Документация решения: Аналитика просмотров видео

# 1. Контекст OLTP базы данных (video\_content\_db)

Описание: OLTP база данных (video\_content\_db) используется для хранения операционных данных о просмотрах видео в реальном времени.

Хранимые данные:  
Таблица Views:  
- view\_id (integer): Уникальный идентификатор просмотра.  
- user\_id (integer): Идентификатор пользователя.  
- video\_id (integer): Идентификатор видео.  
- view\_date (timestamp): Дата и время просмотра.  
Цель: Хранение сырых транзакционных данных для последующей обработки.

# 2. Контекст OLAP базы данных (olap\_video\_analytics)

Описание: OLAP база данных (olap\_video\_analytics) предназначена для аналитической обработки данных о просмотрах видео.

Аналитические вопросы:  
- Как распределены просмотры по пользователям?  
- Как меняется активность просмотров по датам?  
- Какие видео наиболее популярны?

Хранимые данные:  
Таблица dim\_time:  
- time\_id (integer): Уникальный идентификатор времени.  
- date (date): Дата.  
- month (integer): Месяц.  
- year (integer): Год.  
  
Таблица dim\_users:  
- user\_id (integer): Идентификатор пользователя.  
- username (character varying): Имя пользователя.  
  
Таблица dim\_videos:  
- video\_id (integer): Идентификатор видео.  
- title (character varying): Название видео.  
  
Таблица fact\_video\_views:  
- time\_id (integer): Ссылка на dim\_time.  
- user\_id (integer): Ссылка на dim\_users.  
- video\_id (integer): Ссылка на dim\_videos.  
- total\_views (numeric): Общее количество просмотров.

# 3. Общее описание схем, таблиц, ключей, ограничений и связей

Схема OLTP:  
- Таблица Views является основной, содержит транзакционные данные.  
- Отсутствуют явные ссылки на Users и Videos (предполагается, что они могут быть добавлены).

Схема OLAP:  
- Модель звезды с фактовой таблицей fact\_video\_views и размерными таблицами dim\_time, dim\_users, dim\_videos.  
- Связь fact\_video\_views.time\_id → dim\_time.time\_id.  
- Связь fact\_video\_views.user\_id → dim\_users.user\_id.  
- Связь fact\_video\_views.video\_id → dim\_videos.video\_id.  
- Ограничения: Первичные ключи (time\_id, user\_id, video\_id) обеспечивают уникальность записей.  
- Отношения: Иерархическая структура для агрегации данных (время → пользователи → видео).

# 4. Инструкции по запуску скриптов

Подключение к базам:  
- OLTP: video\_content\_db, пользователь postgres, пароль qwerty, хост 127.0.0.1.  
- OLAP: olap\_video\_analytics, пользователь postgres, пароль qwerty, хост 127.0.0.1.

Скрипты создания таблиц:

CREATE TABLE dim\_time (  
 time\_id SERIAL PRIMARY KEY,  
 date DATE,  
 month INTEGER,  
 year INTEGER  
);  
CREATE TABLE dim\_users (  
 user\_id INTEGER PRIMARY KEY,  
 username VARCHAR(50)  
);  
CREATE TABLE dim\_videos (  
 video\_id INTEGER PRIMARY KEY,  
 title VARCHAR(200)  
);  
CREATE TABLE fact\_video\_views (  
 time\_id INTEGER,  
 user\_id INTEGER,  
 video\_id INTEGER,  
 total\_views NUMERIC,  
 PRIMARY KEY (time\_id, user\_id, video\_id),  
 FOREIGN KEY (time\_id) REFERENCES dim\_time(time\_id),  
 FOREIGN KEY (user\_id) REFERENCES dim\_users(user\_id),  
 FOREIGN KEY (video\_id) REFERENCES dim\_videos(video\_id)  
);

ETL процесс:

DO $$  
BEGIN  
 PERFORM dblink\_connect('myconn', 'host=127.0.0.1 user=postgres password=qwerty dbname=video\_content\_db');  
EXCEPTION WHEN OTHERS THEN  
 RAISE NOTICE 'Connection failed: %', SQLERRM;  
 RETURN;  
END $$;  
  
DO $$  
BEGIN  
 -- Загрузка dim\_time  
 INSERT INTO dim\_time (date, month, year)  
 SELECT DISTINCT DATE(view\_date) AS date, EXTRACT(MONTH FROM view\_date) AS month, EXTRACT(YEAR FROM view\_date) AS year  
 FROM dblink('myconn', 'SELECT view\_date FROM Views')   
 AS v(view\_date TIMESTAMP)  
 ON CONFLICT (date) DO NOTHING;  
  
 -- Загрузка dim\_users  
 INSERT INTO dim\_users (user\_id, username)  
 SELECT DISTINCT v.user\_id, 'user' || v.user\_id::text AS username  
 FROM dblink('myconn', 'SELECT user\_id FROM Views')   
 AS v(user\_id INTEGER)  
 ON CONFLICT (user\_id) DO NOTHING;  
  
 -- Загрузка dim\_videos  
 INSERT INTO dim\_videos (video\_id, title)  
 SELECT DISTINCT v.video\_id, 'video' || v.video\_id::text AS title  
 FROM dblink('myconn', 'SELECT video\_id FROM Views')   
 AS v(video\_id INTEGER)  
 ON CONFLICT (video\_id) DO NOTHING;  
  
 -- Загрузка fact\_video\_views  
 INSERT INTO fact\_video\_views (time\_id, user\_id, video\_id, total\_views)  
 SELECT t.time\_id, v.user\_id, v.video\_id, COUNT(\*) AS total\_views  
 FROM dblink('myconn', 'SELECT user\_id, video\_id, view\_date FROM Views')   
 AS v(user\_id INTEGER, video\_id INTEGER, view\_date TIMESTAMP)  
 JOIN dim\_time t ON DATE(v.view\_date) = t.date  
 GROUP BY t.time\_id, v.user\_id, v.video\_id  
 ON CONFLICT ON CONSTRAINT fact\_video\_views\_pkey DO UPDATE  
 SET total\_views = EXCLUDED.total\_views  
 WHERE fact\_video\_views.total\_views != EXCLUDED.total\_views;  
  
 -- Обновление метки времени  
 PERFORM dblink\_exec('myconn', 'UPDATE etl\_log SET last\_load\_timestamp = CURRENT\_TIMESTAMP WHERE table\_name = ''Views''');  
END $$;  
  
SELECT dblink\_disconnect('myconn');

Проверка после выполнения ETL:

SELECT COUNT(\*) FROM dim\_time;  
SELECT COUNT(\*) FROM dim\_users;  
SELECT COUNT(\*) FROM dim\_videos;  
SELECT COUNT(\*) FROM fact\_video\_views;

# 5. Power BI отчёт

Описание: Отчёт создан в Power BI для визуализации аналитических данных.

Визуализации:  
- Заголовок: "Отчёт по просмотрам видео (OLAP-аналитика)".  
- Слайсеры: Фильтр по дате (dim\_time.date), Фильтр по пользователям (dim\_users.username).  
- Визуальные компоненты:  
 - Линейный график: Динамика просмотров по годам (dim\_time.year vs SUM(fact\_video\_views.total\_views)).  
 - Круговая диаграмма: Распределение просмотров по видео (dim\_videos.title vs SUM(fact\_video\_views.total\_views)).  
 - Столбчатая диаграмма: Просмотры по пользователям (dim\_users.username vs SUM(fact\_video\_views.total\_views)).  
Инсайты:  
- Основная активность в 2025 году (5 просмотров).  
- Популярность видео варьируется (например, video1 имеет 2 просмотра).  
- Распределение просмотров по пользователям показывает лидерство user1.