

Wee-Hyong Tok
Rakesh Parida
Matt Masson
Xiaoning Ding
Kaarthik Sivashanmugam

Microsoft® SQL Server® 2012 Integration Services

Przekład: Marek Włodarz

APN Promise, Warszawa 2012

Spis treści

<i>Przedmowa</i>	xiii
<i>Wprowadzenie</i>	xv
<i>O autorach</i>	xx

Część I: Przegląd

1	Ogólna charakterystyka SSIS	3
	Typowe scenariusze wykorzystania SSIS	4
	Konsolidowanie danych z heterogenicznych źródeł	4
	Przenoszenie danych pomiędzy systemami	9
	Ładowanie danych do hurtowni	14
	Czyszczenie, formatowanie lub standaryzowanie danych	18
	Identyfikowanie, przechwytywanie i przetwarzanie zmian danych	19
	Koordynowanie konserwowania, przetwarzania lub analizowania danych	21
	Ewolucja SSIS	24
	Instalowanie SSIS	25
	Funkcje SQL Server wymagane dla integrowania danych	26
	Wydania SQL Server a funkcje Integration Services	28
	Podsumowanie	30
2	Koncepcja SSIS	31
	Przepływ sterowania	32
	Zadania	32
	Ograniczenia pierwszeństwa	34
	Zmienne i wyrażenia	36
	Kontenery	37
	Menedżery połączeń	40
	Pakiety i projekty	42
	Parametry	43
	Dostawcy dzienników	44
	Obsługa zdarzeń	45
	Przepływ danych	46
	Adaptory źródłowe	47
	Adaptory docelowe	48
	Transformacje	49
	Katalog SSIS	50

Wprowadzenie	51
Katalog	52
Foldery	52
Środowiska	53
Odsyłacze	53
Podsumowanie	54
3 Wykonywanie aktualizacji do wersji SSIS 2012	55
Co nowego w SSIS 2012?	55
Uwarunkowania i planowanie aktualizacji	56
Zmiany funkcjonalności SSIS	56
Uwarunkowania i narzędzia	58
Wymagania dotyczące aktualizacji	59
Scenariusze aktualizacji	60
Nieobsługiwane scenariusze aktualizacji	61
Weryfikacja aktualizacji	62
Aktualizacja Integration Services	63
Upgrade Advisor	63
Wykonywanie uaktualnienia	68
Rozwiązywanie problemów z aktualizacją i ręczne aktualizowanie pakietów ..	78
Konwersja na projekty po aktualizacji	80
Podsumowanie	88

Część II: Projektowanie

4 Nowe funkcje narzędzi projektowania SSIS	91
Środowisko projektowania Integration Services	91
Visual Studio	91
Cofanie i powtarzanie zmian	92
Okno Getting Started (Zaczynamy)	92
Przybornik	93
Okno Variables	95
Kontrolka powiększania (Zoom)	96
Automatyczny zapis i przywracanie	97
Ikony statusu	97
Adnotacje	98
Konfiguracja i wdrażanie	98
Zmiany w narzędziu Solution Explorer	98
Zakładka Parameter	99
Konfiguracje Visual Studio	100
Kompilacja projektu	101
Deployment Wizard	102
Project Conversion Wizard	103

Import Project Wizard	103
Nowe zadania i komponenty przepływu danych	104
Change Data Capture	104
Zadanie Expression	106
Transformacja DQS Cleansing	108
Źródła i miejsca docelowe ODBC	108
Przepływ sterowania	108
Uwidacznianie wyrażeń	108
Menedżery połączeń	109
Zadanie Execute SQL	109
Przepływ danych	110
Asystenci połączeń	110
Ulepszone mapowanie kolumn	111
Edytowanie komponentów, które znajdują się w stanie błędu	113
Grupowanie	113
Uprozczone podglądy danych	113
Interfejsy użytkownika transformacji Row Count oraz Pivot	114
Zmiany w źródłach plikowych (Flat File)	115
Skryptowanie	117
Visual Studio Tools for Applications	117
Debugowanie komponentów skryptowych	118
Wsparcie dla .NET 4 Framework	120
Wyrażenia	120
Usunięcie limitu liczby znaków	120
Nowe funkcje wyrażeń	121
Podsumowanie	122
5 Projektowanie zespołowe	123
Ulepszenia wprowadzone w SQL Server 2012	123
Zmiany formatu pakietów	123
Konfiguracje Visual Studio	124
Korzystanie z zarządzania kodem źródłowym w SSIS	125
Łączenie się z Team Foundation Server	125
Dodawanie projektu SSIS do Team Foundation Server	128
Zarządzanie zmianami	132
Zmiany w pliku projektu Visual Studio dla SSIS	135
Najlepsze praktyki	137
Korzystanie z małych, prostych pakietów	137
Jeden pakiet – jeden projektant	138
Spójna konwencja nazewnictwa	138
Podsumowanie	138

6	Tworzenie rozwiązania SSIS	139
	Porównanie modeli wdrażania SSIS	139
	Model wdrażania pakietów	139
	Model wdrażania projektu	141
	Budowanie projektu Integration Services	144
	Tworzenie projektu SSIS	144
	Projektowanie przepływu danych Integration Services	152
	Korzystanie z parametrów i kontenera <i>ForEach</i>	156
	Korzystanie z zadania <i>Execute Package</i>	160
	Kompilowanie i wdrażanie projektu Integration Services	162
	Podsumowanie	163
7	Połączenia SSIS	165
	Opcje połączeń we wcześniejszych wersjach SSIS	165
	Dostawcy technik połączeniowych	166
	OLE DB, ADO.NET oraz ODBC	167
	Nowe opcje połączeń w SSIS 2012	170
	Wprowadzenie do ODBC	170
	Komponenty ODBC w SSIS	172
	ODBC Source	173
	ODBC Destination	178
	Uwarunkowania połączeniowe SSIS	181
	Systemy 64-bitowe i SSIS	181
	Narzędzia SSIS dla architektury 64-bitowej	183
	Łączenie się z innymi źródłami i miejscami docelowymi	187
	Połączenia z Microsoft Excel oraz Access	187
	Połączenia do baz danych Oracle	190
	Tworzenie niestandardowych komponentów	193
	Korzystanie z komponentów skryptowych	195
	Podsumowanie	197
8	Korzystanie z Change Data Capture w SSIS 2012	199
	CDC w SQL Server	199
	Stosowanie CDC w SQL Server	199
	Scenariusze wykorzystania CDC w procesach ETL	201
	Fazy CDC	202
	Komponenty CDC w SSIS 2012	206
	Stan CDC	206
	Zadanie CDC Control	209
	Komponent przepływu danych CDC Source	215
	Komponent CDC Splitter	220
	CDC dla Oracle	221

Wprowadzenie	222
Komponenty obsługujące CDC dla Oracle	223
CDC Service Configuration	223
Oracle CDC Designer	226
Baza danych MSXDBCDC	237
Plik wykonywalny Oracle CDC Service (xdbcscsvc.exe)	240
Obsługa typów plików	243
Komponenty CDC w SSIS	245
Podsumowanie	245
9 Czystczenie i profilowanie danych przy użyciu SSIS	247
Zadanie Data Profiling	247
Transformacja Fuzzy Lookup	252
Transformacja Fuzzy Grouping	258
Transformacja Data Quality Services Cleansing	262
Podsumowanie	268
 Część III: Konfiguracje, zarządzanie i monitorowanie	
10 Konfiguracje w SSIS	271
Podstawy konfiguracji	271
Jak konfiguracje są stosowane	272
Co konfigurować?	272
Konfiguracje w SSIS 2012	273
Parametry	273
Tworzenie parametrów pakietu	274
Tworzenie parametrów projektu	276
Programowe tworzenie parametrów	278
Korzystanie z parametrów	280
Konfigurowanie parametrów w katalogu SSIS	287
Konfigurowanie, weryfikowanie i wykonywanie pakietów i projektów	287
Konfigurowanie wykonania przy użyciu SSMS	287
Konfiguracje w SQL Server Agent, DTEXEC i T-SQL	291
Środowiska SSIS	293
Kolejność aplikowania parametrów	297
Model wdrażania pakietów i kompatybilność wsteczna	297
Model wdrażania pakietów	298
Najlepsze zalecenia dotyczące konfigurowania SSIS	301
Najlepsze zalecenia dla modelu wdrażania pakietów	301
Najlepsze zalecenia dla modelu wdrażania projektów	305
Podsumowanie	306

11	Wykonywanie pakietów SSIS	307
	Uruchamianie pakietów SSIS przy użyciu narzędzia DTEXEC	307
	Lokalizacja pakietów	308
	Konfigurowanie wykonywania pakietów	312
	Zrzuty pamięci	314
	Opcje rejestrowania	315
	Uruchamianie pakietów z katalogu SSIS	318
	Przygotowywanie wykonania	318
	Rozpoczynanie wykonywania pakietu SSIS	321
	Przeglądanie przebiegu wykonywania	324
	Wykonywanie pakietu za pośrednictwem T-SQL	325
	Uruchamianie pakietów za pośrednictwem SQL Server Agent	327
	Tworzenie kroku zadania SSIS	327
	Wykonanie pakietu wdrożonego w katalogu SSIS	329
	Uruchamianie pakietów za pośrednictwem PowerShell	331
	Programowe tworzenie i wykonywanie pakietów SSIS	331
	Podsumowanie	336
12	Magia T-SQL w SSIS	337
	Przegląd procedur składowanych i widoków SSIS	337
	Katalog Integration Services	338
	Właściwości katalogu SSIS	338
	Odpytywanie właściwości katalogu SSIS	339
	Ustawianie właściwości katalogu SSIS	339
	Projekty i pakiety SSIS	340
	Wdrażanie projektu SSIS w katalogu SSIS	340
	Uzyskiwanie informacji o projektach wdrożonych w katalogu SSIS	341
	Konfigurowanie projektów SSIS	343
	Zarządzanie projektami w katalogu SSIS	345
	Uruchamianie pakietów wdrożonych w katalogu SSIS	347
	Środowiska SSIS	352
	Tworzenie środowisk SSIS	352
	Tworzenie zmiennych środowiskowych SSIS	353
	Konfigurowanie projektów SSIS przy użyciu środowisk	354
	Konfigurowanie projektów SSIS przy użyciu wartości odsyłaczy	355
	Wykonywanie pakietów z wykorzystaniem środowisk SSIS	356
	Zarządzanie środowiskami SSIS i zmiennymi środowiskowymi	356
	Podsumowanie	358
13	Magia PowerShell w SSIS	359
	Czym jest PowerShell	359
	PowerShell i SQL Server	360

Zarządzanie SSIS przy użyciu PowerShell	363
Model zarządzania obiektowego SSIS	363
PowerShell i model SSIS Management Object	364
Wykorzystanie T-SQL do zarządzania SSIS z poziomu PowerShell	369
Zalety wykorzystania PowerShell do zarządzania SSIS	370
Podsumowanie	371
14 Raporty SSIS	373
Wprowadzenie do raportów SSIS	373
Przygotowywanie danych	375
Monitorowanie wykonywania pakietów SSIS	376
Integration Services Dashboard	376
Raport All Executions	378
Raporty All Validations i All Operations	379
Wykorzystanie raportów SSIS do rozwiązywania problemów z wykonaniami pakietów SSIS	380
Korzystanie z raportu Execution Performance do identyfikowania trendów wydajności	385
Podsumowanie	388

Część IV: Pogłębione informacje

15 Motor SSIS	391
Motor przepływu sterowania	391
Przegląd	391
Ładowanie	392
Aplikacja parametrów	394
Weryfikacja	394
Wykonanie	396
Motor przepływu danych	404
Przegląd	405
Sterowanie wykonaniem	409
Backpressure	417
Dostrajanie motoru przepływu danych	421
Podsumowanie	423
16 Katalog SSIS	425
Pogłębione informacje na temat katalogu SSIS	425
Tworzenie katalogu SSIS	425
Jednostki wdrożenia w katalogu SSIS	426
Co znajdziemy wewnątrz bazy danych SSISDB?	427
Rozruch instancji SQL Server	430
Katalog SSIS i poziomy rejestrowania	432

Cykl życiowy wykonywania pakietu SSIS	434
Zatrzymywanie wykonania pakietu SSIS	436
Korzystanie z dziennika zdarzeń Application systemu Windows	436
Konserwacja katalogu SSIS i zadania SQL Server Agent	438
Wykonywanie kopii zapasowej i przywracanie katalogu SSIS	440
Tworzenie kopii zapasowej SSISDB	441
Przywracanie bazy danych SSISDB	442
Podsumowanie	444
17 Zabezpieczenia SSIS	445
Ochrona pakietu	445
Kontrola dostępu do pakietu	445
Szyfrowanie pakietów	449
Wrażliwe zmienne i parametry	450
Podpisywanie pakietów	452
Zabezpieczenia katalogu SSIS	453
Przegląd funkcji zabezpieczeń	453
Zarządzanie uprawnieniami	457
Wyzwalacz DDL	465
Wykorzystanie SQL Server Agent	465
Wymagania	466
Tworzenie poświadczeń	466
Tworzenie kont proxy	468
Tworzenie zadania SQL Server Agent	470
Podsumowanie	472
18 Dzienniki SSIS	473
Konfigurowanie opcji rejestrowania	473
Wybieranie kontenerów	473
Wybieranie zdarzeń	476
Dodawanie dostawców dziennika	478
Dostawcy dziennika	481
Pliki tekstowe	481
SQL Server	481
SQL Server Profiler	482
Dzienniki zdarzeń Windows	483
Pliki XML	483
Rejestrowanie w katalogu SSIS	485
Poziomy rejestrowania	485
Dzienniki zdarzeń	487
Informacje kontekstu zdarzenia	487
Zaawansowane zagadnienia rejestrowania	488
Dostosowywanie pól wpisów dziennika	488

Rejestrowanie przy korzystaniu z narzędzia DTEXEC	489
Projektowanie niestandardowego dostawcy dziennika	490
Podsumowanie	491
19 Automatykacja SSIS	493
Wprowadzenie do automatyzacji SSIS	493
Programowe generowanie pakietów SSIS	493
Wykonywanie pakietów sterowane metadanymi	494
Dynamiczne generowanie pakietów	495
Obsługa zdarzeń czasu projektowania	496
Przykłady	498
Wykonanie oparte na metadanych	507
Niestandardowy program uruchamiający pakiet	509
Wykorzystanie PowerShell	513
Korzystanie z PowerShell wraz z SQL Server Agent	516
Alternatywne rozwiązania i przykłady	519
Przykłady w witrynie Codeplex	519
Rozwiązania innych firm	520
Podsumowanie	522

Część V: Rozwiązywanie problemów

20 Rozwiązywanie problemów z awariami pakietów SSIS	525
Wprowadzenie do rozwiązywania problemów	525
Przygotowanie danych	527
Nowe funkcje rozwiązywania problemów z wykonywaniem pakietów	528
Trzy kluczowe kroki podczas rozwiązywania problemów z wykonywaniem pakietów SSIS	530
Ścieżka wykonania	534
Znajdowanie początkowego źródła problemu	534
Rozwiązywanie problemów związanych z wykonaniami pakietów podrzędnych	538
Zdarzenia <i>DiagnosticEx</i>	540
Zadanie <i>Execute Package</i> i ścieżki wykonania	541
Rozwiązywanie problemów z wykonywaniem pakietów za pośrednictwem SQL Server Agent	543
Identyfikowanie wykonań pakietów SSIS realizowanych przez SQL Server Agent	546
Korzystanie z tabel historii SQL Server Agent do identyfikowania kroków zadań SSIS zakończonych niepowodzeniem	547
Podsumowanie	548
21 Najlepsze rozwiązania dotyczące wydajności SSIS	549
Tworzenie strategii poprawy wydajności	549

Technika OWAL	550
Mierzenie sprawności przetwarzania SSIS	552
Mierzenie wydajności systemu	552
Mierzenie wydajności zadań przepływu danych	556
Projektowanie z myślą o wydajności	562
Paralelizowanie projektu	562
Korzystanie z technik optymalizacyjnych SQL Server	567
Masowe ładowanie danych	569
Utrzymywanie operacji SSIS w pamięci	572
Optymalizacja pamięci podręcznej wyszukiwania	573
Optymalizowanie infrastruktury SSIS	577
Podsumowanie	580
22 Rozwiązywanie problemów dotyczących wydajności	581
Profilowanie wydajności	581
Monitorowanie wydajności	582
Przygotowanie danych	583
Poznanie wydajności pakietu SSIS	584
Czas trwania wykonania pakietu SSIS	584
Czas spędzony w każdym zadaniu wchodzącym w skład pakietu	585
Czasy trwania poszczególnych faz komponentu <i>Data Flow</i>	586
Uptywający czas dla faz komponentu <i>Data Flow</i> (czas aktywny kontra czas całkowity)	586
Monitorowanie wydajności wykonywania pakietów SSIS	588
Liczniki wydajności dotyczące wykonania	590
Interaktywna analiza danych wydajności	592
Podsumowanie	599
23 Rozwiązywanie problemów dotyczących danych	601
Rozwiązywanie problemów w środowisku projektowym	601
Zliczanie wierszy	601
Podgląd danych	603
Dane w wyjściu błędów	605
Pułapki i okna debugowania	606
Rozwiązywanie problemów w środowisku produkcyjnym	606
Dane statystyczne wykonania	606
Rozgałęźniki danych	609
Zrzuty błędów	614
Podsumowanie	616