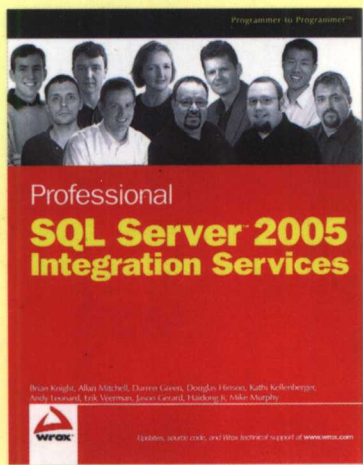


SQL Server 2005深入开发系列

SQL Server 2005

集成服务高级编程



清华大学出版社

 WILEY

Brian Knight 等 著

杨战伟 译

访问www.wrox.com可以获得更新

内容、源代码和Wrox技术支持

SQL Server 2005 深入开发系列

SQL Server 2005 集成服务高级编程

Brian Knight 等著

杨战伟 译

清华大学出版社

北 京

Brian Knight 等
Professional SQL Server 2005 Integration Services
EISBN: 0-7645-8435-9

Copyright © 2006 by Wiley Publishing, Inc.

All Rights Reserved. Authorized translation from the English language edition published by John Wiley & Sons, Inc.

Simplified Chinese translation edition is published and distributed exclusively by Tsinghua University Press under the authorization by John Wiley & Sons, Inc., within the territory of the People's Republic of China only (excluding Hong Kong, Macao SAR and Taiwan). Unauthorized export of this edition is a violation of the Copyright Act. Violation of this Law is subject to Civil and Criminal Penalties.

本书中文简体字翻译版由美国 John Wiley & Sons, Inc. 公司授权清华大学出版社在中华人民共和国境内(不包括中国香港、澳门特别行政区和中国台湾地区)独家出版发行。未经许可之出口视为违反著作权法,将受法律之制裁。未经出版者预先书面许可,不得以任何方式复制或抄袭本书的任何部分。

北京市版权局著作权合同登记号 图字: 01-2006-5907 号

本书封面贴有 John Wiley & Sons 公司防伪标签, 无标签者不得销售。

版权所有, 侵权必究。侵权举报电话: 010-62782989 13501256678 13801310933

图书在版编目(CIP)数据

SQL Server 2005 集成服务高级编程 / (美) 奈特 (Knight, B.) 等著; 杨战伟译. —北京: 清华大学出版社, 2007.1

书名原文: Professional SQL Server 2005 Integration Services

ISBN 978-7-302-14027-6

I. S… II. ①奈… ②杨… III. 关系数据库—数据库管理系统, SQL Server 2006—程序设计
IV. TP311.138

中国版本图书馆 CIP 数据核字 (2006) 第 124208 号

责任编辑: 常晓波

责任校对: 张 剑

责任印制: 王秀菊

出版发行: 清华大学出版社 地 址: 北京清华大学学研大厦 A 座

http://www.tup.com.cn 邮 编: 100084

c-service@tup.tsinghua.edu.cn

社 总 机: 010-62770175 邮购热线: 010-62786544

投稿咨询: 010-62772015 客户服务: 010-62776969

印 刷 者: 清华大学印刷厂

装 订 者: 三河市金元印刷有限公司

经 销: 全国新华书店

开 本: 185×230 印 张: 37.25 字 数: 922 千字

版 次: 2007 年 1 月第 1 版 印 次: 2007 年 1 月第 1 次印刷

印 数: 1~3000

定 价: 69.00 元

本书如存在文字不清、漏印、缺页、倒页、脱页等印装质量问题, 请与清华大学出版社出版部联系
调换。联系电话: (010)62770177 转 3103 产品编号: 022603-01

献给我具有无限耐心的妻子 **Jennifer**

关于作者

Brian Kight 是 SQL Server MVP、MCSE、MCDBA，他是 SQLServerCentral.com 的共同创办人，目前在 Professional Association for SQL Server (PASS) 的董事会任职。他在佛罗里达州的 Jacksonville 开办了一个本地 SQL Server 用户组 (JSSUG)。Brian 是 *SQL Server Standard* 的专栏作家，同时还维护一个针对数据库 Web 站点 SQLServerCentral.com 的专栏。他是 *Admin911:SQL Server* (由 Osborne/McGraw-Hill 出版) 的作者，同时是 *Professional SQL Server DTS* 和 *Professional SQL Server 2005 SSIS* (由 Wiley 出版) 的合著者。Brian 曾在诸如 PASS、SQL Connections 和 TechEd 等大会上做技术演讲。可在以下网站访问他的博客：

www.whiteknighttechnology.com

Allan Mitchell 是一家基于英国 Konesans 的咨询公司的共同所有者，专门研究 ETL 实现和设计。他现在忙于为英国一家顶级投资银行完成一个全国信用卡风险模型项目，同时为客户设计自定义的 SSIS 组件。

Darren Green 是一家基于英国 Konesans 的顾问公司的共同所有者，专门研究 SQL Server，以及 DTS 和 SSIS 解决方案。他管理过从版本 6.5 以来的各种 SQL Server 数据库系统，在 SQL Server 的许多方面拥有丰富经验。他同时还负责一个资源站点 SQLDTS.com 以及 SQLIS.com，同时也是一位 Microsoft MVP。

Douglas Hinson 是任职于 Jacksonville, Florida 的 Hinson&Associates Consulting 公司的一位高级架构师，他对数据库开发和软件开发两方面都很精通。Douglas 专门研究在多种开发环境下概念化并构建针对薪水扣减、收费、付款、索赔处理等操作的保险业后端解决方案。他还开发过后勤和邮政服务应用程序。

Kathi Kellenberger 是一位数据库管理员，任职于 Bryan Cave LLP 公司，该公司是一家总部位于密苏里州圣路易斯市的国际化法律公司。Kathi 多年前在大学里第一次使用 Radio Shack TRS-80 时便和计算机结下了不解之缘。在从事目前这份工作之前她曾经在卫生保健领域工作了 16 年。她目前和她的丈夫 Dennis、上大学的儿子 Andy 和一些宠物居住在伊利诺伊州的 Edwardsville，她的女儿 Denise 则和她比邻而居。工作之余，她会为 SQLServerCentral.com 工作或撰写文章，而其他的空闲时间，她会和她的姐妹们一起徒步旅行、骑自行车或者是在本地的卡拉 OK 厅唱歌。

Andy Leonard 是一位 SQL Server DBA、MSCD 和工程师，他居住在佛罗里达州的 Jacksonville。Andy 管理一个 SQL Server DBA 团队。他善于开发各种类型的企业级解决方案，并热衷于开发工业企业的业务智能解决方案。有关他的更多内容可以访问 www.andyleonard.net，并且可以通过 andy@andyleonard.net 来联系他。

Erik Veerman 是 Solid Quality Learning 公司的顾问，该公司位于佐治亚州的亚特兰大市。自从在 SQL Server 7.0 中第一次发布 DTS 和 OLAP Server 起，Eirk 便开发过基于 Microsoft 产品的业务智能以及基于 ETL 的解决方案，并且与各种客户和各个领域打过交道。他在业界的声誉包括 Microsoft 世界级 BI 解决方案年度大奖，以及 *SQL Server Magazine* 颁发的 Innovator Cup 大奖得主。Erik 领导了 Integration Service 第一个产品实现的 ETL 体系结构和设计工作，并通过 Microsoft 的 SQL Server 2005 reference initiative 即 REAL 项目参与了开发 ETL 标准以及针对 Integration Service 的最佳开发方法。

Jason Gerard 是 Object Future Consulting 公司的董事长，该公司是一家软件开发和咨询公司，位于佛罗里达州的 Jacksonville (www.objectfuture.com)。Jason 是 .NET 和 J2EE 技术方面的专家，并且开发过针对卫生保健、金融和保险领域的企业级应用程序。在工作之余，Jason 则和他的妻子 Sandy、儿子 Jakob 以及他家的懒狗 Tracker 一起共享天伦之乐。

HaidongJi (季海东) 是 MSCD 和 MCDBA，他是一位在伊利诺伊州芝加哥市工作的高级数据库管理员。他负责管理企业 SQL Server 系统以及一些位于 Unix 和 Linux 平台上的 Oracle 以及 MySQL 系统。在从事目前这份工作之前，他是一个专注于 Visual Basic、COM 和 COM+ 以及 SQL Server 等领域的开发人员。他还是 SQLServerCentral.com 网站的专栏作家，这个网站是一个著名的 SQL Server 门户网站。

Mike Murphy 是一位 .NET 开发人员和 MCSD，他以前是一位自动控制系统工程师，目前居住在佛罗里达州的 Jacksonville。Mike 喜欢跟踪计算机技术的最新进展，和 Jacksonville Developer User Group 大会 (www.jaxdug.com) 的同事们交流，并且在时间允许的情况下，驾驶 R/C 直升机。要联系 Mike，可以发电子邮件到 mike@murphysgeekdom.com 或访问 www.murphysgeekdom.com。

致谢名单

责任编辑

Bob Elliott

开发编辑

Brian MacDonald

技术编辑

Slobodan M. Bojanic

James K. Howey

Ted Lee

Runying Mao

Ashwani Nanda

Ashvini Sharma

制作编辑

William A. Barton

文字编辑

出版服务部

编辑经理

Mary Beth Wakefield

制作经理

Tim Tate

副总裁和执委会出版人

Richard Swadley

副总裁和出版人

Joseph B. Wikert

项目协调人

Ryan Steffen

图形和制作专家

Denny Hager

Joyce Haughey

Jennifer Heleine

Barbara Moore

Alicia B. South

质量控制技术员

John Greenough

Brian H. Walls

媒体开发专家

Angela Denny

校对和索引编制

TECHBOOKS 制作服务部

致 谢

首先也是最重要的，我要感谢我的妻子在我编写本书的几个月里细心照料我的两个孩子。在此照例要说的是，如果没有我的妻子 Jennifer，我什么事也干不成。我很遗憾我能献给她的只是一本技术图书。我要感谢我的两个孩子 Colton 和 Liam，他们和父亲——我在一起时是如此耐心。还要感谢 Microsoft 的所有朋友（特别是 Ash）在我编写本书时所提供的技术帮助。在我们优秀的开发编辑 Brian MacDonald 的帮助下，本书的质量得以大大提高。再一次，我必须感谢百事可乐公司在我这段时间的日日夜夜给我提供的充足咖啡因。

——Brian Knight

我要感谢我的妻子，有了她一切都变得可能，还感谢我们的孩子 Eawn，他是世界上最可爱的宝贝，我可以这么说，不可以吗？我还要感谢 Microsoft 的 SSIS 团队，尤其是 Donald Farmer、Ashvini Sharma 和 Kirk Haselden，因为如果没有他们的工作，本书便没有编写的必要。

——Allan Mitchell

我要感谢我的妻子 Teri 是如此的耐心，而且在我专注编写本书时主动减少了外出购物时间。我还要感谢位于 Redmond 的团队为我解答了所有的问题，并且如此慷慨地奉献他们的宝贵时间。

——Darren Green

首先，我要感谢上帝给予我的不断祝福。对于我美丽的妻子 Misty，感谢你在这个项目期间给予我的支持和理解。你是一位我永远可以依赖的优秀妻子和孩子他妈。对于我的儿子 Kyle 和女儿 Mariah，你们是我的灵感，我爱你们。对于我的父母，我要感谢他们灌输给我关于坚持不懈和辛勤工作的价值观。感谢你 Jenny，我的妹妹兼朋友，还要感谢所有其他的家庭成员给予我的爱和支持。谢谢 Brian MacDonald、Ashvini Sharma 和 Allen Mitchell 辛勤阅读了我写的章节内容并提供了你们的建议和观点。我要非常感谢本书的作者团队和 Brian Knight 邀请我加入这个项目并给我这个机会，我十分享受这个过程。

——Douglas Hinson

我要感谢我的家人、朋友和同事给予我的鼓励并和我一起分享我参与此项目的兴奋和喜悦。谢谢 Doug Wilmsmeyer 在十年前建议我学习 VB 和 SQL Server。感谢我的兄弟 Bill Morgan, Jr，他在 1996 年教授我编程逻辑，并让我第一次尝试 ASP 编程。但最重要的，感谢你 Dennis，我的丈夫、我的伴侣、我生命中的挚爱。由于你为我做的一切，让我能够活在我的梦中。

——Kathi Kellenberger

我要首先感谢我优秀的妻子 Christy，她在我投身此项目时便一同参与到其中，而且为本书中我编写的部分做出了很多的贡献。Christy，感谢你的坚定支持！感谢我的儿子 Stevie，你放弃了很多的玩耍时间让你的父亲能够专心于编写本书，感谢可爱的 Emma。还要感谢 Manda 和 Penny 给予我的支持和祈祷。感谢本书团队成员的适应性和灵感，尤其是 Mike Potts、Jason Gerard、Doug Hinson、Mike Murphy 和 Ron Pizur。最后，我要感谢 Brian Knight 提供给我的例子、友谊、领导力以及编写本书部分内容的机会。

——Andy Leonard

有多个原因需要感谢 Microsoft Integration Service 开发团队，首先，感谢你们的远见并创建了一个伟大的产品，它已经在业界引起了很大的反响。其次，感谢 Donald Farmer 和 Ashvini Sharma（他们都在 Microsoft 开发团队）在我 2003 年夏季第一次接触 Integration Service 以来给予我的友好合作，其中包括忍受我经常性的唠叨和各种幼稚的问题；讨论设计方案；和客户一起协同工作来创建成功案例。我在本书中由我负责的部分加入了这些讨论内容和实际的教训。我还要感谢 Mark Chaffin，业界的一位主要贡献者，感谢他介绍我参与到本书的编写中，以及在有关 SSIS 产品实际应用的许多白板设计课程中给我的帮助。

——Erik Veerman

感谢我的妻子 Sandy 能够忍受我这么长时间的深夜挑灯写书。我要感谢我的儿子 Jakob，他总是在我需要时给我带来快乐。十分感谢 Doug Hinson 检查我的工作，感谢 Chad Crisostomo 对我语法方面的批评。感谢 Mike Potts 的支持。最后，感谢 Brian Knight 给我这个机会，还要感谢 Andy Leonard 说服我参与本书的编写。

——Jason Gerard

我要感谢这些年来帮助过我的许多人。感谢我的父母的辛勤工作和坚持不懈，并且在十分困难的环境下给予我们的教育。感谢我的兄弟以及其家庭给我的帮助和关心。感谢 Brian Knight 指引我走向技术写作；我很感谢他的帮助。感谢我们的编辑 Brian MacDonald 在本书中的耐心工作和优秀的编辑指导。最后，感谢 Maria 和 Benjamin，他们无疑是我生命中最宝贵的。Maria，感谢你为我做的一切，也感谢你忍受我的一切。Benjamin，感谢你给我们的生活带来如此多的欢乐和满足感。

——Haidong Ji

我要感谢我的父母 Barb 和 Jim，以及我的兄弟 Tom，感谢他们在我生命中的支持。感谢 Sheri 和 Nichole 对于我的持久信任。我还要感谢 Brian Knight 提供给我这个机会，让我的视野能够拓展到写作领域，我要感谢 Andy Leonard 给予我的动力，最后，十分感谢我所有的朋友和同事。

——Mike Murphy

内容介绍

SQL Server Integration Services (SQL Server 集成服务, SSIS) 自从它诞生以来, 现在正在历经它的第三次而且是最大的一次变革。它从 SQL Server 的一个不起眼的特性转变为提取转换加载 (Extract Transform Load, ETL) 领域的主要成员。这种变革同时也导致了该产品用户对象的变革。曾经是 DBA (数据库管理员) 的特定使用对象现在变成了由 SQL Server 开发人员使用, 而且还有一些可能自己都没有意识到正在使用该产品的不定用户群体。

SSIS 最令人高兴的是它的价格: 如果用户购买了 SQL Server, 则可以免费获得它。相比之下, 许多 ETL 厂商需要收取成百上千美元的费用。此外, SSIS 还是一个很好的平台来进行以后的扩展和集成, 这是许多 ETL 厂商所未提供的。在你经过了最初的学习曲线之后, 将会惊讶于该工具的强大功能, 而且它将为你节省大量的产品开发周期。

本书读者对象

凭借在 SSIS 发展过程中数年的使用经验, 有一个强烈的念头催促我来编写此书。如果你在过去曾经使用过 DTS, 则恐怕需要丢弃原来的旧知识而从头开始。在 SSIS 的这一版本中几乎没有保留 DTS 最初的内容。Microsoft 花了五年的时间, 通过多个版本将 SSIS 环境打成一个全新的企业级 ETL 工具。因此, 如果你认为自己对于 DTS 相当精通, 则需要重新对自己有一个正确的认识。

本书面向开发人员、DBA 以及那些希望使用 SSIS 来转换数据、创建 workflow 或者是维护 SQL Server 的用户。本书是一本专业书籍, 意味着作者假定你对于如何查询 SQL Server 具有基本了解, 而且具有一些初步的编程技能。你并不需要太多的编程技能, 不过这方面的背景有助于你的提高。此外不需要你对 SSIS 以前的版本 (那时还称为 DTS) 有所了解, 但是在本书中一些需要强调性能得到改进的地方我们会提到它。

本书组织结构

本书的前 4 章的组织结构很像是一个说明书的形式, 这部分内容奠定了后面章节的基础。从第 5 章开始, 我们在解释某个特性时将说明如何执行某个任务。SSIS 是一个具有丰富特性的产品。而且本书花了大量篇幅来介绍该产品:

第 1 章介绍了我们将在本书其他所有部分中讨论的基本概念。在该章中我们谈论了 SSIS 体系结构，并概述了用户使用 SSIS 所能够完成的工作。

第 2 章说明了如何通过使用 Import and Export Wizard（导入和导出向导）来导入和导出数据，然后介绍了商业智能开发工具集（Business Intelligence Development Studio, BIDS）。

第 3 章讨论了 SSIS 中各种可用的任务。

第 4 章讨论了如何使用容器在 SSIS 中完成循环，并说明了如何配置每种基本的转换。

在了解了如何配置大多数任务和转换后，第 5 章使用一个比较大的例子将它们组合到一起让你亲身体验 SSIS。

第 6 章介绍了一些高级的任务和转换，它们由于比较复杂，因此在前 3 章并没有深入讨论。

第 7 章介绍了在 SSIS 中使用 Script 任务的一些方式。该章还谈到了表达式。

有时候你可能需要连接到除 SQL Server 之外的系统，第 8 章则说明了如何连接到 SQL Server 之外的系统，例如 Excel、XML 和 Web 服务。

第 9 章演示了如何对 SSIS 进行扩充并使它更可靠。你可以使用本章中的特性来使程序包在出现问题时能够在故障点重新启动。

第 10 章介绍了 Data Flow 缓存体系结构以及如何监视 Data Flow 的执行。

第 11 章说明了如何对 Data Flow 进行性能调整以及一些最佳编程方法。

第 12 章说明了如何将 DTS 2000 程序包移植到 SSIS，以及在必要时如何在 SSIS 下运行 DTS 2000 程序包。此外该章还讨论了元数据管理。

第 13 章讨论了如何使用错误和事件处理功能来处理 SSIS 中的问题。

第 14 章说明了 SSIS 对象模型，以及如何使用它来扩展 SSIS。这章还说明了如何创建你自己的组件，然后在第 15 章为该组件增加了一个用户界面。

第 16 章的内容则是创建一个与 SSIS 接口的应用程序来管理环境，该章还讨论了 WMI 任务。

第 17 章介绍了如何将 SSIS Data Flow 提供给其他的程序，例如 InfoPath、Reporting Service 以及你自己的 .NET 应用程序。

第 18 章介绍了一种软件开发生命周期方法学。该章讨论了 SSIS 如何与 Visual Studio Team Systems 集成。

第 19 章的内容是一个可编程案例分析，其中创建了用于某个银行应用程序的三个 SSIS 程序包。

使用本书所需要的条件

要学习本书，只需安装 SQL Server 2005 和 Integration Services 组件。此外还需要一台可满足运行 SQL Server 2005 而所需最小硬件配置的计算机。另外，可能还希望安装 AdventureWorks 和 AdventureWorksDW 数据库（对于第 14 章和第 15 章的内容，用户将还

需要 Visual Studio 2005 和 C#来运行示例)。

本书约定

为了有助于你对本书有更深入的理解，我们在全书中使用了许多约定。

这样的方框表示重要的、需要牢记的知识，这些知识与周围的内容是直接相关的。

关于当前讨论内容的技巧、提示、窍门和旁白使用楷体字表示。

关于文本的样式：

- 在介绍新术语和重要的词时用楷体字表示它们。
- 键盘按键表示法如下：Ctrl+A。
- 用 Courier 字体表示书中的代码。
- 用两种不同的方法展示代码：

在代码示例中，用灰色背景突出显示新的、重要的代码。

对于当前环境中不那么重要的代码或者以前显示过的代码，不采用灰色背景突出显示它们。

源代码

在阅读本书的例子时，可以选择手工输入所有的代码，或者使用本书附带的源代码文件。本书中使用的所有源代码都可以从 <http://www.wrox.com> 下载。在该站点上，只需定位本书的书名（通过 Search 文本框或者是使用某个书名列表），并单击本书详细信息页面上的 Download Code 链接，来获得本书的所有源代码。

由于许多图书具有类似的书名，你可能发现最容易的方法是通过 ISBN 进行搜索。本书的 ISBN 是 0-7645-9792-2。

一旦下载了代码，就可以使用所喜爱的压缩工具来对它进行解压缩。此外，也可以进入位于 www.wrox.com/dynamic/books/download.aspx 的 Wrox 代码下载主页，来查看针对本书以及所有其他 Wrox 图书的代码。

勘误表

我们尽力确保在本书的内容或代码中不出现错误。但是，错误在所难免。如果在我们

的图书中发现错误，如拼写错误或代码段错误，我们将非常感谢你提供反馈。通过将你发现的错误以勘误表形式发送给读者，可能会为其他读者节省宝贵的时间，同时你也是在帮助我们提供更高质量的内容。

要找到本书的勘误表页面，可以访问 <http://www.wrox.com> 并使用 Search 文本框或书名列表定位本书的书名，然后在图书的详细信息页面上单击 Book Errata 链接。在该页面上，用户可以看到本书当前由 Wrox 编辑提交的所有勘误信息。此外，在 www.wrox.com/misc-pages/booklist.shtml 上可以找到一份完整的图书列表，其中包括指向每本图书勘误表的链接。

如果在 Book Errata 页面上没有找到你所发现的错误，则可以访问 www.wrox.com/contact/techsupport.shtml 并填写这里的表格，将你所发现的错误发送给我们。我们将检查这些信息，如果合适，就会把它们发布到图书的勘误页面，并在图书的后续版本中纠正这些错误。

p2p.wrox.com

为了便于你和作者进行对等交流，建议你加入位于 p2p.wrox.com 的 P2P 论坛。该论坛是一个基于 Web 的系统，可以让你发布有关 Wrox 图书以及相关技术的消息，和其他的读者以及技术用户交流。该论坛提供了一种订阅功能，当论坛中发布了你感兴趣的主题时，可以通过电子邮件发送给你。Wrox 的图书作者、编辑、其他业界专家以及志同道合的读者都有可能出现在该论坛中。

在 <http://p2p.wrox.com> 上，你将会发现各种不同的论坛，它们不仅可以帮助你阅读本书，还可以帮助你开发自己的应用程序。要加入这些论坛，请遵循以下的步骤：

1. 访问 p2p.wrox.com，单击 Register 链接。
2. 阅读使用条款并单击 Agree。
3. 填写加入论坛所需的信息以及任何你愿意提供的可选信息，并单击 Submit。
4. 你将接收到一封电子邮件，它说明了如何验证你的账户并完成加入过程。

你可以在不加入 P2P 的情况下读取论坛中的信息，但是若要发布你自己的消息，则必须加入。

一旦你加入后，就可以发布新消息，并对其他用户发布的消息做出响应，你可以在任何时候阅读 Web 上的消息。如果你希望通过电子邮件方式将特定论坛的新消息发送给你，则可通过论坛列表中的论坛名称，单击 Subscribe to this Forum 图标。

有关如何使用 Wrox P2P 的更多信息，可以阅读 P2P FAQ，来了解关于论坛软件如何工作等问题以及其他与 P2P 和 Wrox 图书有关的许多常见问题的答案。要阅读 FAQ，可以单击 P2P 页面上的 FAQ 链接。

前 言

我第一次开始管理数据转换服务团队是在 2001 年。这之前曾经在分析服务 (Analysis Service) 团队工作。我在 DTS 方面没有多少背景,但是对此产品却非常狂热,而且愿意学习新的知识并充满热情地迎接新的工作。那时我面临的问题是,什么方式能够使我在短时期内迅速熟悉此产品呢?

在我四处寻求此问题的答案时,我的那些新同事们几乎都推荐了“红皮书”,这本书就是 Brian Knight 和 Mark Charrin 所编著的 *Professional DTS* 一书。诚如其言,该书内容广泛、详细,并提供了介绍清晰、易于理解的例子。我认为该书对于任何希望从头开始学习 DTS 的人都是无价之宝。

自那时起,一些年过去了,而且 DTS 也发展成了 SQL Server 集成服务 (SSIS)。这两种产品的哲学基础以及以客户为中心的理念是一致的,当然它们的出身也都是出自 Microsoft 名门。但是 SSIS 是一个完全不同的产品,它是在一个完全不同于 DTS 的领域扮演自己的角色。实际上,DTS 是 SQL Server 的一个很流行的功能,几乎所有需要移动数据和表格的用户都会使用到它。实际上,根据一些调查的结果,超过 70% 的 SQL Server 用户都使用 DTS。由于 DTS 是如此流行,因此可能有人会问为什么还要重新编写该产品来创建 SSIS 呢?

这个问题的答案几乎也就是 SSIS/DTS 团队的定义:倾听客户的意见。我们多次听到很多客户说虽然他们喜欢 DTS,但他们感觉到还需要购买一套作为补充的 ETL 产品,尤其是在高端的企业中。此外,我们还听到有关 DTS 在性能、可伸缩性、复杂性以及可扩展性方面的意见。客户希望从 DTS 获得更多。同时,本书的作者也给我们提供了很多反馈意见。同时我个人也从 Mark Chaffin 获得了很多从 DTS 发展到 SSIS 的反馈。伴随针对 DTS 产品的功能、性能和可伸缩性等方面的改进需求,以及为了适应在开始业界崭露头角的 .NET 和托管代码体系结构。所有这些导致了一个确定的逻辑结论,也就是需要从头开始创建一个新的产品,而不是在旧有的体系结构上对 DTS 进行调整。在我们发布了 SQL 2000 之后,便逐渐开始了将 DTS 提升到更高层次的工作。

幸运的是,对于新的产品究竟应该是怎样的,我们获得了一些很重要的构想和方向。Euan Garden (曾经担任 DTS 项目经理)、Gert Drapers (曾经是 DTS 的架构师/经理)、Jag Bhalla (曾经任职于一家被 Microsoft 所收购的公司) 以及 Bill Baker (负责指导 SQL Server 所有业务智能工作的总经理) 提供了最初的产品开发方向并提出了 SSIS 的发展路线图。DTS 团队仍然是 Management Tools 团队的一部分,而且直到 2001 年它才成为一个单独的团队。这个团队规模很小,但是它具有一个明确而且非常重要的使命:通过开发一个业界

领先的 ETL/数据集成平台来完成 SQL BI “功能栈”。

我在 2001 年夏季开始管理这个担负巨大使命的团队，那时我的工作目标只有一个：一定要完成这个使命！最初的团队规模很小，但是成员都非常有才能。这些人包括 Mark Blaszczak，我曾经见过的最多产的开发人员；Jag Bhalla，一位精通业务的数据仓库业界老手；James Howey，一位十分精通数据管线的资深技术产品经理；Krik Haselden，一位天生的领导者而且是要求严格的开发人员；Ted Lee，一位开发过 SQL Server 以前两个版本、经验丰富老练的开发人员（他是唯一一位真正懂得旧 DTS 代码库的人员！）。我们通过外部雇佣和内部调配创建了开发团队，而且很快补充了大多数的职缺。这里特别要提到的是加入我们团队的新成员，包括 Donald Farmer，一位极有才干而且面向客户的总经理，现在他在很大程度上成为了 SSIS 的等义词；Ashvini Sharma，UI 开发组长，从不气馁而且善于从客户角度考虑问题；还有 Jeff Bernhard，开发经理，他所开发的“宠物计划（pet project）”曾经让我们焦虑一时，但最终显著改进了产品的功能。

在我们接手此产品时，已经发布了 Beta 1 版本。在 Beta 1 之后，我们便着手开发现在所称的 SSIS。从开发过程中的某个阶段开始，我们能感觉到自己所创建的产品已经不再是 DTS，它在各个方面已经超越了 DTS。在经过多次内部辩论之后，我们决定对产品改名。但是叫它什么呢？当时有各种建议的名称（例如，METL），我们对所有候选名称都进行了一番评估。最后，我们选择了一个在命名过程中很早就被建议，简单但却含义丰富的名称：Integration Service（前面使用 SQL Server 作为前缀来表明它是针对 SQL Server 数据的）。

该 DTS 是更大的 SQL BI 组的一部分，而且 SSIS 的设计在很多层面上反映了这一血统。我最早与 DTS 打交道是在 Yukon（SQL 2005）的最初规划期间，当时我是一个小团队的一员，负责为 DTS 设计人员提供有关用户产品体验方面的改进意见。DTS 允许与 OLAP 以及数据挖掘技术深入集成的巨大潜能从一开始就吸引了我（而且这种“超越 ETL”吸引力一直在继续——可以访问 www.byondetl.com）。在本书的第 6 章和第 4 章（该章很好地介绍了新的 Data Flow 任务以及它的组件）对这种集成进行了讨论。SSIS 另外一个相关的关键部分是无论在脚本编程还是创建自定义组件（任务和转换）等方面都具有很好的可扩展性。由 Darren 和 Allen（他还开办了 SQLIS.com 网站，而且是我们的 MVP）所编写的本书第 14 章对此有非常到位的介绍。

我还要补充的是当我以第一人称写此前言而且试图为 SSIS 的开发提供一些建议时，我在团队中的角色只是一个支持人员，而且此产品无疑是团队集体智慧的结晶：辛勤工作、专注、着眼于客户以及谦逊是我们这个团队的特点，实际上，其中的许多成员（Runying Mao、James Howey、Ashvini Sharma、Bob Bojanic、Ted Lee 和 Grant Dickinson）都帮助审校了本书以保证技术上的准确性。在长达五年忙碌紧张的开发过程期间，他们在百忙之中抽出时间来审校了本书！

我猜想在你阅读到本书时我们已经开发完成了 SQL 2005，发布这个我个人认为具有很多优秀特性的伟大产品，是一个漫长但却有价值的旅程。SSIS 是 SQL Server 2005 的一个

重要附加部分。而且本书将帮助你熟练掌握它。SSIS 很容易入门，但它是一个非常深入而且功能丰富的产品，具有一定的复杂性。本书将尽可能地为你展现 SSIS 所提供的广泛价值。我真诚地希望你能够愉快地阅读本书，尽兴地使用 SSIS。

Kamal Hathi
SSIS 产品单元经理

目 录

第 1 章 欢迎使用 SQL Server 集成服务1	
1.1 SQL Server 2005 SSIS 中的新特性.....1	
1.1.1 导入和导出向导.....2	
1.1.2 商业智能开发工具集.....2	
1.2 体系结构.....3	
1.2.1 程序包.....5	
1.2.2 任务.....5	
1.2.3 数据源元素.....6	
1.2.4 数据源视图.....7	
1.3 优先约束.....8	
1.3.1 约束值.....8	
1.3.2 条件表达式.....8	
1.4 容器.....9	
1.5 变量.....10	
1.6 数据流元素.....10	
1.6.1 源.....10	
1.6.2 目的地.....11	
1.6.3 转换.....12	
1.7 错误处理和日志记录.....13	
1.8 SQL Server 2005 的版本.....14	
1.9 小结.....16	
第 2 章 SSIS 工具17	
2.1 Import and Export Wizard.....17	
使用 Import and Export Wizard.....17	
2.2 Package Installation Wizard.....21	
2.3 Business Intelligence Development Studio.....21	
2.4 创建自己的第一个程序包.....23	
2.5 Solution Explorer Window.....24	
2.5.1 Toolbox.....25	
2.5.2 Properties Windows.....26	
2.5.3 Navigation Pane.....26	
2.5.4 其他窗口.....27	
2.6 SSIS Package Designer.....27	
2.6.1 Controller Flow.....28	
2.6.2 Connection Manager.....30	
2.6.3 变量.....31	
2.6.4 数据流.....32	
2.6.5 Event Handler.....33	
2.6.6 Package Explorer.....34	
2.6.7 执行程序包.....34	
2.7 小结.....35	
第 3 章 SSIS 任务36	
3.1 共享属性.....36	
3.2 Execute SQL 任务.....37	
3.3 Bulk Insert 任务.....38	
3.4 使用 Bulk Insert 和 Execute SQL 任务.....41	
3.5 Data Flow 任务.....43	
3.6 Execute Process 任务.....43	
3.7 File System 任务.....44	
3.8 FTP 任务.....45	
3.9 使用 File System 和 FTP 任务.....46	
3.10 Execute Package 任务.....48	
3.11 Script 和 ActiveX 任务.....49	
3.12 Send Mail 任务.....51	
3.13 Message Queue 任务.....52	
3.14 Web Service 任务.....52	