Uncontrolled access to data, with no audit trail of activity and no oversight would be going too far.

—US Naval Officer John Poindexter

A couple of key advantages provided by enterprise systems, which are often overlooked until the auditors start poking around, are process logging, auditing, and data lineage. The concepts are simple: when you move data from system to system and manipulate it along the way, you might need to keep track of where it came from, maintain summary information about data processing, and know which processes touched that data along the way. There are a few reasons for keeping track of this additional information:

- You may need to keep it for troubleshooting purposes, to trace data issues backward through your processes to their source.
- The legal department may have mandated a new legal requirement to document every step of processing in real-time in the event of issues.
- Your business may fall under a regulation such as Sarbanes-Oxley Act, SOX, or Health Insurance Portability and Accountability Act, HIPAA, that requires high levels of security and controls around confidential medical or consumer data.

SSIS provides extensive standard logging functionality that captures runtime information about the status of ETL processing. In addition, it is relatively simple to add auditing capabilities to your enterprise ETL applications with SSIS. In this chapter, you’ll look at ways to take advantage of standard logging and ways to add auditing to your SSIS packages.

Logging

SSIS provides built-in logging capabilities that are easy to enable. In previous chapters, we’ve enabled this feature in some code samples to demonstrate SSIS’s logging capabilities. In this section, we’ll review how to enable and configure logging in your SSIS packages. We’ll also talk about SSIS logging best practices.
Enabling Logging

Logging options in SSIS are accessed at the package level by right-clicking on the control flow and choosing the Logging option from the pop-up context menu, as shown in Figure 13-1.

Figure 13-1. Selecting the Logging option from the pop-up menu

The logging configuration screen allows you to configure SSIS logging options, as shown in Figure 13-2. In this screen, you can set the following logging options:

- The SSIS log provider to use. The most common options are the SSIS log provider for SQL Server, which logs to a SQL Server table, and the SSIS log provider for text files, which outputs logging information to text files. You can, however, choose to log information to the Windows event log, XML files, or SQL Server profiler trace files. The latter options are generally used for specialized troubleshooting tasks.

- The containers for which events should be logged. You can configure logging of specific events at the package, container, and Data Flow task levels. In our example, we’re going to configure at both the package and Data Flow task levels.

- The details you would like logged at the event level. You choose these details from the Details tab.

NOTE: As a general rule, we tend to prefer the SSIS log provider for SQL Server over other logging methods. Because this log provider logs to a SQL Server database table, the logged information is easy to query, simplifying troubleshooting exercises. When you designate the SSIS log provider for SQL Server, it will automatically generate the dbo.sysssislog table and dbo.sp_ssis_addlogentry procedure in your database (if they don’t exist), both of which it marks as system objects.