

USING CA ERWIN MODEL MANAGER FOR DATA WAREHOUSE STRUCTURE MODELING

S. Gribkov, L. Zagorovska

National University of Food Technologies

Key words:

Data model

Data warehouse

Means of collective design

CA Erwin Model

Manager

Article history:

Received 07.09.2014

Received in revised form

10.10.2014

Accepted 30.10.2014

Corresponding author:

S. Gribkov

Email:

sergio_nuft@ukr.net

ABSTRACT

The problems encountered when creating a data warehouse are considered, and it is proposed to use CA Erwin Model Manager to avoid them. The paper describes the use of CA Erwin Model Manager, a software tool for corporate design and modeling, in order to simulate the structure of a data warehouse for VAT "Makaronna fabryka" decision support system.

ВИКОРИСТАННЯ CA ERWIN MODEL MANAGER ДЛЯ МОДЕЛЮВАННЯ СТРУКТУРИ СХОВИЩА ДАНИХ

С.В. Грибков, Л.Г. Загоровська

Національний університет харчових технологій

У статті розглянуто проблеми, що виникають при створенні сховища даних. Для уникнення зазначених проблем запропоновано використовувати засіб колективного проектування складних моделей AllFusion Model Manager, ефективність якого підтверджена при створенні сховища даних системи підтримки прийняття рішень для VAT «Макаронна фабрика».

Ключові слова: *модель даних, сховище даних, засіб колективного проектування, Model Manager.*

В основі сучасних систем підтримки прийняття рішень використовують сховища, що слугують інформаційним джерелом для забезпечення основних функцій даними [1]. Використання сховища даних забезпечує проведення аналізу даних за складом і змінами протягом досить великого проміжку часу, що дає змогу в процесі прийняття рішень використовувати інформацію за попередній період й оцінювати наслідки подібних ситуацій [2, 3]. Крім цього, використання сховища даних прискорює формування звітів як результату виконання складних запитів, що потребує обробки великої кількості записів

зі складними підрахунками. Все це досягається за рахунок централізованого зберігання даних у сховищі даних і використання багатовимірного подання інформації з ігноруванням певних вимог нормалізації, за рахунок дотримання максимум 3-ої нормальної форми, що значно підвищує швидкість опрацювання інформації, оскільки зменшує кількість операцій на вибірку даних із з'єднаних таблиць [3].

Етап проектування сховища даних є одним із найважливіших кроків при створенні інформаційних систем. При проектуванні сховища даних використовують CASE-засіб AllFusion ERwin Data Modeler, призначений для проектування моделей даних на різних рівнях з можливістю використання різних підходів [4]. При проектуванні сховища даних виникають проблеми, що впливають на якість і швидкість розробки, а саме: розмір сховища даних є досить великим і налічує декілька десятків таблиць; склад і кількість колективу розробників, що займаються створенням сховища даних. Обидва фактори взаємозв'язані. Кожен з колективу розробників займається проектуванням певної частини сховища даних, потім частини об'єднуються й таким чином формується повна схема сховища. Однак при цьому виникають такі проблеми: узгодження елементів структури при об'єднанні; зберігання та класифікація проміжних і кінцевих моделей структур тощо.

Для уникнення зазначених проблем пропонується використання засобу колективного проектування складних моделей AllFusion Model Manager, ефективність якого підтверджено при створенні сховища даних системи підтримки прийняття рішень для ВАТ «Макаронна фабрика». AllFusion Model Manager являє собою масштабований багатокористувацький засіб моделювання, який забезпечує ефективну спільну роботу фахівців для створення моделей. Завдяки його використанню забезпечується можливість централізованого зберігання та редагування різних версій моделей, порівняння їх в автоматизованому режимі, створення на їх основі нових моделей шляхом злиття або вичленення складових, а також контроль доступу до різних моделей для проектувальників. AllFusion Model Manager реалізований на основі клієнт-серверної архітектури. Клієнтська частина — інтегрована безпосередньо в AllFusion Data Modeler і AllFusion Process Modeler. Серверна частина (репозитарій) — сховище моделей, організоване у базі даних клієнт-серверної СУБД. AllFusion Model Manager забезпечує захист моделей і надає різний рівень доступу до них, що є незамінним для керівників проектів при координації та контролі за ходом усієї роботи. Засіб підтримує функціональність та синхронізацію моделей процесів і моделей даних, що забезпечує визначення відповідності цих моделей [4]. Model Manager надає адміністратору можливість розмежування доступу користувачів до моделей, у тому числі на рівні бібліотек. При збереженні нової моделі необхідно вказати, до якої бібліотеки вона буде входити.

Для реалізації репозитарію моделей у СУБД MS SQL Server створено БД «KMFV_RepModel». За допомогою майстра AllFusion Model Manager створено структуру репозитарію. AllFusion Model Manager забезпечив роботу з репозитарієм моделей, створених у AllFusion Erwin Data Moduler та AllFusion Process Modeler.

Репозитарій логічно розподілений на бібліотеки (Library), необхідні для збереження всіх моделей. Кожна бібліотека призначена для зберігання моделей певного призначення, наприклад, бібліотека «FunctionModel» призначена для зберігання функціональних моделей. У бібліотеках автоматично створюються версії моделей, які використовуються для відстеження змін у моделях при груповій роботі та забезпечують можливість відмовитися від запропонованих змін. Крім цього, існують позначені версії моделей, що створює користувач. AllFusion Model Manager надає адміністратору можливості розмежування доступу користувачів до моделей, у тому числі на рівні бібліотек. При збереженні нової моделі необхідно вказати, до якої бібліотеки вона буде входити. Підмоделі, або предметні області (Subject Area), забезпечують окремим розробникам більш чітке тематичне уявлення та допомагають зменшити використання системних ресурсів, необхідних для завантаження або вивантаження з репозитарію. Менеджер моделей забезпечує логічне розподілення моделей на підмоделі. Розробники моделей можуть працювати з окремою підмоделлю.

Створення моделі сховища даних розпочали зі створення моделей даних джерел, з яких воно буде наповнюватися. Кожна окрема модель даних відповідає структурі конкретного джерела даних. Створення моделей для джерел, що не використовують СУБД і зберігають інформацію у текстових файлах або у форматах файлів Microsoft Office, відбувалося відразу згідно зі стандартами реляційної бази даних. Для джерел даних, що реалізовані у СУБД, застосовано автоматизований підхід зі створення моделей даних шляхом отримання структури бази даних в окремій моделі із застосуванням зворотного проектування засобами AllFusion Erwin Data Moduler.

Для отримання структури бази даних кожного джерела виконувались такі дії: виклик майстра зворотного проектування Reverse Engineer; обов'язково здійснювався вибір типу моделі з логічним і фізичним рівнем, що дозволяє повністю описати модель даних та її структуру; визначення СУБД, на основі якої реалізована база даних джерела; визначення опцій зворотного проектування таким чином, щоб отримати найбільш повну структуру з усіма об'єктами бази даних (таблиці, представлення, ключі, індекси та зв'язки тощо); встановлення опцій підключення до бази даних із зазначенням прав доступу; здійснення генерації структури та її відображення. Після виконання зазначених дій для кожного джерела отримали модель даних, що відображала структуру бази даних на фізичному рівні. Як джерело даних обрано інформаційні системи й електронні документи, що використовуються в контурі управління макаронним виробництвом. Таким способом були згенеровані моделі даних інформаційних систем: «Змінний журнал виробництва», «Рух продукції», «Лабораторія», «Pavan», «Ремонтно-профілактичні роботи», «Договори та замовлення», «Відділ кадрів», «План виробництва», «Рух сировини та матеріалів», «Облік залишків», «Облік руху некондиційної продукції» та ін.

Побудова моделей даних для джерел, реалізованих у вигляді плоских файлів, наприклад, текстових, здійснювалася в покроковому режимі з обробкою лише мінімально необхідної інформації. Кожна модель даних має

свою назву та занесена до створеної бібліотеки «DD» репозитарію моделей, призначеної для збереження моделей, що відображають структури джерел даних. Назва моделі даних цієї групи починається з «DD_», а далі вказується назва джерела даних, на основі якого вона побудована, наприклад, «DD_MagazineProduction» — модель даних для БД інформаційної системи «Змінний журнал виробництва». Перелік створених моделей, занесених до групи «DD», можна переглянути в менеджері бібліотек AllFusion Model Manager, також існує можливість знищити або перейменувати всю групу чи окрему модель. Для отримання повної інформації про кожну модель у цілому та її складові формувались звіти за допомогою функції Report.

Після створення моделей джерел даних формувалися підмоделі сховища даних. При цьому кожна модель, що відображає джерела даних, була приведена до третьої нормальної форми на фізичному рівні та збережена в окремій бібліотеці моделей «ComponentsSAD». Назви моделей цієї групи починаються з префікса «C_», а далі вказується назва джерела даних, на основі якого вона побудована, наприклад, «C_MagazineProduction» — модель даних для бази даних інформаційної системи «Змінний журнал виробництва».

У результаті виконання описаних процедур отримали набір моделей для побудови стандартного архіву даних. Для подальшого їх об'єднання в єдину модель використана функція «Complect Compare» в додаткових функціях AllFusion ERwin Data Modeler з можливістю порівняння складових моделей, що дозволяє внести до її складу лише необхідні компоненти. На рис. зображено вікно результату порівняння моделей даних інформаційних систем «Лабораторія» та «Змінний журнал виробництва», з якого чітко видно, що в моделі праворуч наявна сутність «Контроль якості сировини» з усіма складовими, а в моделі ліворуч напроти цієї сутності закреслені комірки, що означає її відсутність.

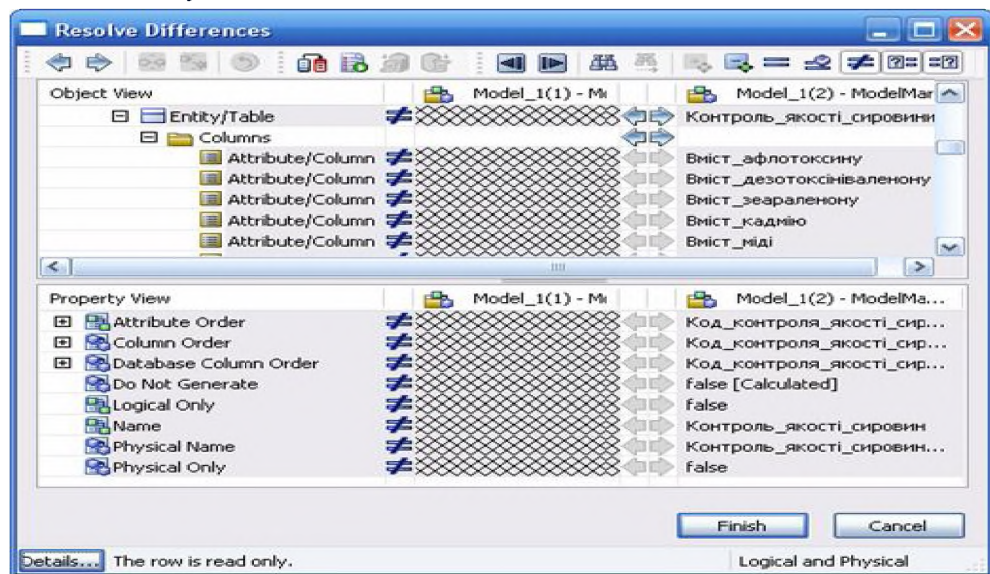


Рис. Вікно порівняння і перенесення складових моделей

За необхідності переносимо сутності чи їхні складові до необхідної моделі, наприклад, переносимо сутність «Контроль якості сировини» до моделі «Змінний журнал виробництва». Після внесення усіх змін та остаточної перевірки зберігаємо нову модель у бібліотеці «SAD» репозитарію моделей. Ім'я моделі формуємо шляхом інтеграції префікса «SAD_» та кроку ітерації, отримавши модель «SAD_1» та «SAD_2» на першому й другому кроці відповідно. Зазначені дії повторюємо до тих пір, поки не об'єднаємо всі моделі в одну, отримавши модель сховища даних.

Для ведення хронології накопичення інформації у сховищі даних створено таблицю «Час надходження даних», що складається з таких полів: «Код часу надходження», «Дата і час», «День», «Місяць», «День тижня», «Квартал», «Рік». Необхідність використання цієї таблиці обумовлена тим, що сховище даних накопичує інформацію з різних джерел, яка потребує, окрім змістовної частини збереження, ще й час її надходження. Накопичення інформації у сховищі даних відбувається автоматично за встановленим розкладом, наприклад, щодня запускаються певні модулі, що збирають інформацію з однієї або кількох оперативних баз даних. Після перетворення до заданого формату та форми інформація записується у сховище даних. Модель сховища даних, побудована описаним способом, складається з великої кількості сутностей, через що представити її на одному аркуші неможливо.

Висновки

Використання AllFusion Model Manager при колективному моделюванні сховища даних забезпечує створення адекватної та несуперечливої моделі даних за найкоротший час. За рахунок репозитарію моделей забезпечується зберігання моделей даних і розподілений доступ до них. Управління версіями моделей в AllFusion Model Manager дозволило розробникам відкривати з репозитарію необхідні версії моделі, блокувати модель для внесення змін іншими розробниками, зберігати версії моделей у репозитарії, аналізувати історію моделі та змін і, якщо потрібно, легко повертати моделі до попереднього стану, порівнювати між собою будь-які дві версії однієї моделі. Спільна робота забезпечується можливістю автоматично об'єднувати різні самостійні моделі. Об'єкти незалежних моделей можна скопіювати в об'єднані моделі даних за допомогою менеджера злиття, заснованого на інструменті ERwin Complete Compare.

Література

1. Грибков С.В. Моделирование структуры системы поддержки принятия решений для предприятий харчової галузі / С.В. Грибков, Л.Г. Загоровська // Харчова промисловість. — 2007. — № 5. — С. 100—102.
2. Rainardi V. Building a Data Warehouse: With Examples in SQL Server, Second Edition / V. Rainardi. — Apress, 2008. — 541 p.
3. Спирли Э. Корпоративные хранилища данных. Планирование, разработка, реализация: пер. с англ. / Э. Спирли. — М.: Вильямс, 2001. — Т. 1. — 400 с.
4. Маклаков С.В. Создание информационных систем с AllFusion Modeling Suite / С.В. Маклаков. — М.: ДИАЛОГ-МИФ, 2005. — 432 с.

ИСПОЛЬЗОВАНИЕ СА ERWIN MODEL MANAGER ДЛЯ МОДЕЛИРОВАНИЯ СТРУКТУРЫ ХРАНИЛИЩА ДАННЫХ

С.В. Грибков, Л.Г. Загоровская

Национальный университет пищевых технологий

В статье рассмотрены проблемы, возникающие при создании хранилища данных. Для избежания проблем предложено использовать программное средство коллективного проектирования и моделирования СА Erwin Model Manager, эффективность которого доказана при создании хранилища данных системы поддержки принятия решения для ОАО «Макаронная фабрика».

Ключевые слова: модель данных, хранилище данных, средство коллективного проектирования, СА Erwin Model Manager.