

Planowanie procesów ETL z wykorzystaniem MS SQL Server 2008 Integration Services

Paweł Potasiński



KAPITAŁ LUDZKI
NARODOWA STRATEGIA SPÓJNOŚCI



WARSZAWSKA
WYŻSZA SZKOŁA
INFORMATYKI

UNIA EUROPEJSKA
EUROPEJSKI
FUNDUSZ SPOŁECZNY



Prelegent.ToString();



- Programista w Asseco Business Solutions S.A.
- Microsoft Certified Trainer od 2004 roku
- Microsoft Most Valuable Professional od 2008 roku
- Autor dwóch książek i licznych artykułów o SQL Server
- Założyciel Polskiej Grupy Użytkowników SQL Server
 - <http://www.plssug.org.pl>
- Blog: <http://zine.net.pl/blogs/sqlgeek>



KAPITAŁ LUDZKI
NARODOWA STRATEGIA SPÓJNOŚCI

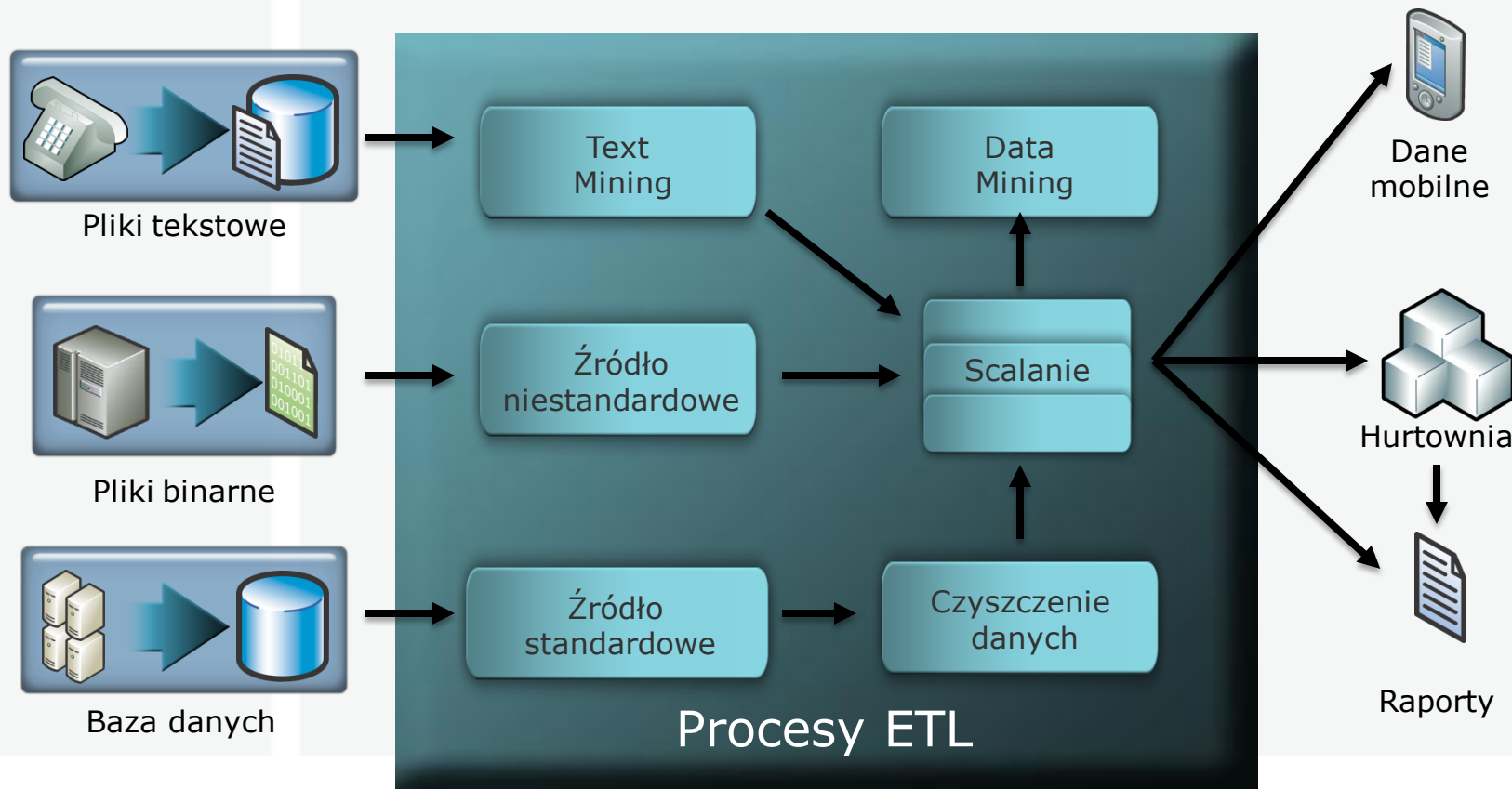


WARSZAWSKA
WYŻSZA SZKOŁA
INFORMATYKI

UNIA EUROPEJSKA
EUROPEJSKI
FUNDUSZ SPOŁECZNY



Procesy ETL



Problemy i rozwiązania

Problemy

- Heterogeniczne źródła danych
- Różne formaty danych
- Różne ziarnistości danych
- Jakość danych
- Okna czasowe dla procesów ETL
- Aktualizacje danych

Rozwiązania

- ← Silne narzędzia ETL
- ← Staging
- ← Staging
- ← Profilowanie i czyszczenie danych
- ← Synchronizacja procesów ETL
- ← Slowly Changing Dimension



KAPITAŁ LUDZKI
NARODOWA STRATEGIA SPÓJNOŚCI



WARSZAWSKA
WYŻSZA SZKOŁA
INFORMATYKI

UNIA EUROPEJSKA
EUROPEJSKI
FUNDUSZ SPOŁECZNY



Integration Services jako platforma ETL

- Obsługa praktycznie dowolnych źródeł danych
- Procesy ETL zapisane w paczkach
- Control Flow i Data Flow
- Spory wybór funkcjonalności (task'ów)
- Zmienne
- Konfiguracje
- Logowanie
- Obsługa zdarzeń
- Repozytorium paczek w bazie systemowej msdb
- Specjalna aplikacja umożliwia łatwe wdrażanie
- Integracja z usługą SQL Server Agent
- Checkpointy
- Obsługa transakcji
- Pełna rozszerzalność
- Skrypty .NET



DEMO:

WYBRANE ROZWIĄZANIA DLA PROCESÓW ETL

- 1. Data Profiling Task**
- 2. Fuzzy Lookup + Fuzzy Grouping**
- 3. Slowly Changing Dimension**



KAPITAŁ LUDZKI
NARODOWA STRATEGIA SPÓJNOŚCI



**WARSZAWSKA
WYŻSZA SZKOŁA
INFORMATYKI**

**UNIA EUROPEJSKA
EUROPEJSKI
FUNDUSZ SPOŁECZNY**



Zasoby

- <http://blogs.msdn.com/mattm/>
- <http://www.sqlis.com>
- <http://blogs.conchango.com/jamiethomson>
- <http://www.kimballgroup.com/>
- <http://www.codeplex.com/kimballscd/>
- <http://www.sqlbi.com>



KAPITAŁ LUDZKI
NARODOWA STRATEGIA SPÓJNOŚCI



WARSZAWSKA
WYŻSZA SZKOŁA
INFORMATYKI

UNIA EUROPEJSKA
EUROPEJSKI
FUNDUSZ SPOŁECZNY



Dziękuję za uwagę

PYTANIA?

pawel.potasinski@sqlpass.org



KAPITAŁ LUDZKI
NARODOWA STRATEGIA SPÓJNOŚCI



WARSZAWSKA
WYŻSZA SZKOŁA
INFORMATYKI

UNIA EUROPEJSKA
EUROPEJSKI
FUNDUSZ SPOŁECZNY

