



UA0300443

ИСПОЛЬЗОВАНИЕ ХРАНИЛИЩ ДАННЫХ в СИСТЕМАХ ДОЗИМЕТРИЧЕСКОГО КОНТРОЛЯ

В. В. Бабенко, М. И. Бабенко, А. С. Казимиров
НПП «Атомкомплексприбор», Украина

Переход к новому этапу преобразования объекта «Укрытие» в безопасный объект и принятие новых норм радиационной безопасности НРБУ-97 ставит перед нами совершенно новые задачи.

Современные концептуальные положения по определению эффективной дозы и обоснованию допустимых уровней, новые требования безопасной работы и гарантированного неперевышения пределов дозы диктуют необходимость постоянного контроля дозового давления на персонал и комплексного подхода при анализе и расчётах полученных доз внутреннего и внешнего облучения; анализе, прогнозировании и планировании дозовых нагрузок и адекватном управлении облучаемостью персонала, оценке целесообразности и достаточности средств радиационной защиты, а также своевременного предупреждения критических ситуаций.

Исходными данными для решения перечисленных выше задач является информация, содержащаяся в разнообразных локальных оперативных системах дозиметрического контроля персонала и рабочих мест (СУБД-источниках), представляющих собой реляционные СУБД, настроенные прежде всего на оперативную обработку коротких и частых транзакций:

- контроль интегрального внутреннего облучения персонала объекта «Укрытие»;
- оперативный контроль ингаляционного внутреннего облучения за рабочую смену на объекте «Укрытие»;
- биофизические обследования персонала;
- индивидуальный контроль аэрозолей в зоне дыхания персонала (пробоотборников, респираторов);
- контроль внешнего облучения тела, хрусталика глаза, кожного покрова, кистей и стоп;
- контроль спецодежды;
- контроль поверхностного загрязнения рабочих мест и площадок;
- контроль воздушной среды рабочих зон объекта «Укрытие».

Решение этих задач лежит на пересечении множества результатов сложных перекрёстных запросов, обращённых к банкам данных СУБД-источников.

Непосредственное обращение к многочисленным и разнообразным базам данных (БД) источников, имеющих двумерное представление данных, в большинстве случаев невозможно из-за следующих причин.

1. Аналитические запросы «конкурируют» с оперативными транзакциями, блокируют данные, вызывая нехватку ресурсов, и не гарантируют достоверность используемой оперативной информации.

2. Структура оперативных БД (обычно третья нормальная форма) состоит из множества «узких» и сложным образом связанных таблиц, не понятных поэтому конечному пользователю. Такая структура не обеспечивает должной скорости выполнения аналитических запросов, не пригодна для комплексного анализа, включающего изучение количественных характеристик данных различных категорий.

3. В БД-источниках используются различные структуры данных, единицы измерения, способы кодирования и т. п.

4. СУБД-источники разработаны (или будут разработаны) на разных платформах аппаратного и программного обеспечения, а информация находится на разных компьютерах и в различных форматах.

Альтернативой непосредственному использованию разнородных источников данных для анализа и расчетов является организация хранилища данных и применение OLAP-технологии для проведения анализа данных и принятия решения.

Хранилища данных

Хранилище данных – это централизованный «склад» различных типов данных, собираемых из разных систем-источников, классифицированных по темам и по времени и организованных таким образом, чтобы их можно было легко и быстро проанализировать. Хранилище оснащено средствами транспортировки, хранения и обслуживания данных, проведения быстрого поиска и интерактивной аналитической обработки.

Хранилища данных отличаются следующие свойства.

- *Предметная ориентированность.* Информация в хранилище организована в соответствии с основными аспектами интересов пользователя. Предметная организация данных способствует как упрощению анализа, так и повышению скорости выполнения аналитических запросов.
- *Интегрированность.* При загрузке в хранилище данные должны быть проверены, очищены и приведены к единому виду. Анализировать такие интегрированные данные намного проще.
- *Привязка ко времени.* Данные, выбранные из оперативных БД, накапливаются в хранилище в виде «исторических слоёв», каждый из которых относится к конкретному периоду времени. Это способствует проведению анализа в разных «временных срезах» и отслеживанию тенденций изменения процессов.
- *Итоговые значения.* Промежуточные и окончательные итоговые значения могут быть заранее просчитаны, что значительно сокращает время анализа информации.

- *Доступ.* Обеспечивается простой, гибкий и быстрый доступ конечного пользователя к информации в интерактивном режиме.
- *Неизменяемость.* Попав в хранилище, данные «залегают» в «свой исторический слой» и уже никогда не меняются. Стабильность данных облегчает их анализ.

Структуры данных хранилища определяются предметной ориентированностью.

Данные в хранилищах организованы вокруг того или иного аспекта предметной области.

Наиболее адекватна такому представлению данных многомерная (пространственная) модель данных. Пространственная модель описывает данные о предметной области как n -мерный гиперкуб или n -мерную таблицу. В ячейках гиперкуба находятся количественные показатели (меры). Каждая ячейка описывается рядом атрибутов, образующих оси координат (измерения).

Одной из типичных структур, применяемых в хранилище данных, является схема звезды, которая является практически проекцией гиперкуба на «реляционную плоскость».

Схема звезды состоит из одной таблицы фактов и нескольких таблиц измерений.

Таблица фактов содержит по одной строке для каждого факта – минимально рассматриваемого атома анализируемого процесса. Полями таблицы фактов, помимо ключей, являются меры. Детальность таблицы фактов выбирается в зависимости от разных факторов.

Таблицы измерений содержат собственно значения измерений, т. е. информацию, характеризующую факты. Несколько особенным и практически необходимым измерением любого хранилища данных является время.

С точки зрения нормализации, таблицы измерений хранят избыточные данные. Но эта избыточность оправдывается двумя соображениями: во-первых, такая схема понятнее конечному пользователю; во-вторых, запросы по БД будут выполняться быстрее за счёт уменьшения количества таблиц, объединяемых в одном запросе.

Архитектура и компоненты хранилища данных

Создание, поддержка, управление и использование хранилища данных – это процесс, цель которого – непрерывная поставка нужной информации всем заинтересованным пользователям. Этот процесс подразумевает постоянное развитие, улучшение, решение всё новых задач и практически никогда не кончается.

Основные компоненты хранилища данных представлены на рисунке.

- Оперативные источники данных могут быть совершенно разными по способу хранения и организации, структуре и месту хранения. Очень часто – это БД разнородных реляционных СУБД.

- Служба переноса и трансформации данных осуществляет преобразование данных – набор операций, которым подвергаются исходные данные по мере их переноса в хранилище: выборка из разных источников, форматирование, фильтрация, интегрирование и загрузка в хранилище.
- Метаданные (данные о данных) – это родословная данных хранилища в целом и каждого элемента информации в отдельности: источник и время «происхождения», данные о модификации, структуре и трансформации, правила организации хранения и выборки и многое другое, записанное в доступной форме. Метаданные хранятся в репозитории отдельно от собственно данных и используются для эффективного взаимодействия различными службами хранилища.
- Реляционное хранилище – это «склад» очищенных и проинтегрированных данных из различных источников, уже доступных для анализа при помощи средств построения отчётов и для подготовки к OLAP-анализу.
- OLAP-хранилище – это «склад» подготовленных к OLAP-анализу данных, где они могут сохраняться в специальных БД многомерных структур – гиперкубах.
- Средства доступа и анализа данных – это программы формирования запросов и генерации отчётов или встроенные приложения для проведения оперативной аналитической обработки данных, например, по OLAP-технологии.

OLAP-технология

Особо следует отметить OLAP-технологии. Это системы интерактивной аналитической обработки данных хранилища. Они предоставляют удобные быстродействующие средства доступа, просмотра и анализа информации по всем аспектам.

OLAP-приложение позволяет:

- консервировать виды данных (обычные отчёты);
- проводить анализы данных сложными методами с помощью сводных таблиц. Сводные таблицы – это интерактивные виды, которые позволяют делать различные информационные «срезы», отфильтровывать данные по определенным признакам или заниматься подробным изучением информации на любом выбранном уровне;
- получать с помощью специальных запросов универсальные отчёты;
- проводить анализ типа «что, если...», предсказывать результаты, создавать прогнозирующие модели и проследживать тенденции изменений и пр.

При этом на всех этапах процесса направление поиска, формирование запроса или параметры анализа определяет непосредственно конечный пользователь в интерактивном режиме.

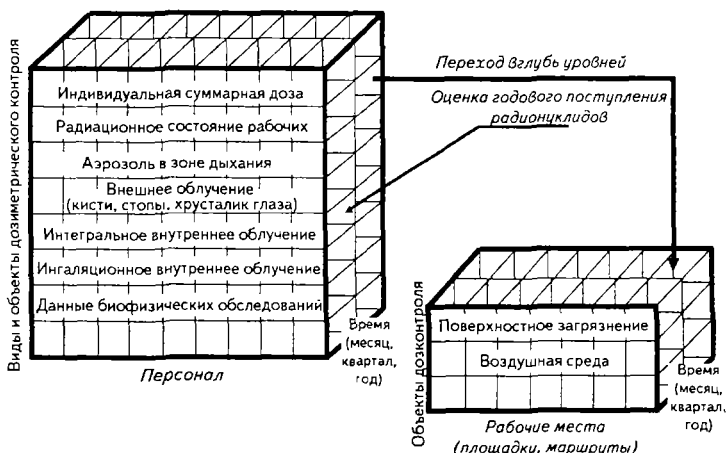
В результате пользователь получает естественную, интуитивно понятную модель данных, организованную в виде многомерной базы данных – гиперкубов. Гиперкуб, в свою очередь, состоит из первичных элементов-кубов.

Осями многомерной системы координат служат *измерения и факты*.

Измерения (время, географическое положение и др.) – это категории, относительно которых будет проводиться анализ данных. Измерения создаются таблицами и столбцами реляционной базы данных. Каждое измерение состоит из набора уровней и членов, которые дают дальнейшее определение данных.

Одним из таких наиважнейших атрибутов является время. С технической точки зрения привязка ко времени означает, что таблицы в явном виде имеют в своём составе «временный ключ» либо данные распределены по нескольким таблицам, каждая из которых относится к определённому временному периоду (году, кварталу и т.п.). Это позволяет получить «временные срезы», что помогает в проведении анализа и отслеживании тенденций в изменении процессов.

Факты (или меры, метрики, вычисления, размеры) – это количественные характеристики данных, которые нужно проанализировать. Факты тоже определяются таблицами и столбцами реляционной базы данных. Пример гиперкуба приведен на рисунке.



Гиперкуб хранилища постоянных данных по внутреннему и внешнему облучению персонала