

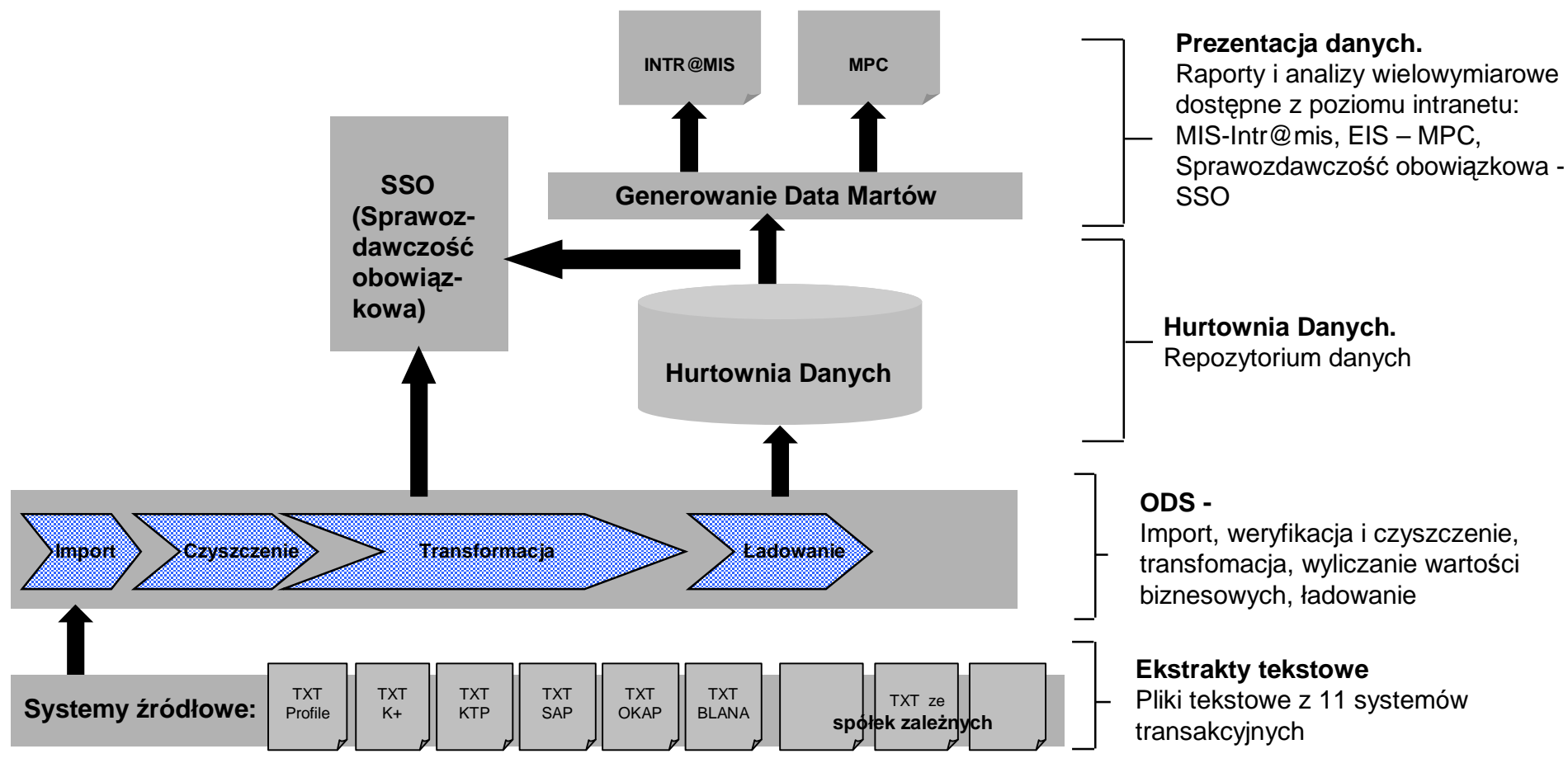
SPD Server w akcji

Zintegrowana platforma sprawozdawcza **aRaMIS**
w Banku BPH SA

Agenda

- ☪ Zintegrowana platforma sprawozdawcza w Banku BPH SA
- ☪ Hurtownia danych – miejsce dla SPD Server.
- ☪ Funkcjonalność repozytorium hurtowni.
- ☪ Korzyści
- ☪ Podsumowanie

Zintegrowana platforma sprawozdawcza w Banku BPH SA



Hurtownia Danych miejsce dla SPD Server

Rola Hurtowni Danych w Banku BPH SA

- § Integracja danych z wielu systemów informatycznych.
- § Unifikacja danych danych.
- § Wzbogacanie danych.
- § Repozytorium dla zintegrowanej platformy raportowej aRaMIS.

Hurtownia Danych miejsce dla SPD Server

- 1 Hurtownia Danych w Banku BPH SA zbudowana jest w oparciu o relacyjny model gwiazdy.
W modelu przechowywane są:
 - ☉ Fakty
 - ☉ Wymiary wersjonowane w czasie
 - ☉ Wymiary wielozmienne.
- 2 Zasilanie odbywa się w cyklach dziennym, dekadowym i miesięcznym.
- 3 Łączny wolumen przechowywanych danych – 8,5TB.
- 4 Liczba rekordów w tabelach faktów od 100 do 400 milionów rekordów.

Repozytorium hurtowni

Główne tabele hurtowni danych zostały umieszczone w repozytorium opartym o klastry SPD Server.

Powstał zestaw makr, który pozwala na uproszczenie wykonywania najczęstszych czynności:

- § Dodanie tabeli składowej do klastra (zasilanie).
- § Wyłączenie tabeli z klastra.
- § Zmiana struktury klastra.

Repozytorium hurtowni

Przechowywanie danych

Fakty:

§ Dane podzielone na partycje zgodnie z czasem zasilania.

Wymiary wielozmienne

§ Pojedyncza partycja odświeżana całościowo.

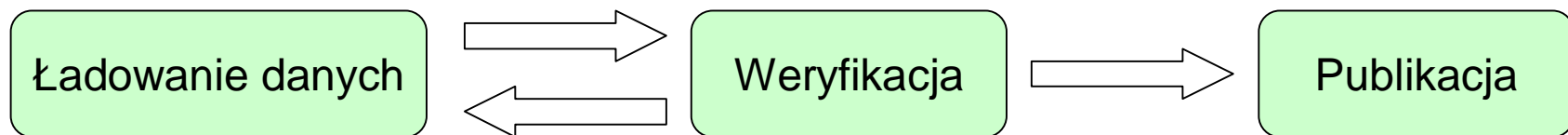
Wymiary wersjonowane w czasie

§ Partycja przechowująca wersje obowiązujące.

§ Partycje przechowujące wersje historyczne podzielone zgodnie z czasem zasilania.

Repozytorium hurtowni

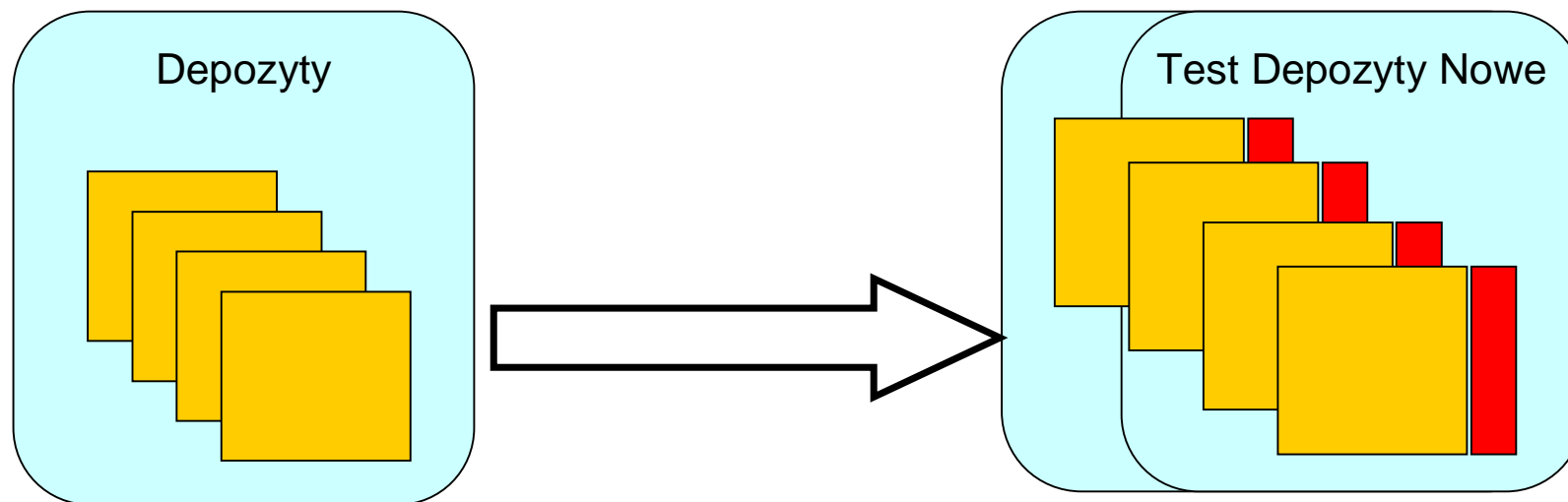
Zasilanie hurtowni.



- Przygotowanie danych do stanu klastra
- Weryfikacja struktury
- Weryfikacja konstrukcyjnej
- Weryfikacja danych
- Indeksowanie
- Sortowanie

Repozytorium hurtowni

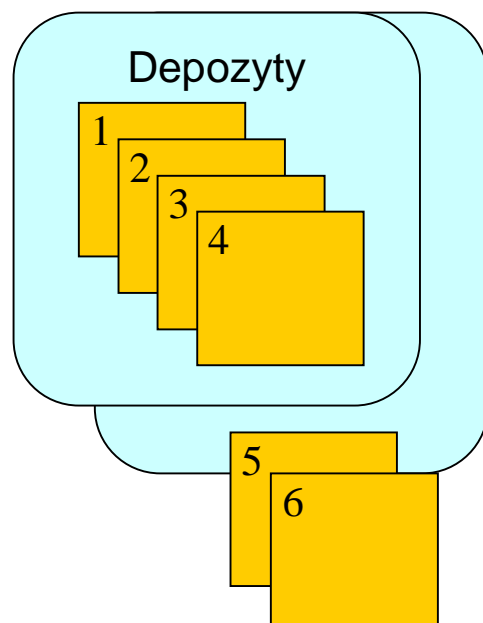
Zmiana struktury klastra.



- Realizacja testów jednostkowych i testów integracyjnych

Repozytorium hurtowni

Zmiana zakresu danych w klastrze.



1. Załadowanie nowych tabel wchodzących w skład klastra.
2. Wybór tabel usuwanych z klastra.
3. Deklaracja klastra wg nowej definicji.

Repozytorium hurtowni

Makra administracyjne

Oprogramowano następujące czynności:

- § Ładowanie i usuwanie danych
- § Wykonywanie kopii zapasowych oraz odzyskiwanie poszczególnych składowych i całych klastrów
- § Publikacja nowych stanów klastrów
- § Zarządzanie i zmiana wersji struktury klastrów
- § Korekta danych przechowywanych w klastrach

Wdrożenie

Modyfikacje w istniejących już elementach hurtowni:

1. Bloki przetwarzania zapisujące dane do repozytorium
 - § Rozbicie dotychczasowych bloków na pary ładowanie – publikacja
 - § Dzięki makrom administracyjnym tworzenie nowych wersji bloków zostało znacznie uproszczone
2. Bloki przetwarzania czytające z repozytorium
 - § Zmiana definicji bibliotek repozytorium

Repozytorium hurtowni

- ☺ W obecnej chwili w klastrach SPD Server przechowywane są dane obejmujące ostatnie dwa lata.
- ☺ W repozytorium umieszczono główne tabele faktów oraz kluczowe wymiary
- ☺ Wolumen przechowywanych danych to 1,5 TB i ciągle rośnie

Korzyści

W porównaniu do dotychczasowej realizacji repozytorium hurtowni danych (opartej o engine BASE) osiągnięto:

- § Skrócenie czasu ładowania o ponad 30%
- § Skrócenie przerwy administracyjnej w dostępie do danych hurtowni
- § Dwuetapowość zasilania – oddzielenie etapu ładowania danych od ich udostępnienia
- § Zmniejszenie przestrzeni potrzebnej do wykonywania kopii zapasowej oraz czasu jej wykonania
- § Elastyczność przy zmianie struktury tabel i zmianie zasięgu czasowego przechowywanych danych

Korzyści

Dodatkowo, w porównaniu do dotychczasowej realizacji repozytorium hurtowni danych (opartej o engine BASE), osiągnięto skrócenie czasu analiz i generowania raportów dzięki:

- § Partycjonowaniu danych – zrównoleglenie odczytu
- § Wykorzystaniu struktury klastrów – zapytania zawężające dane wg dat odczytują tylko niezbędne tabele wchodzące w skład klastra
- § Wykorzystaniu dodatkowych cech SPD Server:
 - § wielowątkowość
 - § Index scan
 - § Indeksy hybrydowe

Problemy

- § Proces automatycznego backupu mocno obciąża zasoby serwera.
- § Struktura składowych klastra musi być identyczna w najdrobniejszych szczegółach (np. labelle kolumn).
- § Złożona obsługa zmian struktury.

Podsumowanie

- ☺ Dwukrotnie krótszy czas zasilania hurtowni
- ☺ Czterokrotnie krótszy czas wykonywania analiz na danych przechowywanych w klastrach SPD Server
- ☺ Rozmiar backupu miesięcznego spadł z 1,5 TB do ok. 30 GB
- ☺ Czas ograniczonej dostępności hurtowni spadł z 26 do 2h