

УДК 681.3

Определение характеристик распределенной базы данных с помощью имитационного моделирования

Е.И.СУКАЧ, О.И.ЕСЬКОВА, Т.Я.КАМОРНИКОВА

Распределенные базы данных (РБД) отличает функциональная и структурная сложность, а процесс их проектирования характеризуется большой длительностью, высокой трудоемкостью и значительными финансовыми затратами. Проектирование и реализация РБД предполагают оперативный доступ пользователей ко всему объему информации, хранящейся в узлах компьютерной сети. Эффективный доступ пользователей в значительной степени определяется организацией РБД [1].

Проектирование РБД требует применения математических методов исследования. С этой целью разработан ряд аналитических методов исследования [2,3], которые частично решают задачи размещения баз данных с целью минимизации затрат на их содержание и взаимодействие с пользователем. Однако они не позволяют наблюдать за объектом имитации в динамике, а также определять узкие места функционирования РБД. Это свидетельствует в пользу имитационных методов исследования.

В качестве инструмента построения ИМ предлагается использовать СМ МІСІС [4]. Она позволяет создавать полностью параметризованные ИМ на высоком уровне детализации, которые могут использоваться для решения разнообразных задач. Благодаря диалоговой технологической оболочке СМ МІСІС исследователи могут оперативным образом вмешиваться в процессы имитации, контролировать и корректировать их реализацию на ПЭВМ.

Ниже описывается имитационная модель (ИМ) РБД, основанная на взаимодействии множества пользователей с информационными ресурсами, размещенными на компьютерах объединенных в сеть. Рассматривается однородная распределенная база данных, которая представляет собой набор файлов, хранящихся в разных узлах информационной сети и логически связанных таким образом, чтобы составлять единую совокупность данных. Согласно концепции "клиент-сервер" предполагается выделение в узле процессов двух типов: клиентов, потребляющих тот или иной вычислительный ресурс, и серверов, которые выполняют требования клиентов, занимаясь управлением разделяемых ресурсов, полностью инкапсулируя их внутри себя.

ИМ РБД включает следующие компоненты:

- узлы обработки информации, которые имеют определенный информационный ресурс (файлы РБД) и обеспечивают выполнение различных запросов пользователей;
- пункты потребителей информационных ресурсов распределенной БД;
- сеть передачи данных.

Узлами обработки информации являются ЭВМ с установленными серверами БД, которые обеспечивают выполнение как собственных, так и удаленных запросов пользователей.

Пункты потребителей информационных ресурсов – это набор терминалов в каждом узле сети с установленным программным обеспечением клиента, с которых поступают запросы в различные узлы РБД.

Сеть передачи данных предназначена для оперативной транспортировки данных в любую часть РБД. Для имитации выбрана многосвязная топология, часто используемая в сетях передачи данных общего пользования. Следует отметить, что ИМ параметризована и структура ее может быть установлена по желанию пользователя. Для этого необходимо изменить значения элементов массива С, определяющего состав узлов РБД и установленные связи между ними.

Число узлов РБД задается параметром N . На каждом узле РБД размещается M независимых файлов, которые будем обозначать F_{ij} , где i -номер узла, j -номер файла. Файлы характеризуются объемом V , который в процессе имитации динамически изменяется – увеличивается при пополнении его информацией и уменьшается при удалении из него информации.

Сгенерированный в узле запрос характеризуется следующим набором параметров:

- номером узла-отправителя;
- номером узла-адресата;
- номером файла РБД, к которому он обращается;
- типом;
- длиной;
- объемом информации для записи в файл или удаления из файла РБД.

В сети передаются сообщения трех типов: запросы на чтение (тип 1), запросы на запись (тип 2), запросы на удаление (тип 3). Количество типов запросов ИМ может быть изменено в процессе имитации. Длина запроса каждого типа определяется некоторыми законами распределения с заданными параметрами, вид которых может быть изменен по желанию исследователя при решении с помощью ИМ конкретной задачи.

Параметр "объем информации для записи в файл или удаления из файла РБД" для запросов типа 1 равен 0. Для запросов типа 2 и 3 он динамически изменяется в соответствии с заданным законом распределения и задает объем информации, на который соответственно будет увеличен или уменьшен соответствующий файл РБД.

Запросы к РБД поступают в соответствии с экспоненциальным законом распределения с заданной частотой L_{kt} , где k -номер узла-отправителя, t -тип генерируемого запроса.

Все запросы, поступающие в любой узел РБД, предполагают доступ к определенному файлу узла. Планирование запросов различных типов по узлам сети определяется для каждого узла-отправителя массивами вероятностей P_{ij} , где t -тип запроса, i -номер узла-адресата, j -номер файла в узле-адресате.

В ИМ отражается механизм блокировки файлов РБД. Запросы на чтение (тип 1), поступающие к некоторому файлу F_{ij} , обрабатываются одновременно. Для этого на обслуживающем приборе, имитирующем узел обработки информации, вводится параметр k , определяющий число каналов. Под числом каналов понимается число одновременно обслуживаемых заявок на приборе. В том случае, если к файлу F_{ij} пришел запрос на запись либо удаление (тип 2 или 3), то число каналов становится равным 1 и все остальные запросы к этому файлу задерживаются в очереди до момента окончания обслуживания запроса типов 2 или 3.

Время обслуживания запросов в узлах обработки информации зависит от типа запроса. Предполагается, что запросы каждого типа выполняются в узле РБД за время, подчиняющееся равномерному закону распределения с заданными параметрами. Параметры закона задаются исследователем для каждого типа запроса и могут быть изменены в ходе имитационного эксперимента.

В процессе имитации в зависимости от изменения объема файлов в узлах РБД динамически изменяются параметры закона распределения времени обслуживания запросов различных типов к соответствующим файлам.

В рассматриваемой системе РБД предполагается использование фиксированной маршрутизации, допускающей несколько возможных путей передачи запросов между узлами сети. Маршрутизация задается массивом T размерности $N \times N \times (N+1)$. При изменении конфигурации сети, которая задается массивом C , происходит корректировка и элементов массива

Время передачи информации по линиям сети передачи данных определяется матрицей размерности $N \times N$.

Ответы на запросы, сформированные в узлах обработки информации, имеют следующие параметры:

- номер узла-отправителя;
- номер узла-адресата;

- тип;
- длину;
- объем информации прочитанной из файла РБД.

Тип ответов устанавливается по типу соответствующего запроса. Длина ответов каждого типа определяется также некоторыми законами распределения с заданными параметрами, вид которых может быть изменен по желанию исследователя.

Параметр "объем информации прочитанной из файла РБД" для ответов типа 2 (сообщение о записи) и типа 3 (сообщение об удалении) равен 0. Для ответов типа 1 он динамически изменяется в соответствии с заданным законом распределения и задает объем информации, который будет передан в узел РБД, пославший запрос на чтение.

Ответы передаются по сети в обратном направлении по тому же маршруту, что и соответствующие запросы.

Как видно из описания, ИМ РБД полностью параметризована и может быть использована для решения следующих задач:

- определения допустимых диапазонов интенсивности запросов;
- определения времени обслуживания различных типов запросов в сети;
- определения границ изменения пропускной способности сети при заданной ее конфигурации для различных составов типов запросов;
- выбора оптимального варианта размещения информационных ресурсов по узлам сети.

Abstract

The model of the database with distributed information resources is considered. The problems of organization of simulation experiments for the given model are surveyed.

Литература

1. Дж. Мартин, Вычислительные сети и распределенная обработка данных, М., 1985.
2. П.И.Иваненко, Т.И.Колмар, Выбор топологии сети и размещение файлов данных в распределенной информационной системе, Упр. системы и машины, № 5 (1993), С.64-70.
3. А.В.Максименков, М.Л.Селезнев, Основы проектирования информационно-вычислительных систем и сетей ЭВМ, М.: Радио и связь, 1991.
4. I.V.Maximey, V.D.Levchyk, E.I.Sukach, Programme technological Complex of Simulation Modeling Proceeding. Intern. AMSE Conference "Applied Modeling & Simulation", Lviv (Ukraine), Sept. 30-Oct. 2, 1993, AMSE press, PP.75-83.

Поступило 22.05.2000