

PROJEKT I IMPLEMENTACJA
HURTOWNI DANYCH W
JEDNOSTCE SŁUŻBY ZDROWIA

ARKADIUSZ WILAMOWSKI

Plan prezentacji

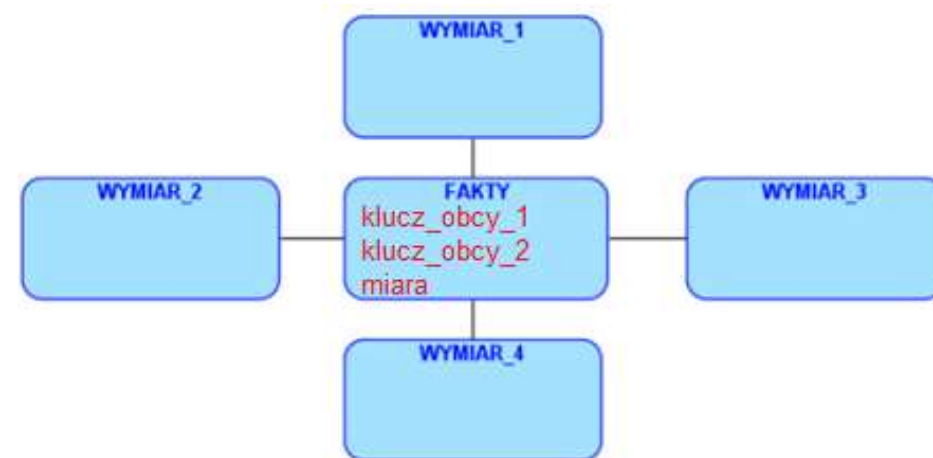
- Wstęp i cel pracy
- Omówienie głównych pojęć
- Zakres prac – szersze omówienie

Wstęp i cel pracy

- Wstęp
- Wybór tematu
- Cel pracy

Omówienie głównych pojęć

- ❑ Różnica między OLTP, a OLAP?
- ❑ Hurtownia danych
 - Definicja i wspólne cechy (nieulotność, zintegrowanie, zróżnicowanie czasowe i tematyczne)
 - Wizja Kimbal vs Inmon
 - Argumenty „za” wyborem
- ❑ Model gwiazdy
- ❑ ETL
- ❑ HIS



Zakres prac

W kolejnych slajdach przedstawię:

1. Analiza wymagań (system źródłowy)
2. Wybór architektury
3. Wybór narzędzi
4. Projekt testowej bazy HIS
5. Projekt bazy hurtowni danych
 - a) obszaru przejściowego
 - b) modelu obszaru hurtowni stricte
6. Przepływ danych HD, ETL i jego struktury pomocnicze
7. Przygotowanie serwerów baz danych
 - a) Konfiguracja sieciowa maszyn wirtualnych
 - b) Nawiązanie połączenia z bazą źródłową (TNS, DBLINK)
 - c) Wygenerowanie struktury bazy HIS
 - d) Wygenerowanie struktury HD
 - e) Wygenerowanie danych testowych
 - f) Przygotowanie przykładowych zapytań
8. Zbudowanie aplikacji przy użyciu narzędzia typu RAD

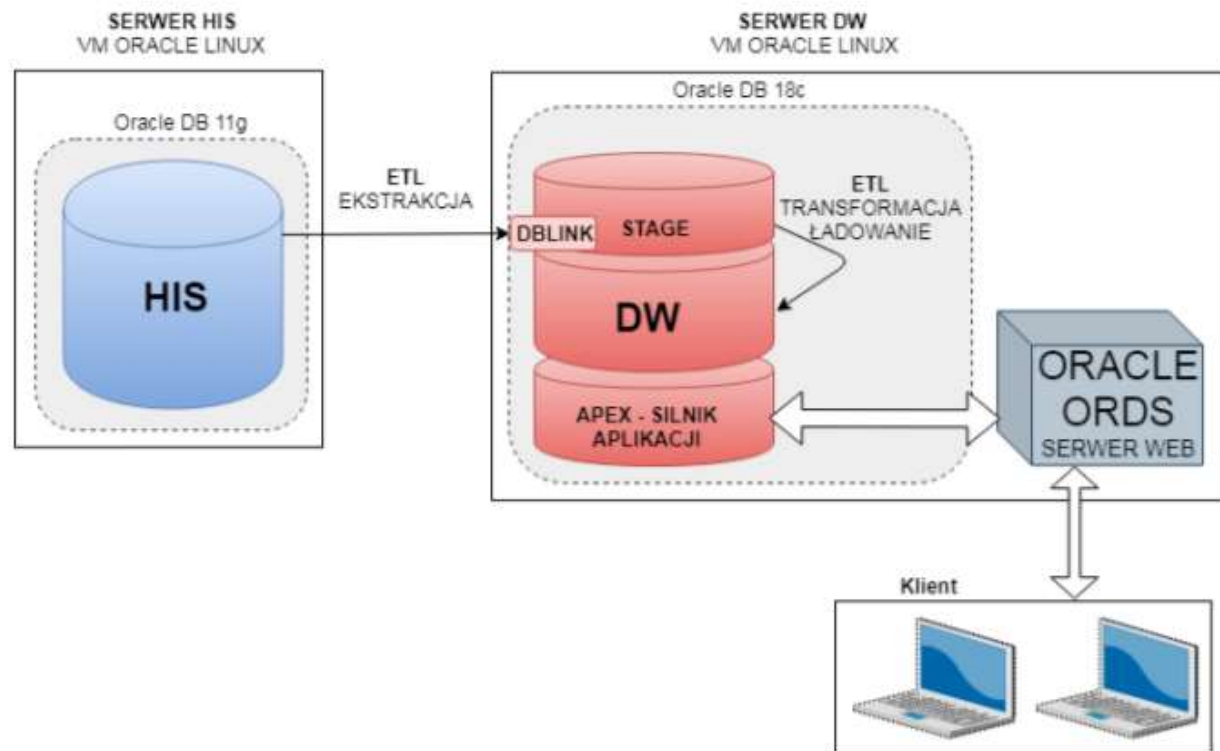
Analiza wymagań

- Opracowanie analizy procedur medycznych, rozliczeń Zespołu Poradni, analiza geograficzna
- System źródłowy to relacyjna baza HIS z Oracle 11g
- Szybka budowa interfejsu użytkownika i nieutrudniająca administracja (dostęp przez przeglądarkę)

Architektura

- Klient-serwer
- 3-warstwowa

Rysunek 4-2 Architektura systemu hurtownii danych



Wybór narzędzi – omówienie stosu technologii

Oracle Linux

RDBMS v 12c (SQL Dev, Data Modeler)

SQL

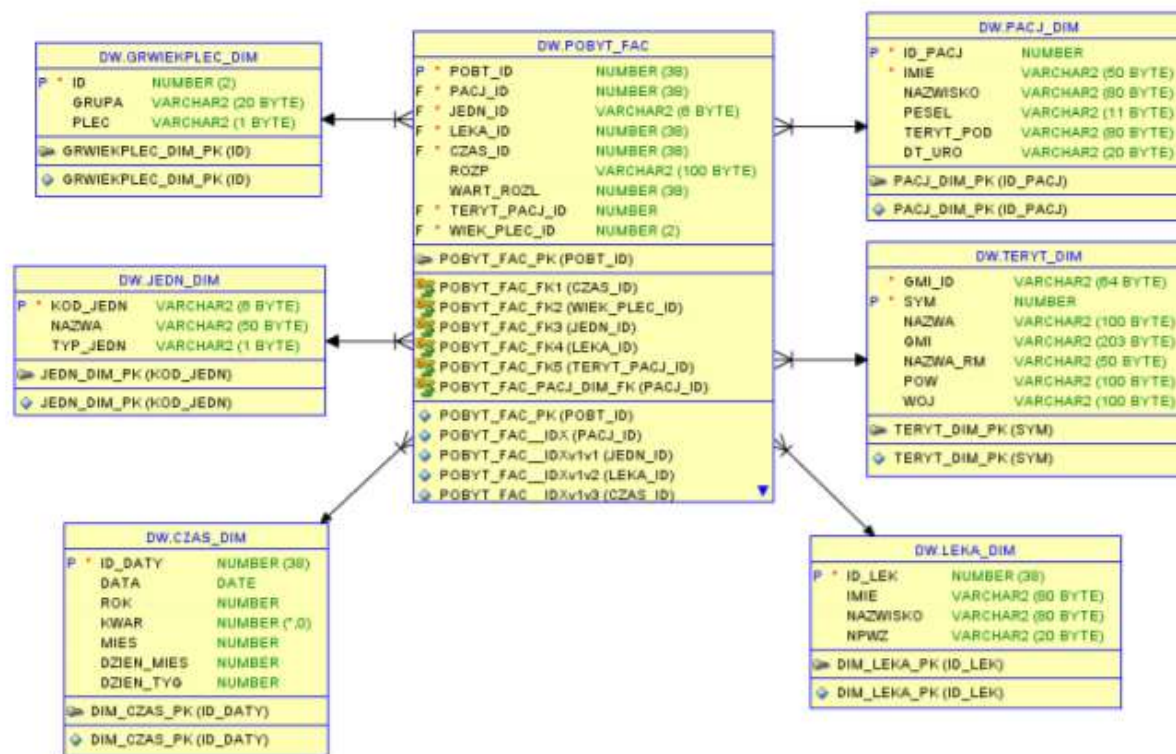
PL\SQL

Oracle Apex (RAD)

Projekt hurtowni danych

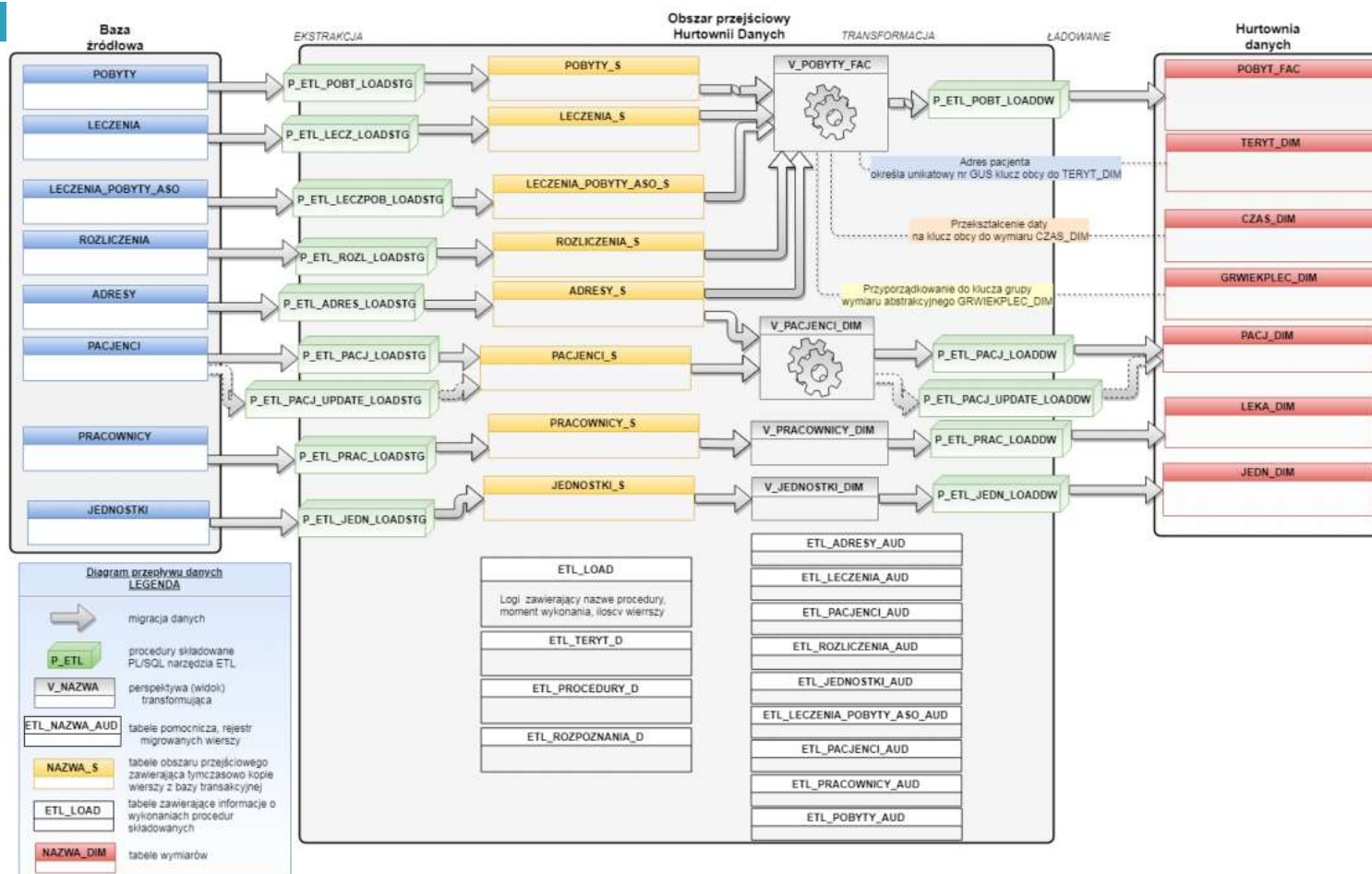
- Model obszaru przejściowego
- Model HD (star)
- Określenie właściwości wymiarów
 - rodzaju
 - granulacji
 - zmienności wymiaru

Rysunek 4-5 Hurtownia danych. Implementacja modelu danych gwiazdy



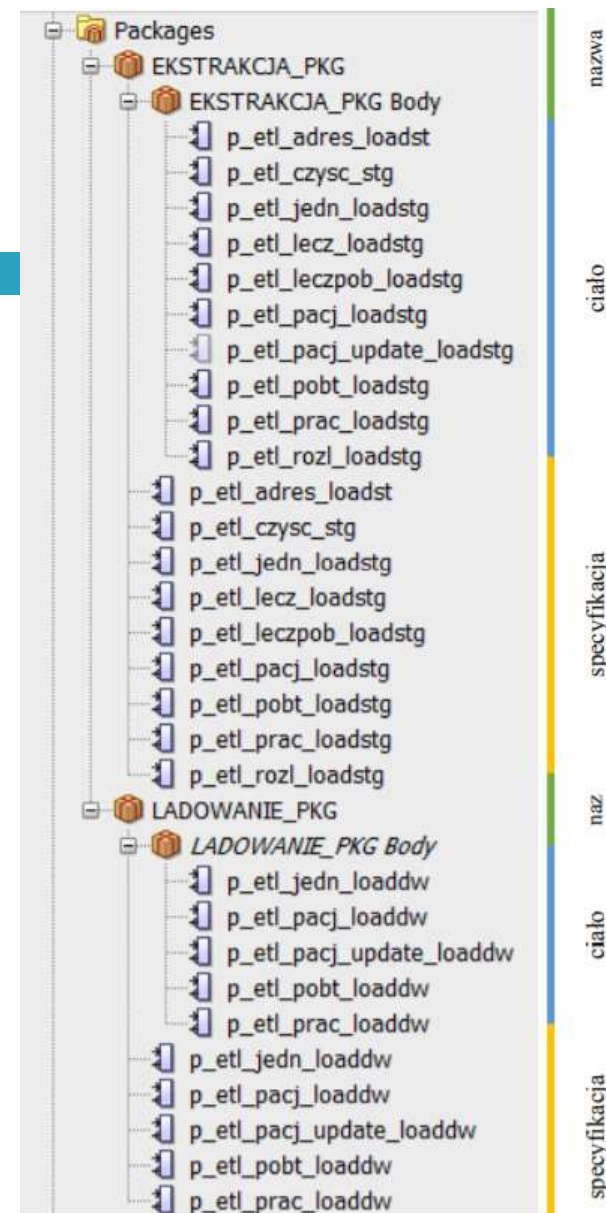
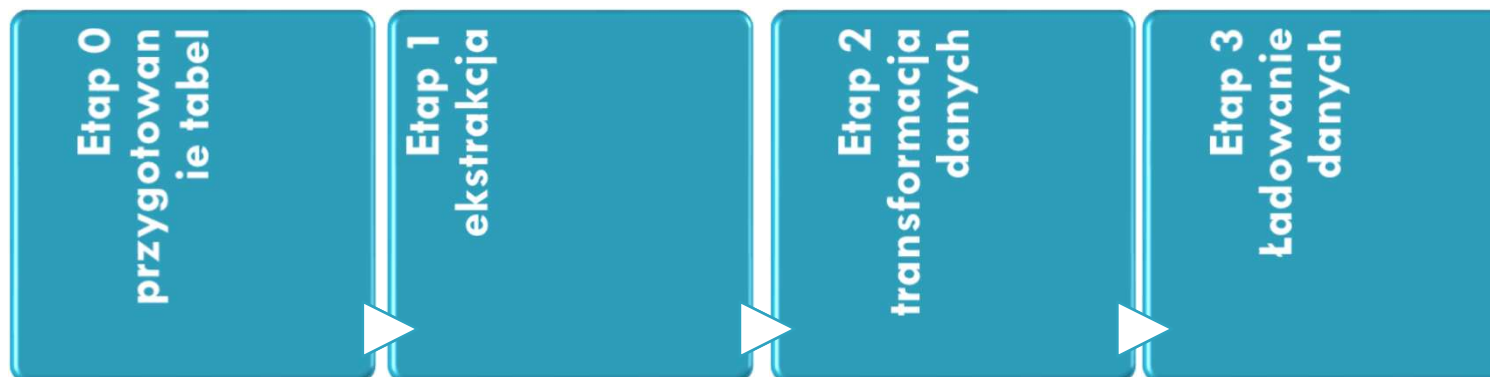
Przepływ danych w HD cz1

- Procedury PL/SQL
- ETL
 - ekstrakcja (DBlink i CTAS)
 - oczyszczanie i transformacja (view)
 - ładowanie
- Struktury pomocnicze



Przepływ danych w HD cz2

- Optymalizacja - paczki/pakiety oraz przetwarzanie wsadowe
- Scheduler (tzw. Job)
- P_START_ETL



Serwery baz danych

- Konfiguracja maszyn wirtualnych
- TNS
- DBLINK

Budowa aplikacji przy użyciu narzędzia typu RAD

- Instalacja i wybór serwera aplikacji
- Konta użytkowników (admin, analityk)
- Budowa interfejsu użytkownika przy użyciu gotowych komponentów
- Użycie zapytań analitycznych

Moduł Analizy Danych

Wymiar ↑	Zasilenie
Jednostki	19/06/16
Pacjenci aktualizacja	19/06/16
Pacjenci	19/06/16
Pobyty	19/06/16
Pracownicy	19/06/16

1 - 5

Przykładowe zapytania analityczne

- Funkcja `DENSE_RANK` pozwala na stworzenie rankingu, oblicza ona kolejne numery wierszy wg. wartości kolumny.

```
SELECT
jedn_id, rozp, COUNT(*) AS ilosc, DENSE_RANK() OVER (PARTITION BY
jedn_id ORDER BY COUNT(*) DESC) AS ranking
FROM pobyt_fac
WHERE jedn_id=:p11_wyb_rozp
GROUP BY jedn_id, rozp
```

Wybierz poradnię
PORCHO

Jedn Id ↑	Rozp	Ilosc	Ranking
PORCHO	S93	708	1
PORCHO	K80	394	2
PORCHO	S52	392	3
PORCHO	S62	390	4
PORCHO	S68	338	5
PORCHO	D22	279	6
PORCHO	K35	226	7
PORCHO	L03	221	8
PORCHO	S61	200	9
PORCHO	S92	199	10
PORCHO	K85	184	11
PORCHO	S01	175	12
PORCHO	B84	158	13
PORCHO	E11.5	147	14

Przykładowe zapytania analityczne

```
SELECT  
C.rok,F.rozp,P.plec ||' '|| P.grupa AS grupa, COUNT(*)  
FROM pobyt_fac F, czas_dim C, grwiekplec_dim P  
WHERE  
F.czas_id=C.id_daty AND  
F.wiek_plec_id=P.ID  
AND rozp =:pl2_wybor_rozp  
GROUP BY (rok,F.rozp,P.plec ||' '|| P.grupa)  
ORDER BY rok DESC
```

Rozpoznania a grupy wiekowe

