



ТЕХНОЛОГИИ ИНТЕГРАЦИИ ИНФОРМАЦИОННЫХ СИСТЕМ И РАСПРЕДЕЛЕННЫХ БАЗ ДАННЫХ

Филатов В.А., Божинский И.А.

Харьковский национальный университет радиоэлектроники

Использование независимых систем автоматизации деятельности организации приводит к несогласованности и семантической неоднородности данных в различных подсистемах. Для эффективного управления современной организацией требуется интегрированная информационная система (ИС), позволяющая работать со всем объемом информации, накопленной в организации.

Задачу интеграции корпоративных приложений разделяют на два класса: интеграция данных и интеграция приложений. Интеграция на уровне данных является предпочтительным способом построения ИС, но она недостаточна в тех случаях, когда логика приложения неотделима от самих данных. Архитектура интегрирующей среды должна поддерживать такие модели интеграции, при которых зависимость между подсистемами минимальна.

Выбор метода интеграции существенно зависит от специфики интегрируемых приложений и доступных технических ресурсов. Основными моделями интеграции являются: передача сообщений, физическая интеграция данных (хранилища данных, репликация данных), логическая интеграция данных, мониторы транзакций и серверы приложений, автоматизация производственных бизнес-процессов. Выбор сценария интеграции, помимо функциональных требований прикладной системы, зависит от таких факторов: качество и актуальность данных, доступность исходного кода прикладных подсистем, качество технической документации, интенсивность работы с подсистемой, особенности сетевого доступа и др.

Под информационным пространством понимают совокупность баз данных (БД), технологий их сопровождения и поддержки, телекоммуникационных систем, обеспечивающих взаимодействие локальных ИС и удовлетворение их информационных потребностей. Доступ к такой информации усложнен с различиями в способах ее хранения, «границы» между источниками данных не только сдерживают использование информации для анализа, но и увеличивают расходы, связанные с процессом создания запросов и транзакций, пересекающих эти «границы».

Объединение ресурсов на основе информационно-коммуникационного взаимодействия ИС выводит их на уровень корпоративных информационных ресурсов под названием Единое Информационное Пространство (ЕИП). ЕИП включает в себя понятие единого пространства данных (ЕПД), которое реализует технологию доступа к удаленным БД, при этом ИС выступают в роли клиента и сервера, взаимодействуя друг с другом по определенному сценарию. В роли информационных ресурсов ИС выступают не только данные, но и различные приложения. В каждой из ИС часть методов обработки данных реализуется в виде модулей в приложении, доступных из других ИС. Такой подход соответствует одноранговой архитектуре взаимодействия ИС и минимизирует дублирование приложений. Распределение приложений по различным



Секция 1. Информационные системы и технологии: опыт создания, модели, инструменты, проблемы

ИС позволяет добиться оптимального баланса загрузки приложений и аппаратных средств и приводит к эффективному использованию информационных ресурсов систем в целом. При этом знание схемы БД необходимо только тому приложению, которое обрабатывает данные из этой БД.

Технология БД относится к числу основных компьютерных технологий и представляет собой совокупность методов и средств определения и манипулирования данными, интегрированными в БД. Важной целью применения технологии БД является создание информационного ресурса, разделяемого между функционально связанными приложениями, с обеспечением независимости внешнего представления БД от способов ее внутренней организации. Для этого используется соответствующий набор технологических инструментов.

В основу технологии БД положено применение реляционной модели данных (РМД), основанной на строгом аппарате реляционной алгебры и математической логики. Технологические операции определения и манипулирования БД выполняются с использованием систем реляционного исчисления. Широкое применение РМД позволило разрешить одну из серьезных проблем достижения модельной однородности БД, создаваемых в средах различных систем управления базами данных (СУБД). В то же время опора на реляционную модель существенно ограничивает возможности определения данных в БД и тем самым предопределяет соответствующие границы применения всей технологии БД.

Интеграционные тенденции формируют проблему построения интегрированных распределенных баз данных (ИРБД), для которых обеспечение схемной однородности на основе РМД оказывается недостаточной. При определении и построении ИРБД реляционный подход должен применяться с учетом классической схемы проектирования БД (необходимо знать, каким образом был выполнен полный цикл этапов моделирования заданной предметной областью в виде реляционных схем интегрируемых БД). Расширение границ применения реляционного подхода позволяет рассматривать схемы БД с учетом возможности их будущей интеграции в ИРБД.

В процессе разработки и эксплуатации ИС часто возникает необходимость интеграции данных из различных источников. При этом выделяют типичные задачи: согласование метамodelей, интерпретация моделей данных, сопоставление классификаторов, справочников и доменов, объединение информации.

Рассматриваются виды интеграции с точки зрения программной поддержки: процедурный, декларативный, смешанный. Кроме этого рассматриваются методы организация доступа к данным с позиции их размещения в файлах БД: консолидация данных, распределение данных, федерализация данных.

Для обеспечения взаимодействия ИС и распределенных БД в рамках ЕИП целесообразно применять распределенные одноранговые архитектуры с промежуточным программным уровнем интеграции. Такой подход позволит достичь требуемого уровня гибкости, открытости и производительности распределенных ИС.