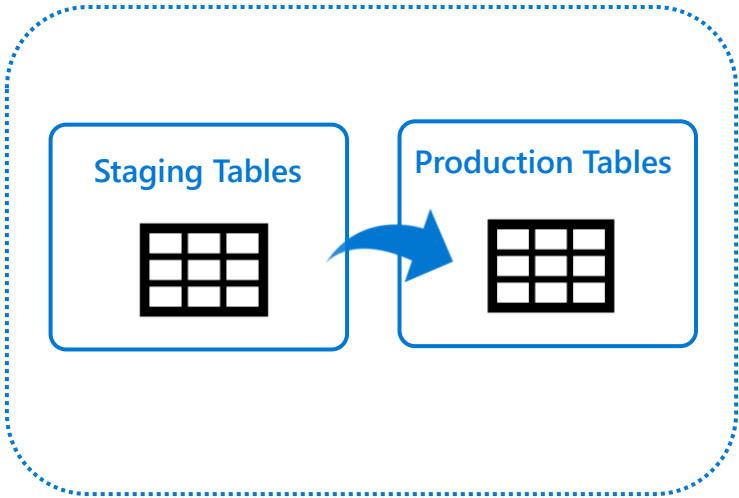


Use PolyBase, the Copy command or the Copy Activity

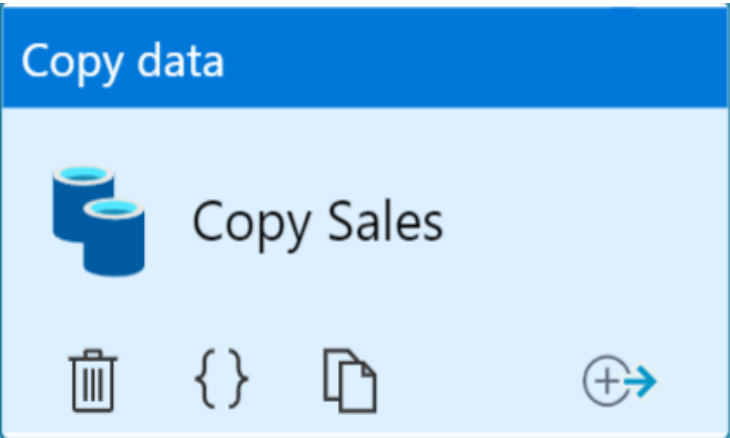
PolyBase



Copy command



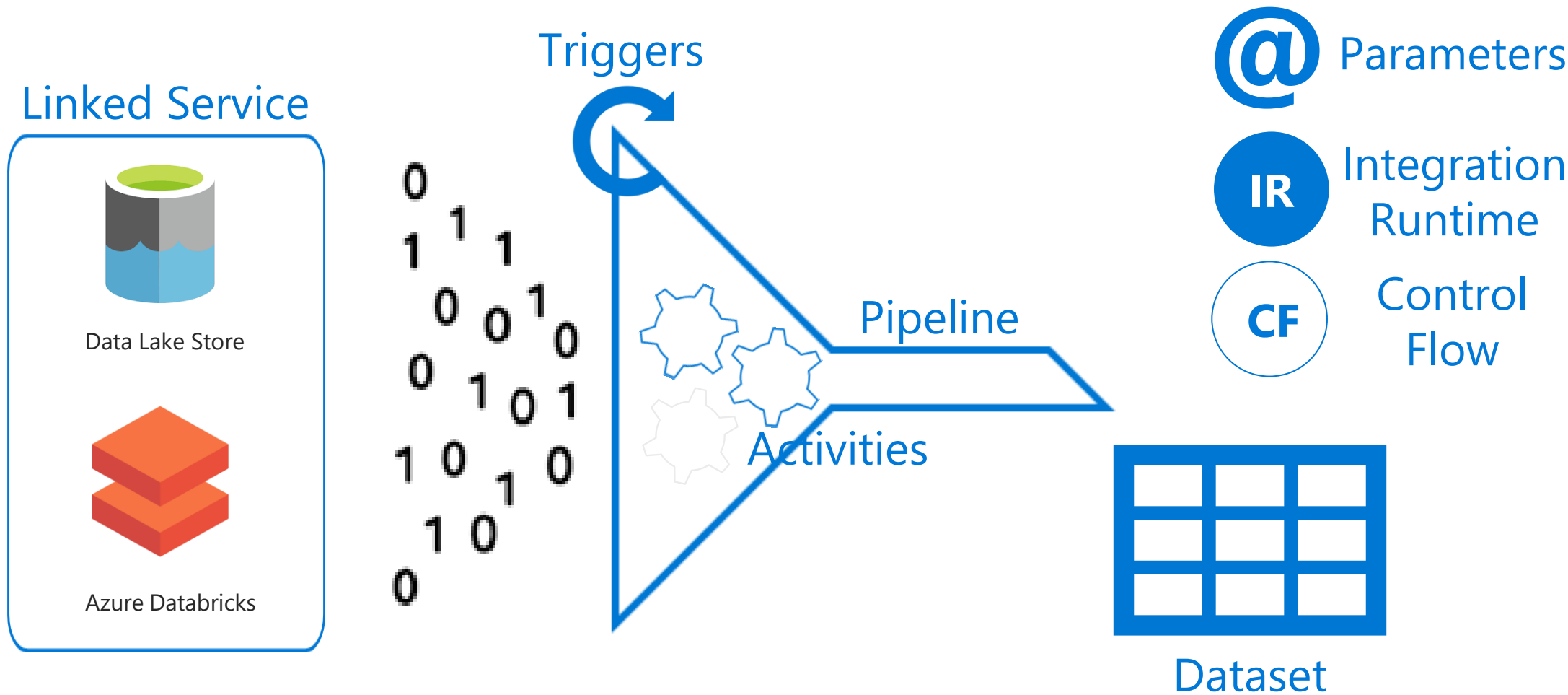
Copy data activity



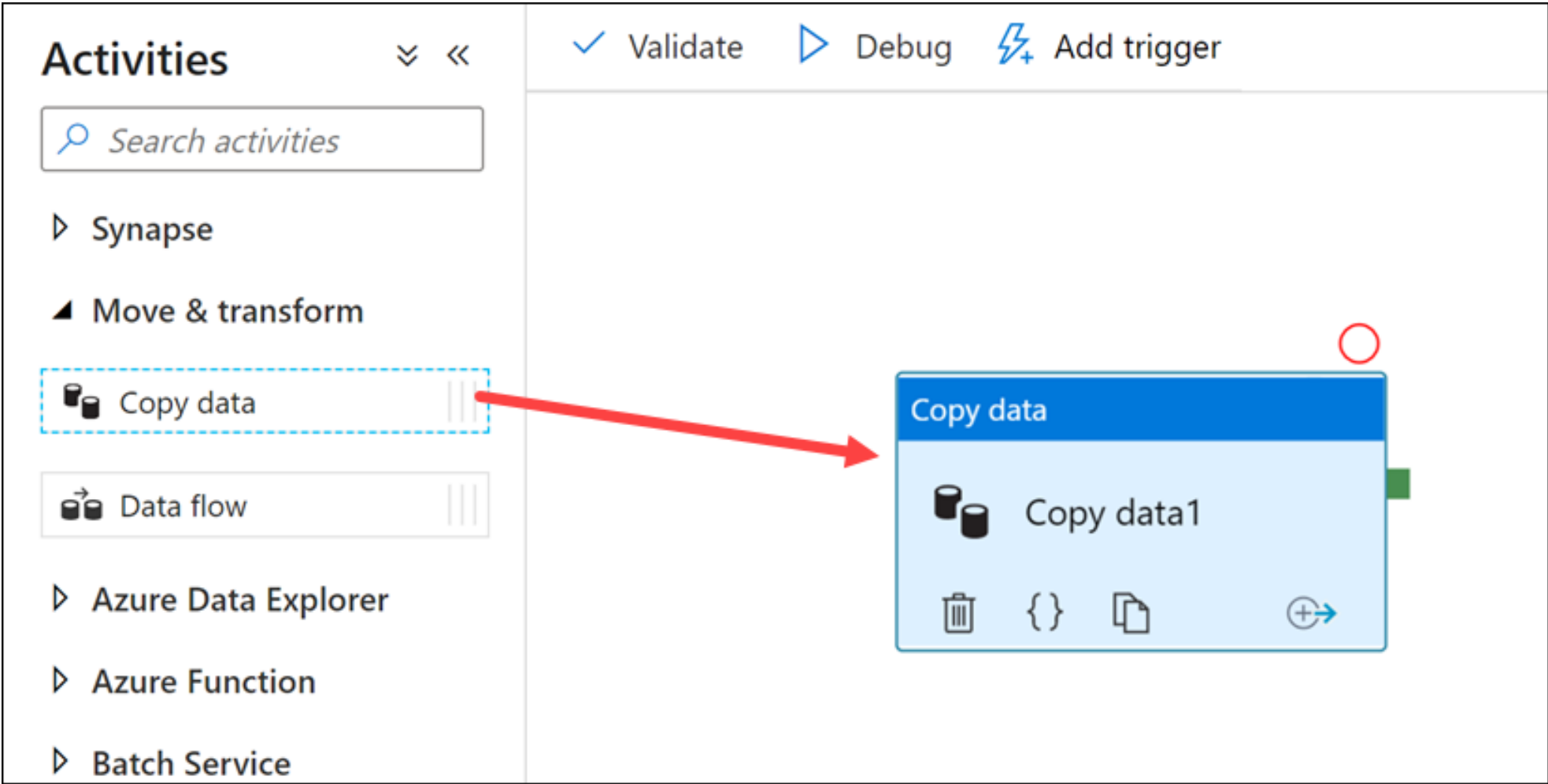
Lesson 02: Petabyte-scale ingestion with Azure Data Factory



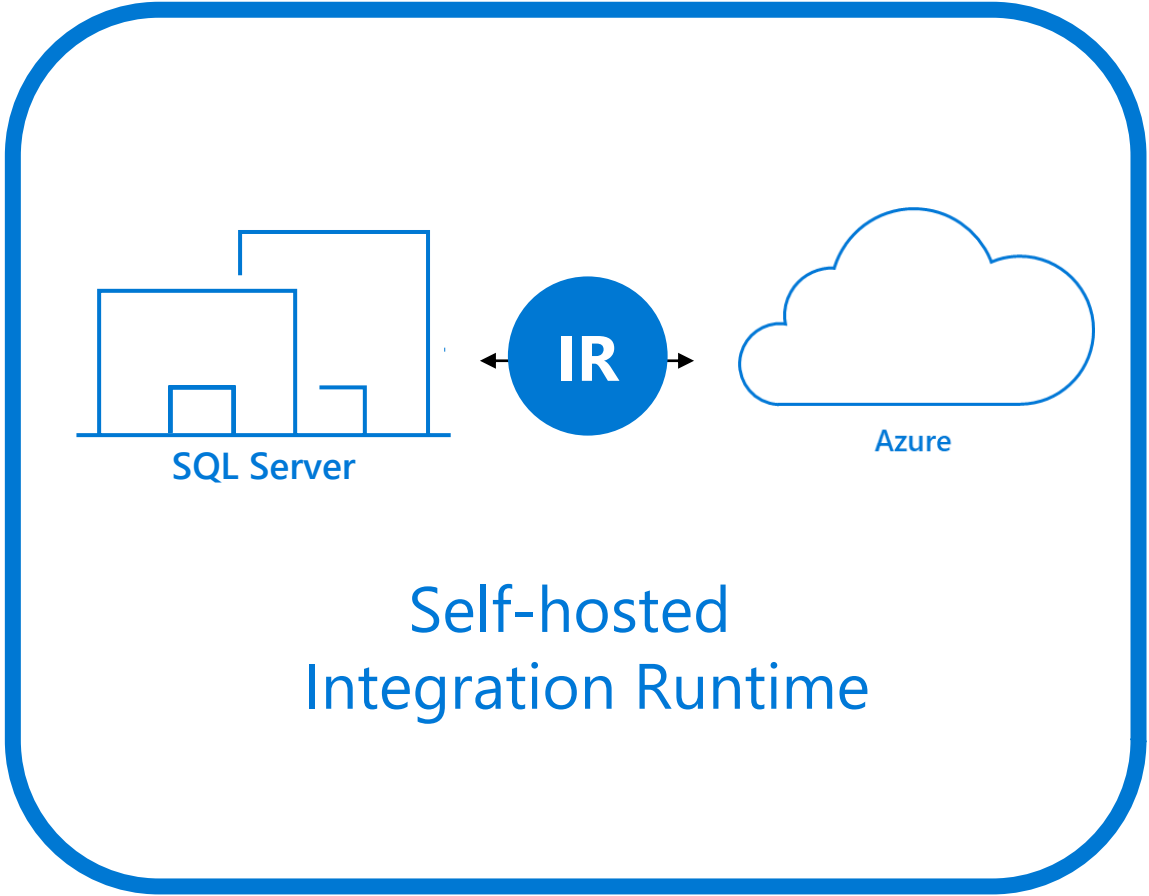
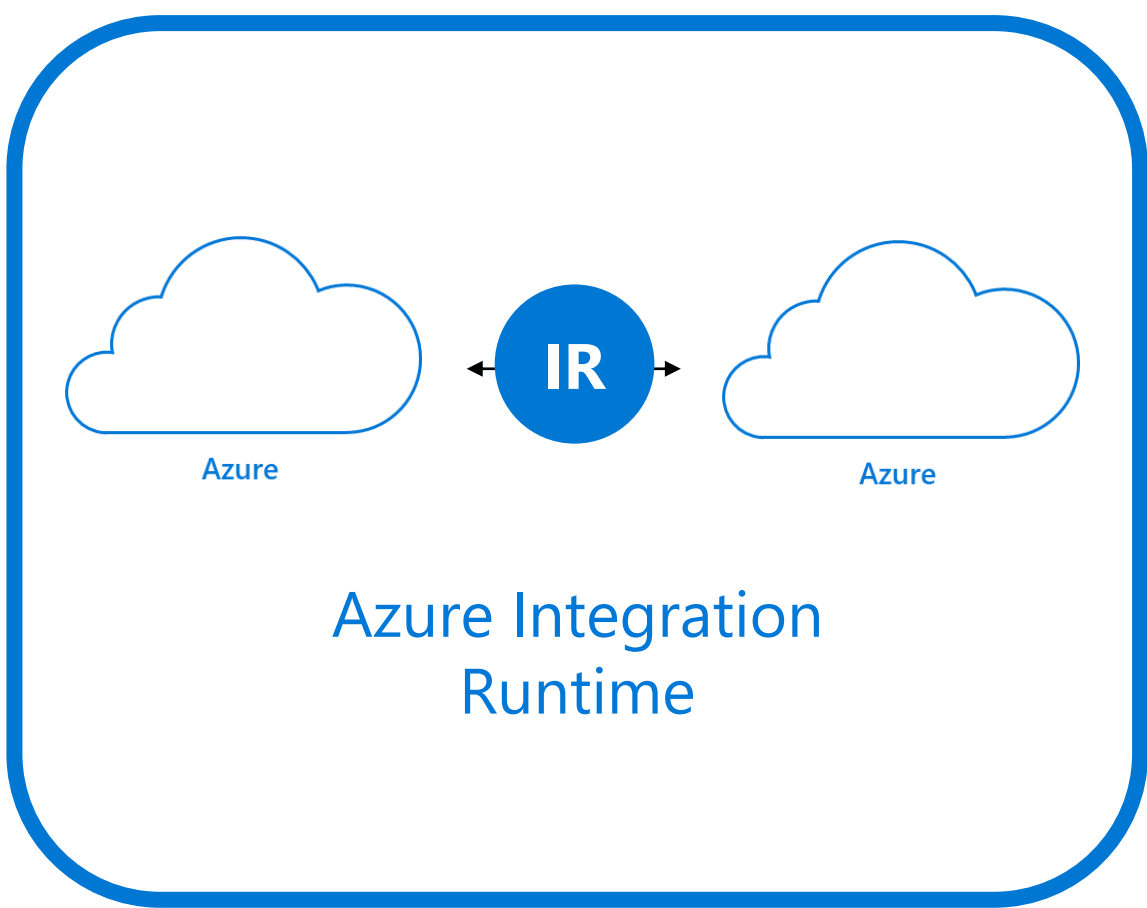
Azure Data Factory/Synapse pipeline revision



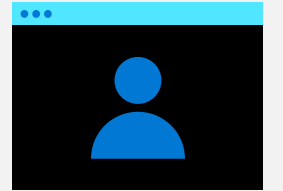
Petabyte-scale ingestion with Azure Data Factory



Understanding integration



Lab: Ingest and load data into the data warehouse



Lab overview

This lab teaches students how to ingest data into the data warehouse through T-SQL scripts and Synapse Analytics integration pipelines. The student will learn how to load data into Synapse dedicated SQL pools with PolyBase and COPY using T-SQL. The student will also learn how to use workload management along with a Copy activity in a Azure Synapse pipeline for petabyte-scale data ingestion.

Lab objectives

After completing this lab, you will be able to:

Use data loading best practices in Azure Synapse Analytics

Petabyte-scale ingestion with Azure Data Factory