

УДК 330.1(075.8)

О.В. ТЕЛЕПНЄВ

Харківський національний університет ім. В.Н. Каразіна, Україна

ОСОБЛИВОСТІ НАКОПИЧУВАЧА ДАНИХ ПОДІЙНО-ОРІЄНТОВАНОЇ СИСТЕМИ ПІДТРИМКИ ЖИТТЄВОГО ЦИКЛУ ПРОДУКТІВ

Розглянуто особливості накопичувачів даних, що застосовуються для різних цілей. Наведено характерні ознаки накопичувача даних подійно-орієнтованої інформаційно-аналітичної системи підтримки життєвого циклу продуктів.

накопичувач даних (Data Warehouse), метадані, багатовимірне моделювання (Dimensional Modeling), життєвий цикл продукту, подія, подійно-орієнтована інформаційно-аналітична система підтримки життєвого циклу складних виробів

Вступ

Подійно-орієнтована інформаційно-аналітична система підтримки життєвого циклу складних виробів (ІАС ЖЦ) допомагає вирішувати багато задач, що встають перед співробітниками організацій, які беруть участь у процесах життєвого циклу виробів. Вона дозволяє консолідувати різноманітні дані про виріб на основі подійно-орієнтованого підходу. Ці дані розміщуються у накопичувачі даних системи, що є однією з її головних частин. У цій статті розглядаються особливості накопичувачів даних як інтегрованого джерела даних для інформаційно-аналітичних систем, та місце накопичувача даних ІАС ЖЦ серед накопичувачів для різних галузей.

Накопичувач даних – інтегроване інформаційне джерело для інформаційно-аналітичних систем

Інформаційно-аналітичні системи з'явилися на підприємствах слідом за системами оперативної обробки даних – СОД (OLTP – On-Line Transaction Processing, оперативна транзакційна обробка) як засіб представлення даних. Спочатку вони були надбудовою над вже існуючими системами та дозволяли отримувати звіти попередньо заданої струк-

тури. Незабаром з'ясувалося, що крім стандартних звітів аналітикам потрібна додаткова інформація, що дозволяє дивитися на бізнес-процес з різних точок зору, а сам процес аналізу даних є ітераційним. Значно більші можливості давали звіти “гнучкої” структури, проте їх використання потребувало від користувачів додаткових знань. Структура баз даних систем оперативної обробки нормалізувалася для забезпечення цілісності даних, що призвело до створення великої кількості маленьких таблиць, описів бізнес-сутностей, охопити яку було важко навіть спеціалісту у сфері баз даних. Орієнтація систем оперативної обробки на процес, а не предметну область, не дозволяла в рамках однієї системи отримати всю інформацію про об'єкт дослідження. Та й створення цих звітів при великому обсязі інформації займало досить багато часу, позбавляючи аналітичний процес оперативності.

Для об'єднання даних про предметну область була запропонована концепція віртуального інтегрованого джерела даних, що опиралася на бази даних розподілених СОД [1, 2], та передбачала транслявання «на льоту» запитів від віртуального джерела до баз даних, емулюючи роботу з даними цієї системи, немов би з накопичувачем даних. У основі віртуального накопичувача – система “представ-

лень” (view-таблиць, що містять результати запитів до “базових” таблиць та, на відміну від них, не містять власних даних) у базі даних, або спеціальні засоби доступу класу Desktop OLAP (Business Objects, Brio Enterprise). Позитивними сторонами цього підходу є відносна простота та незначна ціна реалізації.

Та існує й багато недоліків:

- результатом створення віртуального накопичувача є не він сам, а лише ілюзія його існування;
- оскільки структура збереження даних не змінюється, залишаються проблеми чистоти даних (трансформації даних, інтеграції даних з різних джерел), відсутність історії (віртуальний накопичувач містить лише оперативні дані);
- інтегрований погляд на розподілене віртуальне корпоративне джерело даних можливий лише при функціонуванні усіх СОД, що його формують;
- нормалізація баз даних СОД призводить до звернення у одному запиті до декількох таблиць, що знижує швидкодію. До того ж час обробки запиту до розподіленого накопичувача перевищує ці показники у централізованого джерела;
- виконання складних аналітичних запитів над таблицями СОД значно знижує швидкодію СОД;
- різні СОД можуть мати різні формати та кодування даних. Можливе неузгодження даних пов'язане з несинхроністю моментів оновлення даних, відмінностями у трактуванні подій, помилками введення даних, тощо. Все це не дозволяє формувати несуперечливий погляд на об'єкт аналізу;
- через те, що СОД виконують оперативну обробку, вони містять дані лише за короткий (декілька місяців) інтервал часу. Для аналітичної ж обробки найбільш цікавий погляд на об'єкт аналізу у історичній ретроспективі.

Рішенням, що було позбавлено цих недоліків, стало створення спеціальних інформаційних структур, що отримали назву накопичувач даних (Data Warehouse – склад, сховище даних) . Автор цієї концепції Білл Інмон (W.H. Inmon) визначив його як

“предметно-орієнтовані, інтегровані, незмінювані, підтримуючі хронологію набори даних, організовані задля підтримки управління”, що призначені виступати у ролі “об'єданого та єдиного джерела істини”, яке забезпечує менеджерів та аналітиків достовірною інформацією, необхідною для оперативного аналізу й підтримки прийняття рішень[3]. Концепція накопичувача даних пропонує не лише об'єднаний логічний погляд на дані організації, але й фізичне створення єдиного інтегрованого джерела даних.

Інмон зазначає, що у накопичувачі даних повинні міститися лише атомарні дані (дані про одиничний екземпляр сутності), а будь-які узагальнення (результати усереднення, інтегрування, тощо), що змінюються з часом, необхідно переносити до вітрин даних та аналітичних додатків. “Необхідно, щоб набір параметрів (атрибутів) даних дозволяв характеризувати сутність настільки широко, як це можливо. Іншими словами, атомарні дані у накопичувачі повинні бути настільки повними та репрезентувати якнайбільше класів та категорій даних”.

Основні відмінності накопичувачів даних та баз даних корпоративних систем оперативної обробки даних наступні.

Більшість інформаційних систем оперативної обробки даних:

- є процедурно-зорієнтованими: дані та звіти організовані на основі корпоративних блоків;
- є не інтегрованими: може застосовувати декілька інформаційних систем (додатків) для виконання окремих операцій;
- містять дані, які є постійними у часі та просторі: дані системи є спільними для визначеного часу та місця;
- є змінюваними: містять дані, які часто змінюються.

Накопичувач даних:

- є об'єктно-зорієтованим: дані та звіти організовані на основі об'єктів, значущих для бізнесу (наприклад, продукти, клієнти, співробітники, тощо);

- є інтегрованим: дані до накопичувача поступають з різних систем, попередньо проходячи узгодження метаданих та даних, виправлення помилок – т. з. “очистку даних”;

- містить дані, які є змінними у часі та просторі: дані містять часову й просторову мітку, які дозволяють формувати звітність на цій основі;

- є незмінюваним: дані, які було внесено до накопичувача, зазвичай не змінюються.

Застосування накопичувачів даних забезпечує [4]:

- прискорення оперативних (on-line) запитів та звітів, підвищення їхньої вірогідності;

- збільшення обсягів даних рівня підприємства (інформаційно інтегрованих організацій), завдяки можливості створення деталізованих звітів, які дозволяють виконувати узагальнення будь-якого (з припустимих) рівнів, включаючи й найвищий;

- забезпечення єдиного узгодженого погляду на дані усіма співробітниками організації (з урахуванням прав доступу до інформації);

- покращення технологічності: великий обсяг легкосприйнятної інформації з накопичувача може бути передумовою для удосконалення уявлень про бізнес-процеси.

Однак, є випадки, коли створення накопичувача даних не приносить бажаного ефекту. Часто ці недоліки пов'язані не з технічною складністю реалізації накопичувача, а із особливостями організаційної структури організації, її бізнес-процесів. Такими випадками є:

- можливі обмеження застосування історичних даних з накопичувача в умовах частих змін в організації (структура, бізнес-процеси, продукти), у зовнішньому оточенні організації (наприклад, ринкової кон'юнктури);

- спрощені бізнес-процеси організації дозволяють обмежити аналітичні потреби стандартними звітами, створеними на основі даних з систем оперативної обробки;

- значні затрати часу та матеріальних ресурсів на навчання персоналу й обслуговування накопичувача даних та інформаційних систем, які його використовують чи обробляють дані для нього;

- значний час та висока ціна підготовки даних для накопичувача.

Широке застосування інформаційно-аналітичних систем у різних напрямках ділової активності обумовило створення у організації специфічних джерел даних для аналітики. Свої відмінності мають фінансові накопичувачі, накопичувачі для підтримки роботи із клієнтами, накопичувачі даних про людські ресурси тощо. Поряд з цими спільними для підприємств інформаційними структурами існують також накопичувачі для підтримки специфічних для конкретної організації видів діяльності – накопичувачі даних щодо закупівлі виробів, підтримки у експлуатації програмно-технічних засобів організації, підтримки взаємодії із клієнтами та інші.

Фінансовий накопичувач даних, зазвичай, є першою подібною структурою у організації. Це обумовлено наступними факторами:

- фінансові дані завжди у центрі уваги організації, що збільшує зацікавленість керівництва в успіху створення накопичувача;

- фінансові дані репрезентують невеликий обсяг даних організації;

- фінансові дані описують багато аспектів діяльності організації й її підрозділів та мають спільну одиницю виміру – гроші.

Сфери управління даними про людські ресурси організації, управління закупівлями, управління відносинами із клієнтами (включаючи й підтримку продукції в експлуатації), управління технічним обслуговуванням та ремонтом обладнання організації відрізняються проміж інших незначною частотою трансакцій, порівняно низькою інтенсивністю уведення нової інформації, направленістю на накопичення інформації та тривалим терміном її зберігання. Тому тут часто об'єднують системи операти-

вної обробки даних з системами виконання аналітичних задач.

Деяко осторонь від попередніх накопичувачів стоять накопичувачі даних для виконання інтелектуального аналізу даних (Data Mining/Data Mining and Exploration Data Warehouse). За визначенням Інмона [3], такий накопичувач – “це гібридна форма класичного накопичувача даних, яка є:

- надзвичайно деталізована;
- надзвичайно хронологічна;
- оптимізована для аналізу.

Такий накопичувач є зорієнтованим на вирішення конкретної аналітичної задачі (або задач), й тому потреба в ньому, на відміну від будь-яких інших накопичувачів, зникає після її розв'язання. Отже необхідно, щоб аналітичні задачі, для яких створюється спеціальний накопичувач, виникали часто, а ефект від їхнього вирішення був значним. Типовим прикладом подібного накопичувача є інформаційна структура, що застосовується у аналізі причин та наслідків дефектності у виробництві [5].

Інформаційні структури, описані раніше, є проблемно-зорієнтованими, призначені для інформаційної підтримки окремих напрямків діяльності організації та дають можливість розглядати її під окремим кутом зору. Поряд з ними є значна потреба у створенні інформаційного джерела, що давало б загальний погляд на діяльність організації в цілому. Особливу потребу в цьому мають технічні працівники різних рівнів. Слід зазначити, що така інформаційна структура має доповнювати, а не замінити проблемно-зорієнтовані накопичувачі.

Важливим питанням є вироблення методології гуртування цих даних, вибір набору параметрів тощо. Відповідь на це питання специфічна для організацій різних сфер діяльності. Для тих, що беруть участь у етапах життєвого циклу виробів, є потреба у створенні окремої специфічної інформаційної структури, у якій можна було накопичувати для подальшого аналізу дані про всі значущі події, які тра-

пилися під час цієї діяльності. Для вирішення цієї задачі й пропонується створити накопичувач даних для подійно-орієнтованої інформаційно-аналітичної системи підтримки життєвого циклу продуктів (ІАС ЖЦ).

Метою цієї статті є:

- окреслення місця накопичувача даних для ІАС ЖЦ у інформаційній інфраструктурі підприємства та взаємозв'язків між ним та іншими накопичувачами даних;
- визначення особливостей накопичувача даних для ІАС ЖЦ та обґрунтування доцільності його створення.

Накопичувач даних для ІАС ЖЦ

Накопичувач даних подійно-орієнтованої інформаційно-аналітичної системи підтримки життєвого циклу продуктів (ІАС ЖЦ) є типовим прикладом накопичувача, зорієнтованого на вирішення конкретної аналітичної задачі – інформаційної підтримки процесів життєвого циклу виробів.

Особливістю ІАС ЖЦ та її накопичувача даних є природа даних. Події трапляються в різних організаціях – учасниках процесів життєвого циклу виробів. Отже, для отримання повної картини необхідно збирати у одному місці усі дані які, зазвичай, є доступними лише одній з організацій. Таким чином, накопичувач даних ІАС ЖЦ стає “надорганізаційною” інформаційною структурою, існування якої потребує налагодження інформаційного обміну між усіма зацікавленими організаціями.

Подійно-орієнтований підхід дає можливість застосувати уніфікований набір параметрів [6], який дозволяє описувати усі значущі для організацій події з достатнім для багатьох аналітичних задач ступенем деталізації.

Для збирання цих даних із застосуванням уніфікованого набору запропонована концептуальна модель накопичувача, що надана на рис. 1.

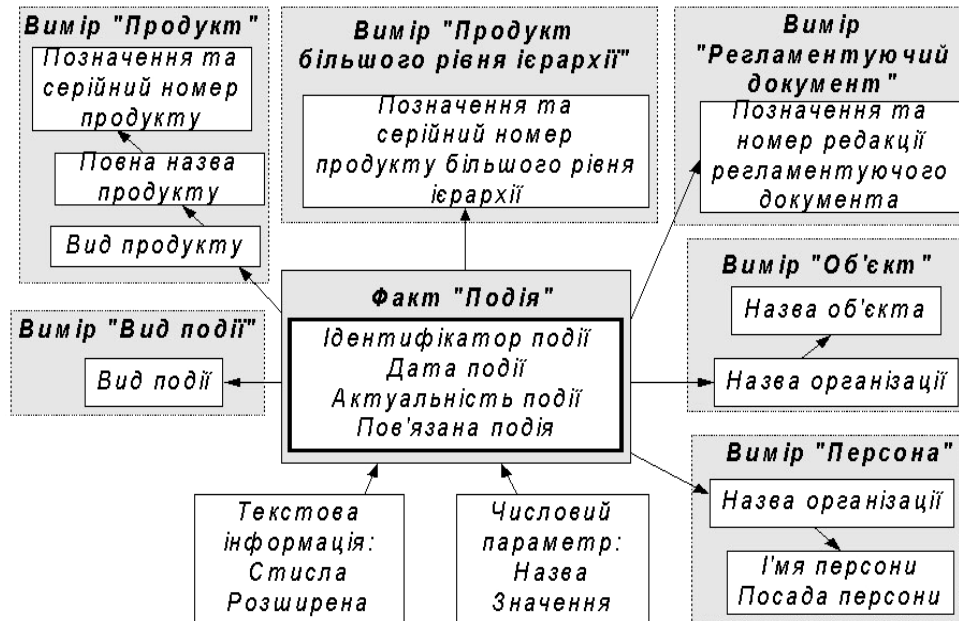


Рис. 1. Концептуальна модель накопичувача даних ІАС ЖЦ

Слід зазначити, що подійно-орієнтований підхід може бути застосований не лише у підтримці життєвого циклу продуктів, але й у інших сферах діяльності організацій, де відбуваються різноманітні події та може бути застосовано поняття життєвий цикл. Прикладом таких сфер можуть бути управління персоналом організації, документообігом, експлуатацією основних фондів організації та інше. Результатом об'єднання цих даних є можливість переходу від аналізу стану продуктів до часткового (фінансова сфера, здебільшого з погляду обмежень доступу, відокремлюється від інших) аналізу діяльності організацій в цілому. Таким чином, накопичувач даних ІАС ЖЦ стає інтегрованим інформаційним джерелом, що дозволяє вирішувати широке коло аналітичних задач:

- отримання статистики про термін створення документів різних типів, їх рух по “маршруту узгодження”;
- здобуття інформації про відмінності нової редакції документа, підстави внесення змін;
- визначення переліку виконаних за період часу робіт з можливістю аналізу по працівникам, підрозділам, об'єктам тощо;

- отримання статистики про термін поставок, динаміку цін на комплектуючі з можливістю аналізу по постачальникам, комплектуючим тощо;
- отримання інформації про склад кожного окремого екземпляру продукції;
- пошук усіх екземплярів продукції, в яких є компоненти з заданим набором параметрів (постачальник, партія поставки тощо);
- пошук різних типів продукції, в яких використано задані компоненти;
- пошук інформації про продукцію, яка поставлена на об'єкти з заданими характеристиками зовнішнього впливу;
- отримання статистики про термін введення у експлуатацію, термін експлуатації, напрацювання на відмову, термін відновлення працездатності;
- отримання даних про причини відмов та методи відновлення працездатності, з можливістю аналізу по організаціях-замовниках, об'єктах, типах продукції, комплектуючих постачальникам, партіях тощо.

Доцільність впровадження такого накопичувача підтверджується великою потребою у нижчеперерахованому:

- об'єднанні розрізаних на цей час даних про значущі для організацій події;

- представленні цих даних у єдиному інтегрованому вигляді для усіх співробітників організацій, що мають у них потребу та відповідні права доступу до інформації;

- виконанні широкого кола задач щодо аналізу даних про події.

Ознаками накопичувача даних подійно-орієнтованої системи є:

- стале зростання обсягів даних й дуже значний термін їхнього зберігання, це обумовлено як суттєвим терміном експлуатації складних продуктів, так наступністю технічних рішень, які застосовуються у нових продуктах;

- надзвичайно велика інтенсивність уведення даних;

- значна кількість користувачів, що одночасно виконують різні аналітичні задачі з використанням даних із накопичувача.

Такий накопичувач можна вважати "елементом колективної пам'яті" організації.

Висновки

В статті розглянуто особливості накопичувачів даних, що застосовуються у різних сферах діяльності, зазначено специфічність накопичувача даних для ІАС ЖЦ. Ці властивості вимагають приділення великої уваги до якості проектування такого накопичувача та програмно-технічного комплексу інформаційно-аналітичної системи в цілому. Наступним

напрямок діяльності є розробка технічного забезпечення ІАС ЖЦ.

Література

1. Гир Д. Федеральный доступ к базам данных. – [Електр. ресурс]. – Режим доступу: <http://www.osp.ru>.

2. Стулов А. Особенности построения информационных хранилищ. – [Електр. ресурс]. – Режим доступу: <http://www.www.osp.ru>.

3. Inmon W.H. Data Warehouse Types. – [Електр. ресурс]. – Режим доступу: <http://www.dbazine.com>.

4. Финансовое хранилище данных или система главной книги. – [Електр.ресурс]. – Режим доступу: <http://www.olap.ru>.

5. Мирошников В.В., Голованов В.В. Многомерный анализ качества в системах управления качеством на основе применения аналитических информационных технологий // Информационные технологии. – 2005. – № 2. – С. 28 – 33.

6. Теплепнев О.В. Подійно-орієнтована інформаційно-аналітична система підтримки життєвого циклу складних виробів. Визначення уніфікованого набору параметрів події // Открытые информационные и компьютерные технологии: Сб. научн. тр. – Х.: Нац. аэрокосм. ун-т «ХАИ». – 2005. – Вып. 27. – С. 185 – 189.

Надійшла до редакції 15.09.2005

Рецензент: д-р техн. наук, проф. Г.Н. Жолткевич, Харківський національний університет ім. В.Н. Каразіна, м. Харків.