



Professional

Microsoft®

SQL Server® 2008 Integration Services

Brian Knight, Erik Veerman, Grant Dickinson, Douglas Hinson, Darren Herbold



Updates, source code, and Wrox technical support at www.wrox.com

Contents

Introduction

Chapter 1: Welcome to SQL Server Integration Services

SQL Server SSIS Historical Overview

What's New in SSIS

Getting Started

Architecture

Precedence Constraints

Containers

Variables

Data Flow Elements

Error Handling and Logging

Editions of SQL Server

Summary

Chapter 2: The SSIS Tools

Import and Export Wizard

Business Intelligence Development Studio

Creating Your First Package

The Solution Explorer Window

The SSIS Package Designer

Package Installation Wizard

Management Studio

Summary

Chapter 3: SSIS Tasks

SSIS Task Objects

Looping and Sequence Tasks

Scripting Tasks

Analysis Services Tasks

Data Flow Task

Data Preparation Tasks

RDBMS Server Tasks

Workflow Tasks

SMO Administration Tasks

Summary

Chapter 4: Containers

Task Host Containers

Sequence Containers

Groups

For Loop Container

Foreach Loop Container

Summary

Chapter 5: The Data Flow

The Data Flow

Data Viewers

Sources

Destinations

Transformations

Data Flow Example

Summary

Chapter 6: Using Expressions and Variables

The Paradigm

Understanding Data Types

Using Variables

Working with Expressions

Summary

Chapter 7: Joining Data

The Lookup Component

The Merge Join Component

Contrasting to the Relational Join

New Lookup Features

Building the Basic Package

Using the Lookup Component

Cache Connection Manager and Transform

Summary

Chapter 8: Creating an End-to-End Package

Basic Transformation Tutorial

Typical Mainframe ETL with Data Scrubbing

Looping and the Dynamic Task

Summary

Chapter 9: Scripting in SSIS

Scripting?

Getting Started in SSIS Scripting

[Using the Script Task](#)
[Using the Script Component](#)
[Essential Coding, Debugging, and](#)
[Troubleshooting Techniques](#)
[Summary](#)

[Chapter 10: Loading a Data Warehouse](#)

[Data Profiling](#)
[Data Extraction](#)
[Dimension Table Loading](#)
[Fact Table Loading](#)
[SSAS Processing](#)
[Master ETL Package](#)
[Summary](#)

[Chapter 11: Using the Relational Engine](#)

[Data Extraction](#)
[SQL Server 2008 Change Data Capture](#)
[Data Loading](#)
[Summary](#)

[Chapter 12: Accessing Heterogeneous Data](#)

[Excel and Access](#)
[Oracle](#)
[XML and Web Services](#)
[Flat Files](#)

ODBC

Other Heterogeneous Sources

Summary

Chapter 13: Reliability and Scalability

Restarting Packages

Package Transactions

Error Outputs

Scaling Out

Summary

Chapter 14: Understanding and Tuning the Data Flow Engine

The SSIS Engine

SSIS Data Flow Design and Tuning

Pipeline Performance Monitoring

Summary

Chapter 15: Source Control and Software Development Life Cycle

Introduction to Software Development Life Cycles

Versioning and Source Code Control

Code Deployment and Promotion from Development to Test to Production

Summary

Chapter 16: DTS 2000 Migration

[Managing DTS 2000 Packages within SQL Server Management Studio](#)

[Running DTS 2000 Packages under SSIS](#)

[Migrating DTS 2000 Packages to SSIS](#)

[Using the Package Migration Wizard](#)

[Third-Party Migration Solution](#)

[Summary](#)

[Chapter 17: Error and Event Handling](#)

[Precedence Constraint](#)

[Event Handling](#)

[Breakpoints](#)

[Error Rows](#)

[Logging](#)

[Summary](#)

[Chapter 18: Programming and Extending SSIS](#)

[The Sample Components](#)

[The Pipeline Component Methods](#)

[Building the Components](#)

[Using the Components](#)

[Upgrading to SQL 2008](#)

[Summary](#)

[Chapter 19: Adding a User Interface to Your Component](#)

[Three Key Steps](#)

[Building the User Interface](#)

[Further Development](#)
[Other Considerations](#)
[Summary](#)

[Chapter 20: External Management and WMI Task Implementation](#)

[External Management of SSIS with Managed Code](#)
[Application Object Maintenance Operations](#)
[Package Log Providers](#)
[Package Configurations](#)
[Windows Management Instrumentation Tasks](#)
[Summary](#)

[Chapter 21: Using SSIS with External Applications](#)

[InfoPath Documents](#)
[ASP.NET Applications](#)
[Winform .NET Applications](#)
[Summary](#)

[Chapter 22: Administering SSIS](#)

[Package Configuration](#)
[Deployment Utility](#)
[The Package Store](#)
[Management Studio](#)
[Running Packages with DTExecUI](#)
[Security](#)

[Command-Line Utilities](#)

[Scheduling a Package](#)

[Proxy Accounts](#)

[64-Bit Issues](#)

[Performance Counters](#)

[Summary](#)

[Chapter 23: Case Study: A Programmatic Example](#)

[What You Will Take Away](#)

[Background](#)

[Business Problem](#)

[Solution Summary](#)

[Solution Architecture](#)

[Data Architecture](#)

[Case Study Load Packages](#)

[Case Study Invoice Matching Process](#)

[Creating a Parent Driver Package](#)

[Summary](#)

[Index](#)

[Advertisement](#)

**Professional
SQL Server® 2008
Integration Services**

Brian Knight
Erik Veerman
Grant Dickinson
Douglas Hinson
Darren Herbold



Wiley Publishing, Inc.

Professional SQL Server® 2008 Integration Services

Published by

Wiley Publishing, Inc.

10475 Crosspoint Boulevard

Indianapolis, IN 46256

www.wiley.com

Copyright © 2008 by Wiley Publishing, Inc., Indianapolis,
Indiana

Published simultaneously in Canada

ISBN: 978-0-470-24795-2

Library of Congress Cataloging-in-Publication Data

Professional Microsoft SQL server 2008 integration services /
Brian Knight . . . [et al.].

p. cm.

Includes index.

ISBN 978-0-470-24795-2 (paper/website)

1. SQL server. 2. Database management. I. Knight, Brian.

QA76.9.D3P7662 2008

005.75'85—dc22

2008025018

No part of this publication may be reproduced, stored in a retrieval system or transmitted in any form or by any means, electronic, mechanical, photocopying, recording, scanning or otherwise, except as permitted under Sections 107 or 108 of the 1976 United States Copyright Act, without either the prior written permission of the Publisher, or authorization through payment of the appropriate per-copy fee to the Copyright Clearance Center, 222 Rosewood Drive, Danvers, MA 01923, (978) 750-8400, fax (978) 646-8600.

Requests to the Publisher for permission should be addressed to the Legal Department, Wiley Publishing, Inc., 10475 Crosspoint Blvd., Indianapolis, IN 46256, (317) 572-3447, fax (317) 572-4355, or online at <http://www.wiley.com/go/permissions>.

Limit of Liability/Disclaimer of Warranty: The publisher and the author make no representations or warranties with respect to the accuracy or completeness of the contents of this work and specifically disclaim all warranties, including without limitation warranties of fitness for a particular purpose. No warranty may be created or extended by sales or promotional materials. The advice and strategies contained herein may not be suitable for every situation. This work is sold with the understanding that the publisher is not engaged in rendering legal, accounting, or other professional services. If professional assistance is required, the services of a competent professional person should be sought. Neither the publisher nor the author shall be liable for damages arising herefrom. The fact that an organization or Website is referred to in this work as a citation and/or a potential source of further information does not mean that the author or the publisher endorses the information the organization or Website may provide or recommendations it may make. Further, readers should be aware that Internet Websites listed in this work may have changed or disappeared between when this work was written and when it is read.

For general information on our other products and services or to obtain technical support, please contact our Customer Care Department within the U.S. at (800) 762-2974, outside the U.S. at (317) 572-3993 or fax (317) 572-4002.

Trademarks: Wiley and the Wiley logo, Wrox, the Wrox logo, Wrox Programmer to Programmer, and related trade dress are trademarks or registered trademarks of John Wiley & Sons, Inc. and/or its affiliates, in the United States and

other countries, and may not be used without written permission. Microsoft and SQL Server are registered trademarks of Microsoft Corporation in the United States and/or other countries. All other trademarks are the property of their respective owners. Wiley Publishing, Inc. is not associated with any product or vendor mentioned in this book.

Wiley also publishes its books in a variety of electronic formats. Some content that appears in print may not be available in electronic books.

About the Authors

Brian Knight, SQL Server MVP, MCSE, MCDBA, is the co-founder of SQLServerCentral.com and JumpstartTV.com. Brian is a Principal Consultant and owner of Pragmatic Works. He runs the local SQL Server users' group in Jacksonville (JSSUG) and was on the Board of Directors of the Professional Association for SQL Server (PASS). Brian is a contributing columnist for SQL Server Standard and also maintains a regular column for the database website SQLServerCentral.com and does regular webcasts at Jumpstart TV. He has authored nine SQL Server books during the past 10 years. Brian has spoken at conferences like PASS, SQL Connections, and TechEd, and many Code Camps. You can find his blog at <http://www.pragmaticworks.com>. Brian spends weekends practicing to be a professional cage fighter and practicing for next season's *American Idol*.

Erik Veerman is a Mentor for Solid Quality Mentors focusing on training, mentoring, and architecting solutions on the SQL Server BI platform. His industry recognition includes Microsoft's Worldwide BI Solution of the Year and *SQL Server Magazine's* Innovator Cup winner. Erik has designed dozens of BI solutions across a broad business spectrum — telecommunications, marketing, retail, commercial real estate, finance, supply chain, and information technology. His experience with high-volume multi-terabyte environments and SQL Server 64-bit has enabled clients to scale their Microsoft-based BI solutions for optimal potential. As an expert in OLAP design, ETL processing, and dimensional modeling, Erik is a presenter, author, and instructor. He led the ETL architecture and design for the first production implementation of Integration Services (SSIS) and helped drive the ETL standards and best

practices for SSIS on Microsoft's SQL Server 2005 reference initiative, Project REAL. Erik is also co-author of *Professional SQL Server 2005 Integration Services* and *Expert SQL Server 2005 Integration Services*, and lead author for the MS Press Training Kit *SQL Server 2005 Business Intelligence Implementation and Maintenance*. As a resident of Atlanta, GA, Erik participates in the local Atlanta SQL Server User's Group, a PASS chapter.

Grant Dickinson is a Program Manager at Microsoft, focusing on designing technologies that enable customers and partners to create innovative and scalable Business Intelligence solutions. Grant has helped qualify, architect, and implement BI solutions across a broad range of industries, including a solution that was once one of the largest Microsoft-based data warehouses in the world. He has designed and provided expertise into product features across the Microsoft BI stack, including technologies in SSIS, SQL Server, and Office. Grant helped develop the Microsoft best-practices ETL reference implementation, Project REAL, and has spoken at conferences around the world. Grant is currently focused on data quality and stewardship in the Master Data Management space. Grant was born in Zimbabwe and spent much of his youth in Southern Africa. He gained a BSc Computer Science at the University of the Witwatersrand in Johannesburg, and today he lives in Seattle with his wife and family.

Douglas Hinson splits his time between database and software development for financial applications in the logistics and insurance industries. Douglas specializes in conceptualizing, reengineering, and developing back-end solutions that connect business operational and financial functions. As a result, he has an extensive background in SQL Server and financial applications, and fits in some technical writing on the side. He has coauthored several Wrox books, including *SQL Server 2005 Performance Tuning*,

SQL Server 2005 CLR Programming, and the previous edition of this book, *SQL Server 2005 Integration Services*.

Darren Herbold, MCDBA, MCSE is a dedicated consultant who is passionate about delivering business value to his clients. A principal consultant at Pragmatic Works Consulting (www.PragmaticWorks.com) and a graduate from Florida State University, he has expertise in Business Intelligence, Database Administration, and .NET Software Development. His main focus is on the SQL Server stack, where he delivers training, mentoring, and develops Data Warehouse, ETL, and Reporting solutions for his clients. He has developed a robust .NET application development framework and a code-generation tool that saves clients an average of 30 to 40 percent off of development time. Darren has also created BI and software solutions for clients such as Microsoft, Post Properties, and the University of South Florida. Visit his blog at: <http://pragmaticworks.com/community/blogs/>.

Credits

Executive Editor

Bob Elliott

Development Editor

Brian MacDonald

Technical Editors

Douglas Laudenschlager

Carla Sabotta

Michael A. Entin

Ranjeeta Nanda

Ritu Kothari

Feng Guo

Neal Graves

Devin Knight

Production Editor

Kathleen Wisor

Copy Editor

Kim Cofer

Editorial Manager

Mary Beth Wakefield

Production Manager

Tim Tate

Vice President and Executive Group Publisher

Richard Swadley

Vice President and Executive Publisher

Joseph B. Wikert

Project Coordinator, Cover

Lynsey Stanford

Proofreader

Nancy Carrasco

Indexer

Melanie Belkin

Acknowledgments

As always, I must thank my wife and best friend for supporting me for the past 10 years of marriage. I've been fortunate to have found a woman who doesn't fall asleep immediately when copyediting my technical writing. Thanks to my three children: Colton, Liam, and Camille for allowing their daddy to be distracted sometimes with this book when they wanted to play. Thanks also to all the wonderful co-authors, who truly made this book possible. Once again, I must thank the Pepsi Cola Company for inventing Mountain Dew, which drove the late night writing. Lastly, thanks to my sensei and song writer Sensei Yoshi, who has helped me win my first cage fighting match and is preparing me for next year's *American Idol*.

—**Brian Knight**

First of all, I'd like to thank my kids, Meg, Nate, Kate, and Caleb, for being patient with me through the sometimes tiresome process of writing. And of course, my wife, Amy, is amazing! Also thanks to Brian Knight and the other authors who helped make this book so valuable. And thanks to Andy Leonard for his help. There's no one better at knowing how SSIS integrates with Visual Studio Team System. Thanks go to the producers of coffee beans, because without caffeine, I honestly don't know if my contribution to this book would have been worth anything!

—**Erik Veerman**

To my beautiful Heidi, thank you for your support, encouragement, and understanding, and for being the best wife and friend I could ask for. Jessica and Anna, my two wonderful little girls, I thank God every day that we have the gift of you in our lives. I cherish the times when you two sat on my lap while I (ungainly) authored this book, waiting

for me to finish so we could go and play outside. You are so little but you give so much. Bruce, you are a brother and uncle-extraordinaire. Dad, Mom, Mae, Pai (and the whole fandamily); though we are spread around the world, your legacy keeps our faith, values, and family strong. Hatch, Dives, Cyril, Neil, Len, Dave, Donald, Erik, and Henk, thanks for your friendship, mentorship, and passion during these years in Microsoft. Finally, my gratitude to Brian; it's been fun presenting and writing with you — good job on another great book!

—***Grant Dickinson***

Thanks to God for the blessing of being able to do what I love for a living. To my beautiful wife, Misty, thank you for being so supportive and understanding during this project, as always. Kyle and Mariah, thanks for being so patient with your Dad while he was putting this project together. A big thanks to the Wrox and Microsoft Tech editors and our Jacksonville area SQL Server guru, Brian Knight, who has come through again with a great cast of authors and a reworked, well-crafted guide to SQL Server Integrated Services 2008.

—***Douglas Hinson***

I'd first like to thank my beautiful wife, Ashley, and my wonderful children, Sydney and Kiley, for all their patience and support during this process. I want to also thank Brian Knight for giving me the awesome opportunity to be involved in this project. He's been a fantastic mentor, friend, and overall great guy. Mt. Dew played a pivotal role in this too, for obvious reasons. I also would like to thank all the great folks at KBX Boxing Gym in Alpharetta for teaching me to dig deep when I would much rather vomit and pass out. Go Krav Maga! Lastly, I want to thank the fine staff at Wiley Publishing for their support and guidance in this endeavor.

—***Darren Herbold***

Introduction

SQL Server Integration Services (SSIS) was released to the market in SQL Server 2005 and took the Extract Transform Load (ETL) market by surprise. In SQL Server 2008, SSIS has focused on maturing the product and improving the product's scalability and performance by an astonishing 70% in some cases. If you're new to SSIS, you've picked a fantastic field to become involved in! The one consistent skill needed in today's technical job market is ETL. If a company wants to establish a partnership with another company, they'll need to communicate data back and forth between the two companies. If your company wants to launch new products, they'll need a way to integrate those products into their website and catalog. All of these types of tasks are going to require the skillset you are developing and will learn in this book.

Companies that had never used SQL Server before are now allowing it in their environment because SSIS is such an easy-to-use and cost-effective way to move data. SSIS competes with the largest ETL tools on the market, like Data Stage and Ab Initio, at a tiny fraction of the price. SQL Server 2008 now offers more components that you use to make your life even easier and the performance scales to a level never seen on the SQL Server platform.

The best thing about SSIS is its price tag: free with your SQL Server purchase. Many ETL vendors charge hundreds of thousands of dollars for what you will see in this book. SSIS is also a great platform for you to expand and integrate into, which many ETL vendors do not offer. Once you get past the initial learning curve, you'll be amazed with the power of the tool, and it can take weeks off your time to market.

Who This Book Is For

Having used SSIS since the beta stages of SQL Server 2005 and through its evolution into its current form, the idea of writing this book was quite compelling. If you've never used SSIS before, we spend the first chapters focusing on lowering your learning curve on this product. If you've used SSIS in the past, we've added quite a bit of new content that is specific to SQL Server 2008 and to take your skills to the next level. If you're an SSIS 2005 user, luckily, this is an incremental release, and you won't have to completely relearn your skills.

This book is intended for developers, DBAs, and casual users who hope to use SSIS for transforming data, creating a workflow, or maintaining their SQL Server. This book is a *professional* book, meaning that the authors assume that you know the basics of how to query a SQL Server and have some rudimentary programming skills. Not much programming skill will be needed or assumed, but it will help with your advancement. No skills in the prior release of SSIS (called DTS then) are required, but we do reference it throughout the book when we call attention to feature enhancements.

What This Book Covers

Whether you're new to SSIS or an experienced SSIS developer, there's something for you in this book. This book takes you from the architecture and basics of SSIS all the way through to developing hard-core SSIS solutions to solve many of the industry's common business scenarios. The book is tutorial based, meaning that it teaches you through simple examples.

By the time you've completed this book, you'll know how to load and synchronize database systems using SSIS by

using some of the new SQL Server 2008 features. You'll also know how to load data warehouses, which is a very hot and specialized skill. Even in warehousing, you'll find features in the new 2008 release that you'll wonder how you lived without!

How This Book Is Structured

After discussing the architecture of SSIS, we'll start with the basics by introducing the fundamental concepts of SSIS: The Data Flow and Control Flow. We'll then build through the various other features, including the warehousing and scripting, and proceed to advanced topics like programming and extending the engine. We'll conclude with a case study that helps to tie everything together. SSIS is a very feature-rich product, and it took a lot to cover the product.

Chapter 1, "Welcome to SQL Server Integration Services," introduces the concepts that we're going to discuss throughout the remainder of this book. We talk about the SSIS architecture and give a brief overview of what you can do with SSIS.

Chapter 2, "The SSIS Tools," shows you how to quickly learn how to import and export data by using the Import and Export Wizard and then takes you on a tour of the Business Intelligence Development Studio (BIDS).

Chapter 3, "SSIS Tasks," goes into each of the tasks that are available to you in SSIS. These tasks are the building blocks for your SSIS workflow and are much like Lego block programming.

Chapter 4, "Containers," covers how to use containers to do looping in SSIS and describes how to configure each of the basic transforms.

Chapter 5, "The Data Flow," dives into the data flow components in SSIS. These components are where typical

ETL developers will spend 75% of their time when loading a database.

Chapter 6, “Using Expressions and Variables,” instructs you how to use the obscure expression language in SSIS by showing you many example use cases and how to solve them through the language.

Chapter 7, “Joining Data,” focuses on how to join systems together, whether those systems are two flat files or database platforms. Much of the chapter is spent showing the Lookup Component, which is where much of the work into SSIS 2008 went.

Now that you know how to configure most of the tasks and transforms, Chapter 8, “Creating an End-to-End Package,” puts it all together with a large example that lets you try out your SSIS experience.

Chapter 9, “Scripting in SSIS,” shows you some of the ways you can use the Script Task in SSIS.

Chapter 10, “Loading a Data Warehouse,” covers how to load a data warehouse from the ground up through example. Even smaller companies now are finding that to compete they need to make their data work for them by employing a data warehouse. We show how to load dimension and fact tables in this chapter and some of the common issues.

Chapter 11, “Using the Relational Engine” focuses on how to synchronize systems incrementally. Generally, it’s too inefficient to completely purge and load a system daily or monthly. This chapter shows you some of the new SQL Server 2008 features like Change Data Capture that help you make this synchronization a smooth process.

Sometimes you connect to systems other than SQL Server. Chapter 12, “Accessing Heterogeneous Data,” shows you how to connect to systems other than SQL Server like Excel, XML, and Web services.

Chapter 13, “Reliability and Scalability” demonstrates how to scale SSIS and make it more reliable. You can use the features in this chapter to show you how to make the package restartable if a problem occurs.

Chapter 14, “Understanding and Tuning the Data Flow Engine,” explains the architecture of the SSIS data flow engine in detail and how to tune your SSIS packages for maximum efficiency.

Chapter 15, “Source Control and Software Development Life Cycle,” introduces a software development life cycle methodology to you. It speaks to how SSIS can integrate with Visual Studio Team System.

Chapter 16, “DTS 2000 Migration” shows how to migrate DTS 2000 packages to SSIS and if necessary, how to run DTS 2000 packages under SSIS. It also discusses third-party management to convert packages.

Chapter 17, “Error and Event Handling,” discusses how to handle problems with SSIS with error and event handling.

Chapter 18, “Programming and Extending SSIS,” shows the SSIS object model and how to use it to extend SSIS. The chapter goes through creating your own task, and then Chapter 19, “Adding a User Interface to Your Component,” adds a user interface to the discussion.

Chapter 20, “External Management and WMI Task Implementation,” walks through creating an application that interfaces with the SSIS to manage the environment. It also discusses the WMI set of tasks.

Chapter 21, “Using SSIS with External Applications,” teaches you how to expose the SSIS Data Flow to other programs like InfoPath and your own .NET applications.

Chapter 22, “Administering SSIS,” shows you how to deploy and administer the packages that you’ve worked so hard to develop. We cover the SSIS service, how to run packages and schedule packages and some of the challenges you’ll see with 64-bit systems.

Chapter 23 is a programmatic case study that creates three SSIS packages for a banking application.

What You Need to Use This Book

To follow this book, you will only need to have SQL Server 2008 and the Integration Services component installed. You'll need a machine that can support the minimum hardware requirements to run SQL Server 2008. You'll also want to have the AdventureWorks2008 and AdventureWorksDW2008 databases installed. Instructions for accessing these databases can be found in the ReadMe file on this book's Web site.

Conventions

To help you get the most from the text and keep track of what's happening, we've used a number of conventions throughout the book:

- We *highlight* new terms and important words when we introduce them.
- We show keyboard strokes like this: Ctrl+A.
- We show filenames, URLs, and code within the text like SO: persistence.properties.
- We present code in two different ways:

In code examples we highlight new and important code with a gray background.

The gray highlighting is not used for code that's less important in the present context or that has been shown before.

Source Code

As you work through the examples in this book, you may choose either to type in all the code manually or to use the source code files that accompany the book. All of the source code used in this book is available for download at <http://www.wrox.com>. Once at the site, simply locate the book's title (either by using the Search box or by using one of the title lists) and click the Download Code link on the book's detail page to obtain all the source code for the book.

Because many books have similar titles, you may find it easiest to search by ISBN; this book's ISBN is 978-0-470-24795-2.

Once you download the code, just decompress it with your favorite compression tool. Alternatively, you can go to the main Wrox code download page at www.wrox.com/dynamic/books/download.aspx to see the code available for this book and all other Wrox books.

Errata

We make every effort to ensure that there are no errors in the text or in the code. However, no one is perfect, and mistakes do occur. If you find an error in one of our books, like a spelling mistake or faulty piece of code, we would be very grateful for your feedback. By sending in errata, you may save another reader hours of frustration, and at the same time you will be helping us provide even higher-quality information.

To find the errata page for this book, go to <http://www.wrox.com> and locate the title using the Search box or one of the title lists. Then, on the book details page, click the Book Errata link. On this page you can view all errata that has been submitted for this book and posted by Wrox editors. A complete book list including links to each

book's errata is also available at www.wrox.com/misc-pages/booklist.shtml.

If you don't spot "your" error on the Book Errata page, go to www.wrox.com/contact/techsupport.shtml and complete the form there to send us the error you have found. We'll check the information and, if appropriate, post a message to the book's errata page and fix the problem in subsequent editions of the book.

p2p.wrox.com

For author and peer discussion, join the P2P forums at p2p.wrox.com. The forums are a Web-based system for you to post messages relating to Wrox books and related technologies and to interact with other readers and technology users. The forums offer a subscription feature to e-mail you topics of interest of your choosing when new posts are made to the forums. Wrox authors, editors, other industry experts, and your fellow readers are present on these forums.

At <http://p2p.wrox.com> you will find a number of different forums that will help you, not only as you read this book, but also as you develop your own applications. To join the forums, just follow these steps:

- 1.** Go to p2p.wrox.com and click the Register link.
- 2.** Read the terms of use and click Agree.
- 3.** Complete the required information to join as well as any optional information you wish to provide and click Submit.
- 4.** You will receive an e-mail with information describing how to verify your account and complete the joining process.

You can read messages in the forums without joining P2P, but in order to post your own messages, you must

join.

Once you join, you can post new messages and respond to messages other users post. You can read messages at any time on the Web. If you would like to have new messages from a particular forum e-mailed to you, click the Subscribe to this Forum icon by the forum name in the forum listing.

For more information about how to use the Wrox P2P, be sure to read the P2P FAQs for answers to questions about how the forum software works, as well as many common questions specific to P2P and Wrox books. To read the FAQs, click the FAQ link on any P2P page.

Chapter 1

Welcome to SQL Server Integration Services

SQL Server Integration Services (SSIS) is the anchor in a trilogy of products that make up the Microsoft SQL Server Business Intelligence (BI) platform. SSIS along with Analysis Services and Reporting Services round out a platform that clearly puts Microsoft on the map in the enterprise Business Intelligence arena. In its simplest form, SSIS is an enterprise-level extract, transform, and load (ETL) development tool. However, SSIS is not just a fancy wrapper around an import wizard. In a drag-and-drop development environment, ETL developers can snap together intricate workflows and out-of-the-box data-cleansing flows that rival custom coding and expensive third-party tools. For your edge cases, the model is easily extensible and custom components can be developed in .NET languages to simply snap into the framework. However, custom coding most likely will not even be necessary. With the latest version of SSIS, novice developers can use the embedded Visual Studio Tools for Applications (VSTA) development environment to custom code workflow tasks and data pipeline transformations in VB or C# .NET languages.

When we put together the first edition of this book, we were blown away by the new architecture and capabilities of SSIS. SSIS was a big change from the Data Transformation Services (DTS) product that it replaced and there has been much to learn. Since the first edition of SSIS, we have collectively racked up many years of experience converting