

ДОСЛІДЖЕННЯ МЕХАНІЗМУ МІГРАЦІЇ ДАНИХ, ЩО ЗБЕРІГАЮТЬСЯ В РЕЛЯЦІЙНИХ СУБД

Міграція даних – процес перенесення даних з вихідної БД в цільову, схеми яких можуть відрізнятися. Причин, з яких організації починають проекти з міграції даних, може бути досить багато: від оновлень додатків і впровадження нових корпоративних систем до повномасштабної реструктуризації внаслідок злиття компаній.

На даний момент не існує єдиного підходу до вирішення проблеми міграції даних. Фірма, яка розробляє ПЗ, не завжди підтримує можливість збереження накопичених даних при переході на більш пізню версію програмного продукту. Розробникам доводиться аналізувати всі зміни та кожен раз створювати нові процедури перенесення даних для конкретного випадку зміни схеми БД.

Найпростішим варіантом реалізації механізму міграції даних (ММД) є створення проміжної програми. Вона повинна зв'язуватися з вихідною і цільовою БД і виконувати необхідні перетворення.

Інший варіант міграції даних (кращий при переході на нову СУБД) – попереднє завантаження старих даних в тимчасові таблиці нової СУБД. Сучасні СУБД (наприклад, Oracle) зазвичай містять спеціальні утиліти, які дозволяють здійснювати дуже швидко завантаження зовнішніх даних різних форматів. В такому випадку модуль міграції пишеться засобами мов, вбудованих в СУБД, наприклад, PL/SQL або Java.

Кілька популярних засобів для перенесення даних: Migration Architect, TRUmigrate, Flyway.

ММД повинен гарантувати перегляд і перенесення всіх записів з вихідної БД, цілісність даних, а також, що перенесення записів буде здійснюватися з урахуванням всіх залежностей. Запис може бути перенесений тільки в тому випадку, якщо він є незалежним, або записи, від яких залежать дані, вже перенесені в цільову БД. Тому основним правилом, що визначає порядок перегляду та перенесення, є залежності між записам, які визначаються за зовнішніми ключам. ММД реалізується у вигляді процедури, яка спочатку викликається для незалежних записів, а далі рекурсивно і для залежних записів від цього запису. Порядок пересування між ними повинен починатися з таблиці, що містить незалежні записи. Тоді основний порядок, визначений на множині таблиць схеми БД – залежність між таблицями. При такому підході схема БД вважається поданою у вигляді орієнтованого графа. Як вершини виступають таблиці, а в якості дуг – залежності між таблицями, причому дуга виходить з незалежної таблиці (вершини) і входить в залежну таблицю (вершину). Для кожної незалежної вершини будується своє дерево алгоритмом перегляду вглиб. Суміщенням дерев перегляду для всіх незалежних вершин виходить загальне дерево перегляду. Це і буде порядок пересування, причому ММД можна починати з будь-якої незалежної вершини (таблиці). Сформована послідовність і необхідна інформація про схему вихідної БД для виконання ММД (інформація про таблиці та зв'язки між ними) зберігається в файлі конфігурації. Далі цей файл аналізується і використовується при перенесенні даних.

Додаток, що виконує механізм міграції даних написано в середовищі NetBeans IDE 6.8 на мові Java.