



Опыт разработки и внедрения решений репликации данных из Oracle в Greenplum

Рябов Алексей, Павлов Дмитрий

Tinkoff.ru

Эволюция механизмов загрузки данных

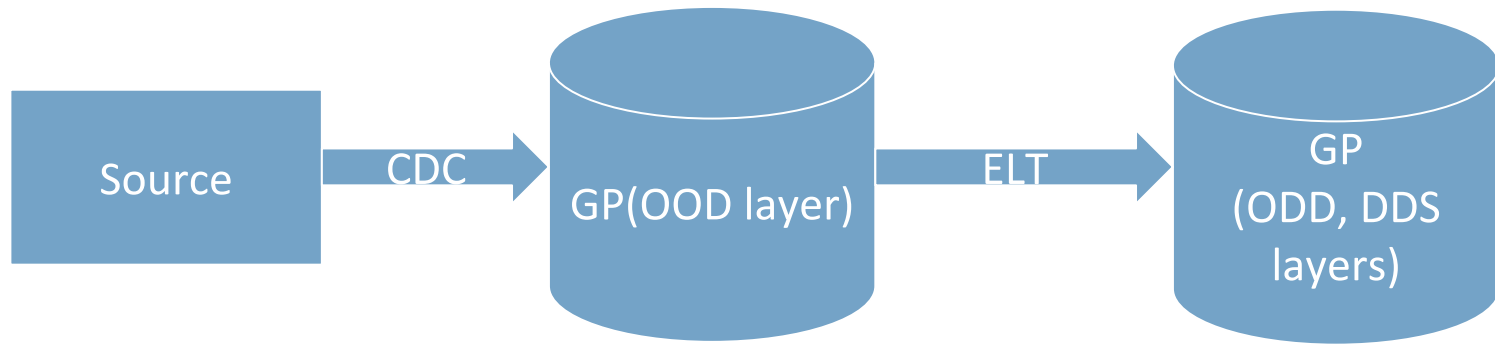


- До 2012 года: чтение прямо из источников
- 2012 – 2014: репликация источников прямо в GreenPlum
- 2014 – н.в: использование ODS/O2G

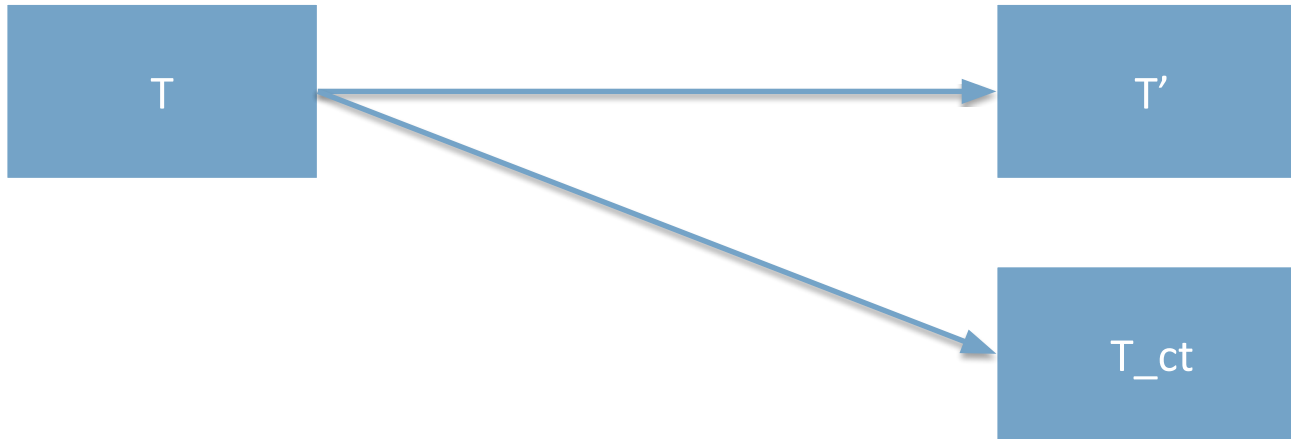
Этап 1



Этап 2. Загрузка данных



Этап 2. Структура таблиц в GP



- T' – реплика исходной таблицы источника в GP
- T_ct – change-table, таблица в GP с историей изменений таблицы T в источнике
- Таблицы T и T_ct имеют дополнительные служебные поля
- Таблицы T_ct секционированы по дате операции

Этап 2. Применение изменений в GP



**Основная проблема – огромное количество операций на источниках.
Вариант решения – «схлопывание по ключу»**

Применение изменений в таблицы-реплики. Пример.

Лог изменений

Op. number	key	f1	f2	f3
1	1	f11		
2	1		f2	
3	1	f111	f22	
4	1			
5	1			
6	1	f1111		
7	1			f333

Порядок применения изменений в реплику.

- данные попадают в т.н. net-таблицу. (внутренняя таблица, для загрузки используются external table)
- Данные разбиваются на батчи и «схлопываются» последовательно по ключу. В у казанном примере результат «схлопывания» будет следующий:

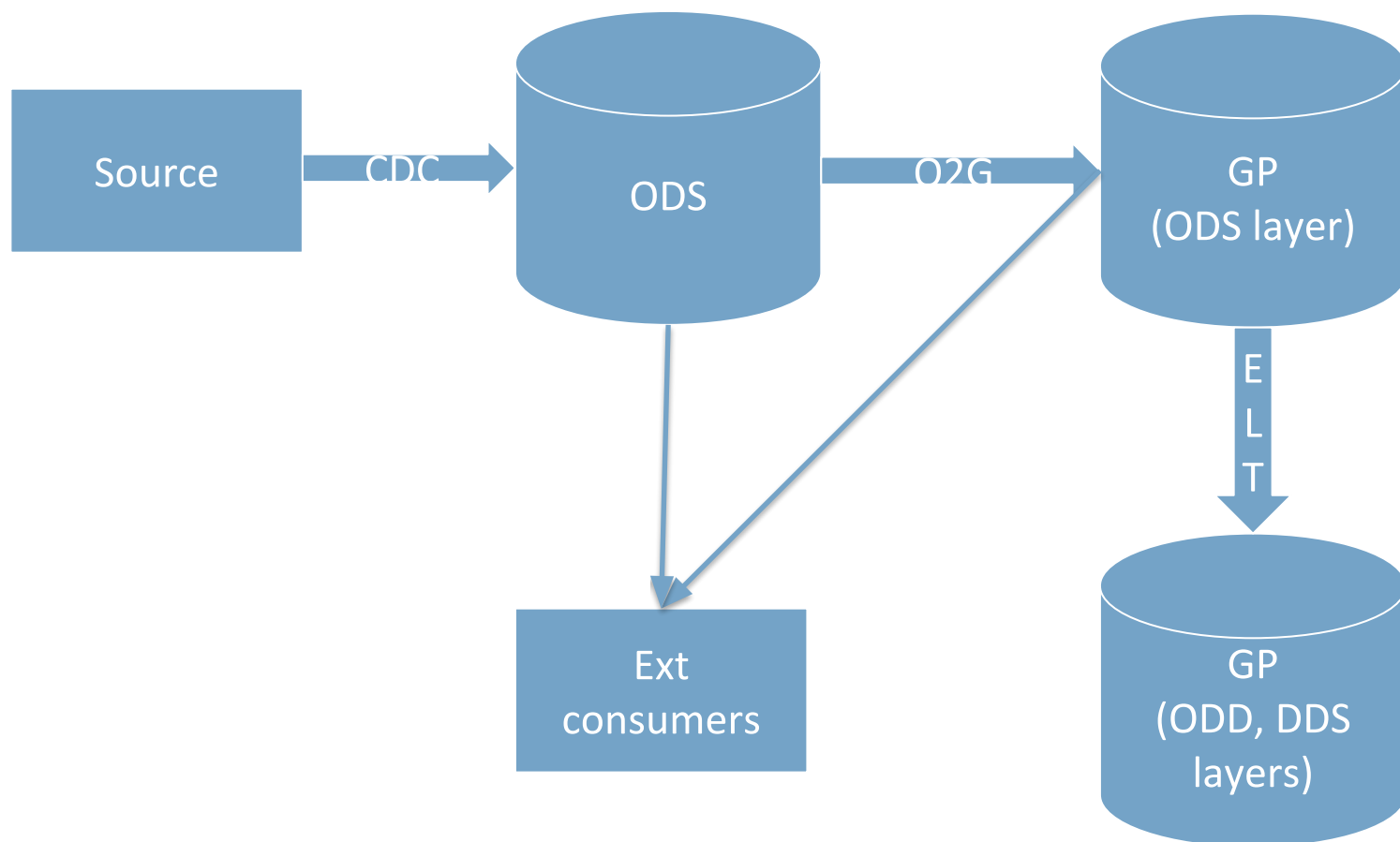
key	f1	f2	f3
1	f1111	f22	f333



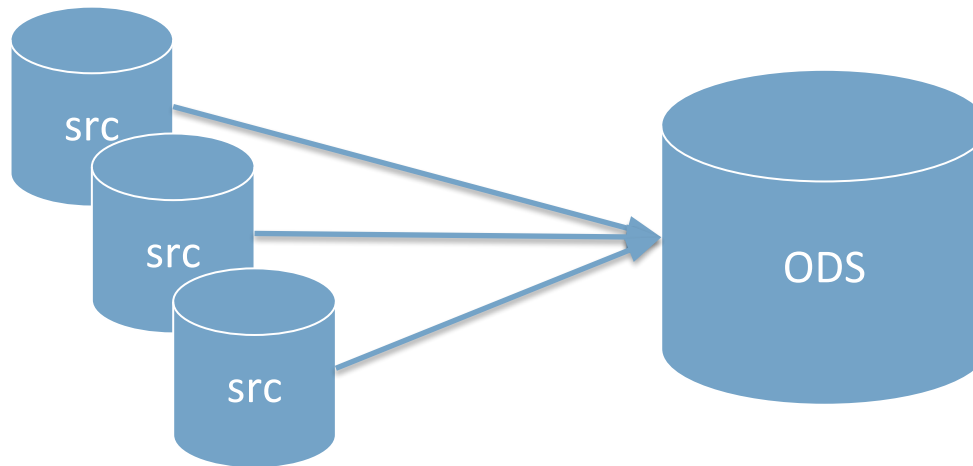
Этап 2. Причины отказа

- Огромная нагрузка на ввод/вывод в GP, как следствие - частый выход из строя дисков
- Большая часть транзакций в GP (80% от общего количества) – от репликатора
- Задержка при переходе в режим «row-by-row»
- ВОХ-решение: не было multitarget
- ВОХ-решение: плохая кастомизация

Этап 3 (Текущий, общая архитектура)

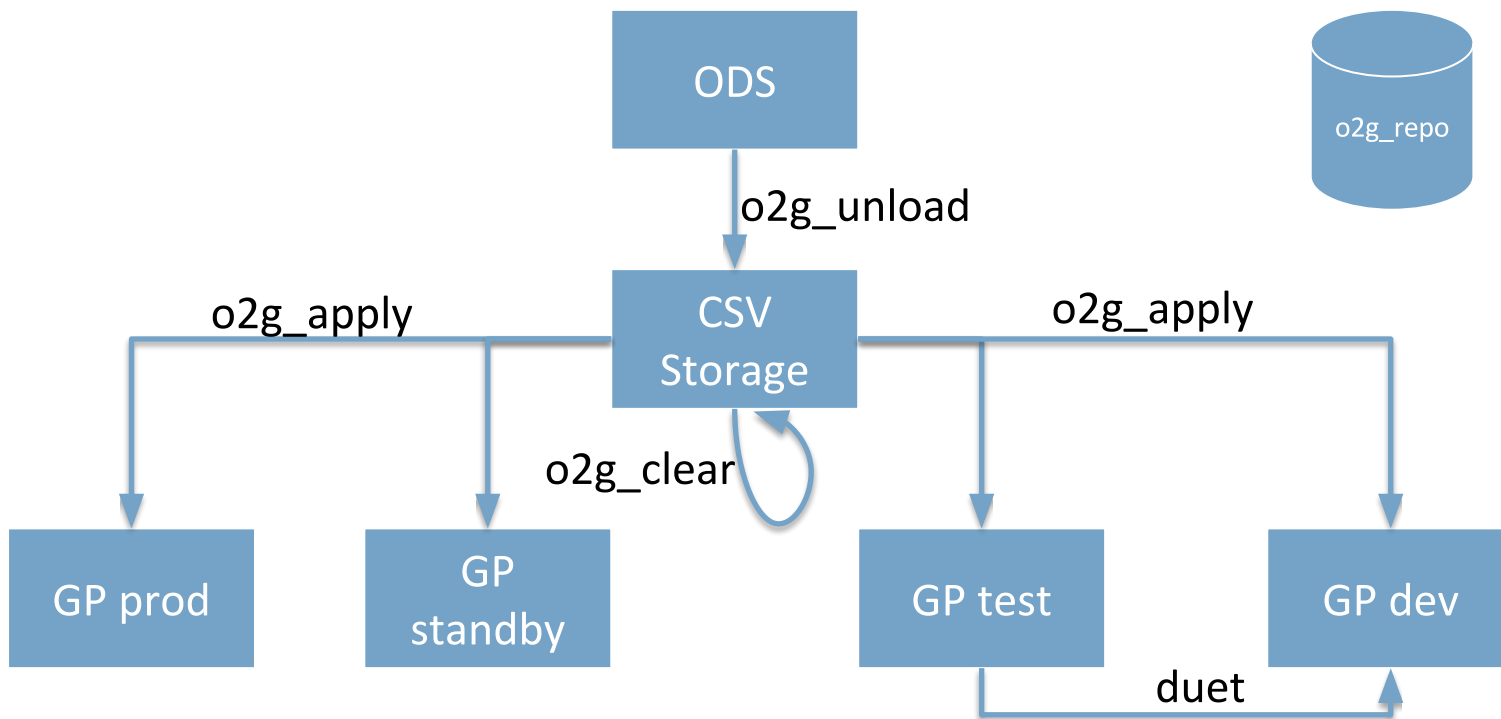


Этап 3. ODS



Репликация работает в режиме «Real time data warehousing» – когда несколько источников реплицируются в один приемник.

Этап 3. Архитектура O2G



Этап 3. 3 типа актуальности данных



- ➔ Online (в ODS). Online операционная отчетность
- ➔ 3-ех часовая задержка. Регулярная операционная отчетность
- ➔ Раз в сутки. Регулярные ETL-процессы, построение модели данных DWH



Этап 3. Преимущества

- Полностью лег на архитектуру ELT процессов, разработанных во время жизни этапа 2
- Снижена нагрузка на ввод/вывод GP
- Позволило сжимать данные и хранить их поколonoчно – экономия места, оптимизация ввода/вывода
- Multitarget – все контуры не зависимы и ничего не знают друг о друге
- Другие потребители данных ODS (шина...)



Количественные характеристики

- кол-во таблиц в ODS: 1600 из 27 различных источников
- общий размер ODS в Oracle: 36 Тб
- общий размер реплик ODS в GreenPlum: 7,7 Тб
- задержка репликации источник->ODS: близка к 0
- количество DML операций в сутки: 1,5В
- задержка репликации ODS->Greenplum: до трех часов (регулируется администратором)
- Полностью самостоятельная система, минимум времени для поддержки

Планы на будущее



- Оставляем в ODS только _ct таблицы
- Обработку изменений выносим в отдельный in-memory движок



Тинькофф

Спасибо за внимание

[Tinkoff.ru](https://tinkoff.ru)