

# Contents

<b>Acknowledgments</b>	<b>xi</b>
<b>Foreword</b>	<b>xxiii</b>
<b>Introduction</b>	<b>xxv</b>
<b>Who This Book Is For</b>	<b>xxv</b>
<b>How This Book Is Structured</b>	<b>xxv</b>
<b>What You Need to Use This Book</b>	<b>xxvi</b>
<b>Conventions</b>	<b>xxvii</b>
<b>Source Code</b>	<b>xxvii</b>
<b>Errata</b>	<b>xxvii</b>
<b>p2p.wrox.com</b>	<b>xxviii</b>
<b>Chapter 1: Welcome to SQL Server Integration Services</b>	<b>1</b>
<b>What's New in SQL Server 2005 SSIS</b>	<b>1</b>
Import and Export Wizard	2
The Business Intelligence Development Studio	3
<b>Architecture</b>	<b>3</b>
Packages	5
Tasks	5
Data Source Elements	6
Data Source Views	8
<b>Precedence Constraints</b>	<b>9</b>
Constraint Value	9
Conditional Expressions	9
<b>Containers</b>	<b>10</b>
<b>Variables</b>	<b>10</b>
<b>Data Flow Elements</b>	<b>11</b>
Sources	11
Destinations	12
Transformations	13
<b>Error Handling and Logging</b>	<b>14</b>
<b>Editions of SQL Server 2005</b>	<b>16</b>
<b>Summary</b>	<b>17</b>

<b>Chapter 2: The SSIS Tools</b>	<b>19</b>
<b>Import and Export Wizard</b>	<b>19</b>
Using the Import and Export Wizard	19
<b>Package Installation Wizard</b>	<b>25</b>
<b>Business Intelligence Development Studio</b>	<b>25</b>
<b>Creating Your First Package</b>	<b>27</b>
<b>The Solution Explorer Window</b>	<b>28</b>
The Toolbox	29
The Properties Windows	30
Navigation Pane	31
Other Windows	32
<b>The SSIS Package Designer</b>	<b>32</b>
Controller Flow	33
Connection Managers	36
Variables	37
Data Flow	38
Event Handlers	39
Package Explorer	40
Executing a Package	41
<b>Summary</b>	<b>41</b>
<b>Chapter 3: SSIS Tasks</b>	<b>43</b>
<b>Shared Properties</b>	<b>43</b>
<b>Execute SQL Task</b>	<b>44</b>
<b>Bulk Insert Task</b>	<b>46</b>
<b>Using the Bulk Insert and Execute SQL Tasks</b>	<b>48</b>
<b>Data Flow Task</b>	<b>51</b>
<b>Execute Process Task</b>	<b>51</b>
<b>File System Task</b>	<b>52</b>
<b>FTP Task</b>	<b>54</b>
<b>Using the File System and FTP Task</b>	<b>55</b>
<b>Execute Package Task</b>	<b>59</b>
<b>Script and ActiveX Tasks</b>	<b>60</b>
<b>Send Mail Task</b>	<b>62</b>
<b>Message Queue Task</b>	<b>63</b>
<b>Web Service Task</b>	<b>63</b>
<b>WMI Data Reader and Event Watcher Task</b>	<b>65</b>
<b>XML Task</b>	<b>68</b>
<b>SQL Server Analysis Services Execute DDL and Processing Tasks</b>	<b>70</b>
<b>Data Mining Query Task</b>	<b>71</b>

---

<b>The Expression Page</b>	<b>72</b>
<b>Summary</b>	<b>72</b>
<b>Chapter 4: Containers and Data Flow</b>	<b>73</b>
<b>Containers</b>	<b>73</b>
Task Host Containers	73
Sequence Containers	73
For Loop Container	74
Foreach Loop Container	78
<b>Sources</b>	<b>80</b>
OLE DB Source	80
Excel Source	83
Flat File Source	84
Raw File Source	84
XML Source	84
Data Reader Source	84
<b>Destinations</b>	<b>84</b>
Data Mining Model Training	86
DataReader Destination	86
Dimension and Partition Processing	87
Excel Destination	87
Flat File Destination	88
OLE DB Destination	88
Raw File Destination	89
Recordset Destination	89
SQL Server and Mobile Destinations	90
<b>Transformations</b>	<b>90</b>
Aggregate	91
Audit	93
Character Map	94
Conditional Split	94
Copy Column	97
Data Conversion	97
Data Mining Query	98
Derived Column	99
Export Column	100
Fuzzy Grouping and Lookup	101
Import Column	101
Lookup	101
Merge	102
Merge Join	103

# Contents

---

Multicast	104
OLE DB Command	105
Percentage and Row Sampling	105
Pivot and Unpivot	106
Row Count	107
Script Component	107
Slowly Changing Dimension	107
Sort	107
Term Extraction and Lookup	108
Union All	109
<b>Data Flow Example</b>	<b>110</b>
<b>Data Viewers</b>	<b>116</b>
<b>Summary</b>	<b>117</b>
<b>Chapter 5: Creating an End-to-End Package</b>	<b>119</b>
<b>Basic Transformation Tutorial</b>	<b>119</b>
Creating Connections	120
Creating the Tasks	122
Creating the Data Flow	123
Completing the Package	124
Saving the Package	125
Executing the Package	125
<b>Typical Mainframe ETL with Data Scrubbing</b>	<b>125</b>
Creating the Data Flow	129
Handling Dirty Data	129
Finalizing	133
Handling More Bad Data	134
<b>Looping and the Dynamic Task</b>	<b>138</b>
Looping	139
Making the Package Dynamic	140
<b>Summary</b>	<b>142</b>
<b>Chapter 6: Advanced Tasks and Transforms</b>	<b>143</b>
<b>Execute SQL Task</b>	<b>144</b>
Variables	145
Expressions	146
Using SQL Output Parameters to Change Runtime Settings	148
<b>Import Column</b>	<b>150</b>
Import Column Example	151
Import Column Example Using File Iteration	155

<b>Using Temp Tables in SSIS Package Development</b>	<b>157</b>
<b>Export Column</b>	<b>159</b>
<b>Row Count</b>	<b>160</b>
<b>OLE DB Command</b>	<b>162</b>
<b>Term Extraction</b>	<b>165</b>
<b>Term Lookup</b>	<b>171</b>
<b>Fuzzy Lookup</b>	<b>173</b>
<b>Fuzzy Grouping</b>	<b>178</b>
<b>Pivot Transform</b>	<b>182</b>
<b>Unpivot</b>	<b>186</b>
<b>Slowly Changing Dimension</b>	<b>189</b>
<b>Database Object-Level Tasks</b>	<b>201</b>
Transfer Database Task	201
Transfer Logins Task	202
Transfer Master Stored Procedures Task	203
Transfer SQL Server Objects Task	204
<b>Summary</b>	<b>205</b>
<b>Chapter 7: Scripting in SSIS</b>	<b>207</b>
<hr/>	
<b>Scripting Overview</b>	<b>207</b>
<b>Expressions</b>	<b>208</b>
Dynamic Properties	208
Expressions in Tasks	211
<b>Script Tasks</b>	<b>213</b>
The Dts Object	218
Accessing Variables	219
Events	220
Logging	222
Script Task Debugging	222
<b>Using .NET Assemblies</b>	<b>225</b>
<b>Structured Exception Handling</b>	<b>227</b>
<b>Script Component</b>	<b>228</b>
Using the Script Component	229
Debugging the Script Component	235
<b>Summary</b>	<b>237</b>
<b>Chapter 8: Accessing Heterogeneous Data</b>	<b>239</b>
<hr/>	
<b>Excel</b>	<b>240</b>
Exporting to Excel	240
Importing from Excel	244

# Contents

---

<b>Access</b>	<b>250</b>
Understanding Access Security	250
Configuring an Access Connection Manager	251
Importing from Access	253
Using a Parameter	255
<b>Oracle</b>	<b>259</b>
Client Setup	259
Importing Oracle Data	259
<b>Web Services</b>	<b>262</b>
The Hyperlink Extractor Service	262
The Currency Conversion Service	265
<b>XML Data</b>	<b>272</b>
<b>Summary</b>	<b>278</b>
<b>Chapter 9: Reliability and Scalability</b>	<b>279</b>
<b>Restarting Packages</b>	<b>279</b>
Simple Control Flow	280
Containers within Containers and Checkpoints	285
Variations on a Theme	288
Inside the Checkpoint File	290
<b>Package Transactions</b>	<b>292</b>
Single Package, Single Transaction	293
Single Package, Multiple Transactions	296
Two Packages, One Transaction	298
Single Package Using a Native Transaction in SQL Server	299
<b>Error Outputs</b>	<b>301</b>
<b>Scaling Out</b>	<b>304</b>
Scale Out Memory Pressures	304
Scale Out by Staging Data	305
<b>Summary</b>	<b>310</b>
<b>Chapter 10: Understanding the Integration Services Engine</b>	<b>311</b>
<b>The Integration Services Engine: An Analogy</b>	<b>311</b>
<b>Understanding the SSIS Data Flow and Control Flow</b>	<b>312</b>
Comparing and Contrasting the Data Flow and Control Flow	312
SSIS Package Execution Times from Package Start to Package Finish	314
<b>Enterprise Workflows with the Control Flow</b>	<b>315</b>

---

<b>Enterprise Data Processing with the Data Flow</b>	<b>317</b>
Memory Buffer Architecture	317
Types of Transformations	318
Advanced Data Flow Execution Concepts	329
<b>Summary</b>	<b>339</b>
<b>Chapter 11: Applying the Integration Services Engine</b>	<b>341</b>
<b>That Was Then: DTS</b>	<b>341</b>
DTS Solution Architecture	342
Common DTS Processing Practices	343
DTS Limitations	345
<b>This Is Now: Integration Services</b>	<b>346</b>
Integration Services Design Practices	347
Optimizing Package Processing	360
<b>Pipeline Performance Monitoring</b>	<b>366</b>
<b>Summary</b>	<b>369</b>
<b>Chapter 12: DTS 2000 Migration and Metadata Management</b>	<b>371</b>
<b>Migrating DTS 2000 Packages to SSIS</b>	<b>371</b>
<b>Using the Package Migration Wizard</b>	<b>373</b>
<b>Running DTS 2000 Packages under SSIS</b>	<b>380</b>
<b>Package Metadata and Storage Management</b>	<b>384</b>
Managing SSIS Packages within SQL Server Management Studio	386
Managing DTS 2000 Packages within SQL Server Management Studio	388
<b>Summary</b>	<b>389</b>
<b>Chapter 13: Error and Event Handling</b>	<b>391</b>
<b>Precedence Constraints</b>	<b>391</b>
Precedence Constraints and Expressions	393
Multiple Constraints	397
<b>Event Handling</b>	<b>398</b>
Events	398
Inventory Example	399
Event Bubbling	406
<b>Breakpoints</b>	<b>408</b>
<b>Checkpoints</b>	<b>411</b>
<b>Logging</b>	<b>412</b>
<b>Summary</b>	<b>416</b>

<b>Chapter 14: Programming and Extending SSIS</b>	<b>417</b>
<b>The Sample Components</b>	<b>417</b>
Component 1: Source Adapter	418
Component 2: Transformation	418
Component 3: Destination Adapter	419
<b>The Pipeline Component Methods</b>	<b>419</b>
Design-Time	419
Runtime	423
Connection Time	425
<b>Building the Components</b>	<b>425</b>
Preparation	426
Building the Source Adapter	432
Building the Transform	443
Building the Destination Adapter	454
<b>Debugging Components</b>	<b>461</b>
Design-Time	462
Runtime	463
<b>Summary</b>	<b>466</b>
<b>Chapter 15: Adding a User Interface to Your Component</b>	<b>467</b>
<b>Three Key Steps</b>	<b>467</b>
<b>Building the User Interface</b>	<b>468</b>
Adding the Project	468
Implementing IDtsComponentUI	471
Setting the UITypeName	475
Building the Form	476
<b>Further Development</b>	<b>481</b>
Runtime Connections	482
Component Properties	484
Handling Errors and Warnings	486
Column Properties	488
<b>Summary</b>	<b>490</b>
<b>Chapter 16: External Management and WMI Task Implementation</b>	<b>491</b>
<b>External Management with Managed Code</b>	<b>491</b>
<b>Application Object Maintenance Operations</b>	<b>492</b>
Package Maintenance Operations	493
Server Folder Maintenance	494
Package Role Maintenance	495

Package Monitoring	495
A Package Management Example	496
<b>Package Log Providers</b>	<b>505</b>
Specifying Events to Log	506
Programming Log Providers	507
<b>Package Configurations</b>	<b>509</b>
Creating a Configuration	510
Programming the Configuration Object	511
Configuration Object	512
<b>Windows Management Instrumentation Tasks</b>	<b>512</b>
WMI Reader Task Explained	513
WMI Event Watcher Task	514
WMI Data Reader Example	515
WMI Event Watcher Task Example	521
<b>Summary</b>	<b>522</b>
<hr/> <b>Chapter 17: Using SSIS with External Applications</b>	<hr/> <b>523</b>
<b>RSS In, Reporting Services Report Out</b>	<b>524</b>
<b>InfoPath Document</b>	<b>532</b>
<b>ASP.NET Application</b>	<b>540</b>
<b>Summary</b>	<b>545</b>
<hr/> <b>Chapter 18: SSIS Software Development Life Cycle</b>	<hr/> <b>547</b>
<b>Introduction to Software Development Life Cycles</b>	<b>548</b>
Software Development Life Cycles: A Brief History	548
Types of Software Development Life Cycles	549
<b>Versioning and Source Code Control</b>	<b>550</b>
Microsoft Visual SourceSafe	550
Team Foundation Server, Team System, and SSIS	565
MSF Agile and SSIS	570
The Project Portal	573
Putting It to Work	573
<b>Code Deployment and Promotion from Development to Test to Production</b>	<b>587</b>
The Deployment Wizard	588
Import a Package	589
<b>Summary</b>	<b>591</b>

<b>Chapter 19: Case Study: A Programmatic Example</b>	<b>593</b>
<b>Background</b>	<b>593</b>
<b>Business Problem</b>	<b>594</b>
<b>Solution Summary</b>	<b>594</b>
<b>Solution Architecture</b>	<b>595</b>
<b>Data Architecture</b>	<b>599</b>
File Storage Location Setup	599
Bank ACH Payments	599
Lockbox Files	600
PayPal or Direct Credits to Corporate Account	601
Case Study Database Model	601
Database Setup	602
<b>Case Study Load Package</b>	<b>610</b>
Naming Conventions and Tips	611
Package Setup and File System Tasks	612
Lockbox Control Flow Processing	616
Lockbox Validation	622
Lockbox Processing	631
ACH Control Flow Processing	635
ACH Validation	639
ACH Processing	643
E-mail Payment Processing	645
E-mail Data Flow Processing	648
Testing	651
<b>Case Study Process Package</b>	<b>651</b>
Package Setup	652
High-Confidence Data Flow	654
Medium-Confidence Data Flow	659
Interpreting the Results	663
<b>Running in SQL Agent</b>	<b>664</b>
<b>Summary</b>	<b>665</b>
<b>Index</b>	<b>667</b>