

КОМПЛЕКСНЫЕ СТРУКТУРЫ МНОГОМЕРНЫХ БАЗ ДАННЫХ

САЕНКО В.И., АЛЬРАВАДЖБА М.

Предложена обобщенная структура многомерных баз данных, названная комплексной структурой, позволяющая повысить эффективность функционирования распределенной системы. Приводятся результаты анализа ее достоинств и недостатков.

1. Описание проблемы

Структура и организация баз оказывает значительное влияние на функционирование информационной системы. Особенно ощутимо это проявляется в распределенной системе и системе реального времени.

Как показано в [1], наиболее перспективной технологией, используемой при построении сложной информационной системы, являются многомерные базы данных (MDDB), представляющие собой обобщение обычных реляционных баз данных \mathbf{Bn} (NRDB). Существует несколько структур баз данных типа MDDB: многомерные гиперкубы \mathbf{Bc} (HCDB), плотные базы с радиальной структурой \mathbf{Bd} (DRDB-DIS - dense relation database with Data in Star), многомерные массивы (MWA - multiway array), базы данных типа хранилища данных \mathbf{Bw} (DWDB - Data Warehouse database). Все они имеют различные представления и способы организации. Основной вопрос - как выбрать для конкретной системы самую лучшую структуру, учитывая, что каждая из баз является оптимальной только по определенному критерию.

2. Многомерные базы данных (MDDB)

HCDB основана на представлении сложных информационных объектов в виде абстрактной многомерной дискретной структуры (гиперкуба), задаваемой как $\mathbf{Bc} = \{b_{ci}\}$, $b_{ci} = b_{ci}(x_1, x_2, x_3, \dots, x_n)$, b_{ci} - элемент базы данных; x_1, x_2, \dots, x_n - координаты гиперкуба. \mathbf{Bc} в общем случае совпадает с гипотетической базой данных, эквивалентной дискретному множеству, образованному декартовым произведением доменов $D_1 \times D_2 \times \dots \times D_k$, где домен D_i есть множество, совпадающее с множеством различных дискретных элементов, образованное выборочным атрибутом A_i , используемым в какой-либо реляционной (NRDB) базе данных системы \mathbf{Bn} . В качестве ортообразующих переменных гиперкуба выступают атрибуты-ключи основных баз. Гиперкуб определен для многомерно-го дискретного метрического пространства с ортогональной метрикой: $\|\mathbf{x}\| = |x_{11} - x_{1j}| + |x_{21} - x_{2j}| + \dots + |x_{n1} - x_{nj}|$.

Каждый элемент b_{mi} соответствует определенному кортежу исходных NRDB. Если соответствующего кортежа не существует, элемент равен 0, если кортеж существует, то элементу присваивается ненулевое значение. Предлагается два варианта представления этих значений: первый - элементу присваивается значение, равное 1 ($b_{ci} \in \mathbf{Z}$, $\mathbf{Z} = \{0, 1\}$), второй вариант - элементу присваивается некоторое числовое значение, равное $k_1 k_2 k_3 k_4 k_5$, где k_i - номер уровня i -го атрибута. Во втором варианте для каждой переменной хранится ключ разрядности, соответствующий

ший числу уровней, которые переменная принимает в NRDB. Запросы формируются в терминах отношений значений координат гиперкуба $b_{c1} q_1 b_{c2} q_2 \dots q_n b_{cn}$, как выбор множества ненулевых элементов при заданных m -местных отношениях на множестве атрибутов. База данных будет задана в виде множества ненулевых значений дискретной функции. Дискретное представление допускается использовать для любых типов данных, так как в реальной базе данных все имеющиеся значения для числовых атрибутов можно заменить дискретным набором этих значений путем соответствующего нормирования.

Недостаток представления в виде HCDB - высокая размерность при явном хранении всех элементов. Достоинство - такое представление обеспечивает контроль целостности информации при построении распределенной базы.

DRDB-DIS предполагает, что для заданной разреженной NRDB, со схемой (A_1, A_2, \dots, A_m) , формируются DRDB \mathbf{Bd}_i меньшей арности, со схемами $(E_{1i}, E_{2i}, \dots, E_{mi})$, где $mi < m$. $E_{pi} \in \mathbf{A}$, $\mathbf{A} = \{A_1, A_2, \dots, A_m\}$. Исходные запросы $A_1 q_1 A_2 q_2 \dots q_m A_m$ представляются в виде $E_{ji} q_1 E_{1k} q_2 \dots q_r E_{rp}$. Так как $A_1 = E_{ji}$, $A_2 = E_{1k}, \dots$, то они остаются теми же самыми, а атрибуты нормальной базы данных NRDB заменяются атрибутами плотных баз.

Недостаток - наличие дополнительных процедур связи баз данных. Достоинство - компактность хранения, удобство при построении распределенной структуры базы.

MWA основан на формировании многомерно-го массива и может рассматриваться как альтернатива DRDB-DIS.

DWDB будем рассматривать как расширение NRDB. В системе вводится новое представление $\mathbf{Bw} = F(\{\mathbf{Bn}\}, p)$, $\{\mathbf{Bn}\}$ - множество \mathbf{Bn} , p - аналоговый или дискретный параметр, т.е. $p = \langle \text{время} \rangle, \langle \text{регион} \rangle, \langle \text{класс} \rangle, \langle \text{группа} \rangle, \dots$ - (в частном случае $p = t$, где t - время). b_{wi} - элемент DWDB зависит от нового параметра, т.е. $b_{wi} = b_{wi}(p)$, DWDB не меняет схему основной базы, а добавляет новую переменную, эквивалентную атрибуту. Запросы для DWDB делаются по этой новой переменной (например, найти такие значения p , при которых выполняются заданные отношения между атрибутами, $A_{ij} q_1 a^*$). Наборы реляционных баз в DWDB $\{\mathbf{Bn}_i\}$ образуют кубы пространства $2, 5D$ ($b_{wi} \in 2, 5D$, где $2, 5D$ - дискретное пространство размерностью 2.5), являются частным случаем MDDB. Хранение информации может осуществляться в виде сечений, т.е. реляционных таблиц NRDB $\mathbf{Bn} = \{b_{ni}\}$.

3. Комплексные структуры

Считаем, что функциональность системы \mathbf{S} задается набором функций $\mathbf{F} = \{F_1, F_2, \dots, F_n\}$, которые она обеспечивает (например, вывод информации в соответствии с заданным форматом, поиск информации по заданному правилу, редактирование, доставка информации удаленному пользователю, разрешение конфликта одновременного доступа пользователей). Среди них будем выделять активные функции $\{F_{ai}\}$, т.е. функции, которые должны решаться в текущий момент времени, и отложенные $\{F_{ri}\}$. Все они делятся на специальные $\{F_i^s\}$, т.е. отражающие локальные цели, и общие $\{F_i^c\}$, характеризующие и обеспечивающие работу всей системы. Будем полагать, что базы

данных должны обеспечивать все функции системы. Таким образом, объект задается набором функций $F = \{ \{F_a^{s_i}\}, \{F_r^{s_i}\}, \{F_a^{c_i}\}, \{F_r^{c_i}\} \}$. Степень обеспечения активных функций $\{ \{F_a^{s_i}\}, \{F_r^{s_i}\} \}$ отражает текущее состояние системы, степень обеспечения отложенных функций $\{ \{F_a^{c_i}\}, \{F_r^{c_i}\} \}$ совместно с обеспечением активных функций отражает общую функциональность системы (рис.1).



Рис.1. Функциональная структура

Оптимизационные задачи в распределенной системе относятся к многокритериальным, поэтому их решение требует компромиссных подходов. Как было показано, каждый из типов баз данных оптимален только по некоторым критериям. При работе с базами данных возникает несколько противоречий, связанных с различным способом их представления. Базы данных HCDB хорошо отражают целостность системы, но имеют громадные объемы, увеличивают время доступа к данным и неэкономны с точки зрения заполнения информацией. Базы данных DRDB имеют хорошие показатели плотности хранения информации, но плохо согласуются с контролем целостности и временем реализации запросов. Базы HCDB, DRDB вообще не позволяют осуществить анализ данных по шкале времени.

Проблема состоит в обеспечении решения всех функций с учетом критериев и ограничений системы. Для ее решения предлагается описывать объект несколькими базами данных, являющимися отображением одних и тех же данных, каждая из которых удовлетворяет какому-либо критерию. Все базы представляют связанную структуру, будем называть ее комплексной структурой CSDB.

Технология представления комплексной структуры (CSDB) сводится к нескольким этапам. Разреженная физическая NRDB B_n заменяется на плотные DRDB-DIS B_d . Поддержка реализации общих функций F^c (независимо, активные они или отложенные) осуществляется с помощью либо NRDB, либо DRDB. Для описания характеристик всей системы формируется одна или несколько логических баз HCDB, а для активных специальных функциональных задач формируется база данных выходных форм. Выходным формам ставятся в соответствие множе-

ства запросов, под которые в свою очередь формируются виртуальные временные базы (VRDB) B_v . Виртуальные реляционные базы $B_v = \{b_{v_i}\}$ используются для оперативной работы, они являются подмножеством главной NRDB или DRDB. VRDB создаются в соответствии с существующими наборами формул, определяющих унарные, бинарные и m-местные отношения на кортежах физических баз данных и/или компонентах NRDB:

$$\{ (s_i[k] \ q_2 s_j[l]), (s_i[k] \ q_2 a), (s_i[k] \ q_1 s_i[l]), (s_i[k] \ q_m s_j[l] \ q_m s_t[r] \dots q_m s_f[h]) \},$$

где $s_i[k]$ – k – компонент s_i кортежа; q_1, q_2, q_m – соответственно унарное, бинарное и m-местное отношения; a – константа; q_m – заданный тип отношения между элементами множества. Для обеспечения функций статического многомерного или динамического (on-line) межбазового анализа данных создаются базы данных с новыми размерностями – DWDB. Эти базы не являются временными и потребляют значительные ресурсы всей системы.

Образование новых баз B_v осуществляется в соответствии с требованиями обеспечения новых функций и m-местных отношений на имеющемся множестве элементов MDDB. Совокупность этих баз наиболее полно характеризует информационный объект.

Комплексная структура, таким образом, представляется несколькими связанными описаниями: плотными базами данных DRDB $\{B_d\} \in B_m$, логической структурой $B_c = \{b_{c_i}\}$ (полное описание гиперкуба HCDB), набором виртуальных баз $\{B_v\} \in B_m$, расширяемой базой DWDB $B_w = F(\{B_d\}, p)$:

$$T = \{DRDB, HCDB, DWDB, VRDB\} = \{ \{B_d\}, B_c, B_w, \{B_v\} \}.$$

Она является целостной структурой, содержащей избыточное описание исходного набора данных. Все базы связываются логически и синхронизируются между собой. Размещение баз осуществляется в рамках существующей распределенной транспортной системы.

4. Преимущества представления комплексными структурами

Преимущества ощущаются в распределенных системах реального времени. Для распределенной системы выделим ряд наиболее существенных характеристик: целостность базы, распределенность, время доступа к базе данных, достоверность данных, расходуемые ресурсы на хранение данных, живучесть (надежность).

Целостность обеспечивается процедурами контроля с помощью баз HCDB. Всю систему характеризует множество ненулевых элементов $\{b_i\}$. HDDB также обеспечивает удобную форму кодирования элементов базы, введение кодов проверки и защиты данных. Целостность сводится к сохранению существующего набора ненулевых элементов, позволяя восстанавливать исходные базы в случае разрушения отдельных участков данных.

Распределенность данных понимается как распределенность виртуальных баз VRDB, плотных (упакованных) баз DRDB и архивов B_a . Архивы B_a могут быть включены в состав структуры T. Комплексная структура дает некоторый выигрыш, так как

распределению подлежат только DRDB. Виртуальные базы могут генерироваться в соответствии со схемой распределения приложений клиента, существующим графиком и существующей схемой размещения серверов. Распределенность указанных баз рассматривается как послойная распределенность, обеспечивающая заданный коэффициент избыточности системы.

Время доступа к данным в системе с комплексной структурой имеет лучшие значения, поскольку для оперативной работы используются виртуальные базы, хотя это вносит отдельный элемент риска потери информации в случае выхода из строя системы. Для предотвращения такой потери целесообразно ввести специальный механизм синхронизации всех баз. Синхронизация баз данных основана на том, что изменение в базе сопровождается формированием специального сообщения-подтверждения и автоматической коррекцией всех баз. При этом в системе устанавливается допустимое время синхронизации, т.е. минимальное время запаздывания внесения изменений. Таким образом, изменения вносятся сначала в виртуальную базу и только спустя некоторое время они внесутся во все базы. После внесения изменений будут переданы отчеты: локальный – для вносящего изменения и общий – для администратора баз.

Достоверность рассматривается с двух точек зрения: имеющихся и новых данных. Достоверность имеющихся данных обеспечивается наличием послойно распределенных структур. Сложности в обеспечении достоверности новых данных имеют место в связи с тем, что возникает неопределенность, что считать истиной при внесении ложного изменения в данные одной из баз. Для разрешения этого противоречия вводятся приоритетное занесение данных (только для определенного ранга пользователя) и подтверждение, т.е. выдача статистики после занесения всех изменений пользователю. Данные считаются истинными при совпадении во всех копиях. В этой ситуации случайное изменение не приводит к ошибкам в базах.

Живучесть или устойчивость к разрушению обеспечивается послойным распределением баз, обеспечивающих заданный коэффициент избыточности. Чем выше значение этого коэффициента, тем больше устойчивость системы к разрушению данных.

Расход ресурсов при использовании комплексной структуры значительно выше, чем при обычном использовании баз. Особенно это проявляется по отношению к ресурсам сервера. Учитывая значительное падение цен на компьютерную технику, в целом это приводит к повышению эффективности системы.

Основные выводы. Комплексная структура позволяет, во-первых, осуществлять разбиение систе-

мы данных для обеспечения требований распределенности, во-вторых – реализовать запросы, соответствующие любому допустимому n -местному отношению на множестве элементов MDDB, в-третьих – выполнять оптимизацию по нескольким критериям, в-четвертых – поддерживать компактную форму хранения информации за счет плотных баз данных, в-пятых – обеспечивать содержательность пользовательского интерфейса и высокие скорости доступа за счет виртуальных баз. Комплексная структура и связь ее с основными параметрами, характеризующими качество информационной системы, представлена на рис. 2.

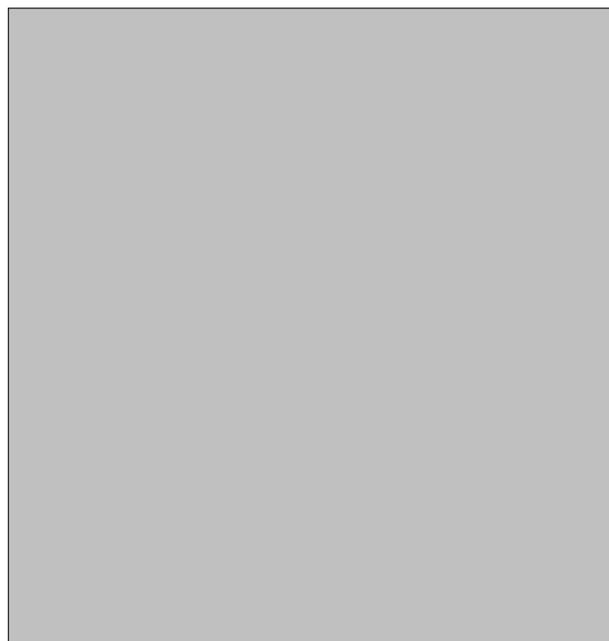


Рис. 2. Схема взаимосвязей структуры

Литература: 1. Саенко В.И., Клименко А.В., Альраваджба М. Использование новых технологий в информационных корпоративных системах // Радиотехника и информатика. 1997. №1. С. 111-114.

Поступила в редколлегия 27.03.98

Саенко Владимир Иванович, канд. техн. наук, доцент кафедры информационных управляющих систем ХТУ-РЭ. Научные интересы: администрирование, мониторинг и управление процессами в компьютерных сетях. Увлечение: филателия. Адрес: 310726, Украина, Харьков, пр.Ленина, 14, тел. 409-451.

Альраваджба М., аспирант кафедры информационных управляющих систем ХТУРЭ. Научные интересы: управление процессами в компьютерных сетях, сетевые технологии. Увлечение: туризм. Адрес: 310726, Украина, Харьков, пр.Ленина, 14, тел. 409-451.