Backup Methods

by Sophia Tutorial

☰ WHAT'S COVERED

This tutorial explores various database backup strategies in two parts:

1. Introduction
2. Three Options

1. Introduction

There are many different approaches for database backups. In part, the best approach is dependent on the organization, the size of the database, and how often the database is used. The most common options within most databases are a full database backup, a differential backup, and an incremental backup.

2. Three Options

The simplest type of backup is a full database backup. This provides a copy of the entire database and allows us to restore the data to the point in time when the database backup was made. Note that you can have database transactions running even when the database is still in the process of backing up the data. However, database backup input and output operations can slow down some of those transactions.

Running a full database backup is an option that a lot of organizations use on a nightly basis. It is a good plan if the database size is relatively small, because each backup file is small and can easily be restored independently if there are any issues. However, if the database is large, it can take a lot of time and space to create full backups on a nightly basis. Besides being the simplest, this type of backup is also the fastest to restore, as it only takes one step.

Differential backups are another approach. They contain only the data that has changed since the last full backup. Differential backups are cumulative, not incremental. This means that the differential backup will save all of the changes since the last full backup, even if there are other differential backups run since the last full backup.

For example, imagine that we ran the full backup only Sunday, and ran the differential backup every day. The differential backup on Monday would only contain Monday's data. The differential backup on Wednesday would contain Monday, Tuesday, and Wednesday's data, even though the differential backup ran on Monday and Tuesday. Friday's differential backup would contain all of the data from Monday to Friday. As such, if we used a full backup and differential backup strategy, we would need at most 2 restores to restore the data: the

full backup, and the most recent differential backup. Since the differential backups only contain the data changed since the last full backup, they can be created a lot faster than a full backup.

The incremental backup is similar to a differential backup, but the key difference is that each incremental backup only contains the data from the last backup, regardless of whether it is a full backup, differential backup, or incremental backup. It is not cumulative.

For example, imagine that we ran the full backup on Sunday and ran incremental backups every day. The incremental backup on Monday would have Monday's data. The incremental backup on Tuesday would only have Tuesday's data. Saturday's incremental backup would only have Saturday's data. These incremental backups run in the least amount of time and take the least amount of space. However, they can take the longest to restore as each incremental backup has to be applied one after another. Moreover, if the data from one incremental backup is missing, the incremental backup files after that are not useful because we would be missing the changes.

☑ SUMMARY

There are three primary database backup strategies that can be implemented depending on the organization and the use of the database.

Source: Authored by Vincent Tran