

Разработка параллельной СУБД в оперативной памяти* для кластерных систем

Е.В. Аксенова

В настоящее время одним из наиболее эффективных средств обработки сверхбольших баз данных являются параллельные СУБД, реализованные на аппаратной платформе кластерных вычислительных систем. Параллельная СУБД использует принцип фрагментного параллелизма, который заключается в следующем. Каждое отношение (таблица) базы данных разбивается на горизонтальные фрагменты, распределяемые по процессорным узлам кластера. Запрос к базе данных выполняется в виде нескольких параллельных процессов (агентов), каждый из которых обрабатывает свой фрагмент отношения. Полученные фрагменты сливаются в результирующее отношение.

В настоящее время в области СУБД в оперативной памяти ведутся интенсивные научные исследования и коммерческие разработки. В силу того, что оперативная память современных компьютеров позволяет разместить базу данных размером в десятки гигабайт, СУБД в оперативной памяти потенциально способны обрабатывать запросы существенно быстрее, чем традиционные дисковые СУБД. В данной работе представлены начальные результаты разработки параллельной СУБД в оперативной памяти, создаваемой на основе прототипа параллельной СУБД Омега для многопроцессорных вычислительных систем с кластерной архитектурой [<http://omega.susu.ru>].

Основные различия между СУБД в оперативной памяти и дисковыми СУБД заключаются в алгоритмах оптимизации запросов, управлении буферным пулом, организации индексных структур. Одной из основных задач при разработке параллельной СУБД в оперативной памяти является организация экономичного хранения данных в оперативной памяти и создание эффективных алгоритмов управления этими данными.

Предлагается техника на основе использования указателей и специальных доменных таблиц, представленная на Рис. 1. Для хранения значений каждого атрибута реляционного отношения используется отдельная таблица в оперативной памяти. В этой таблице вместо значений атрибутов хранятся указатели на системную таблицу соответствующего домена. Таблица домена представляет собой хэш-таблицу, в которой записаны все имеющиеся на данный момент значения на данном домене. Данная организация хранения позволяет экономично использовать оперативную память при работе с большими базами данных.

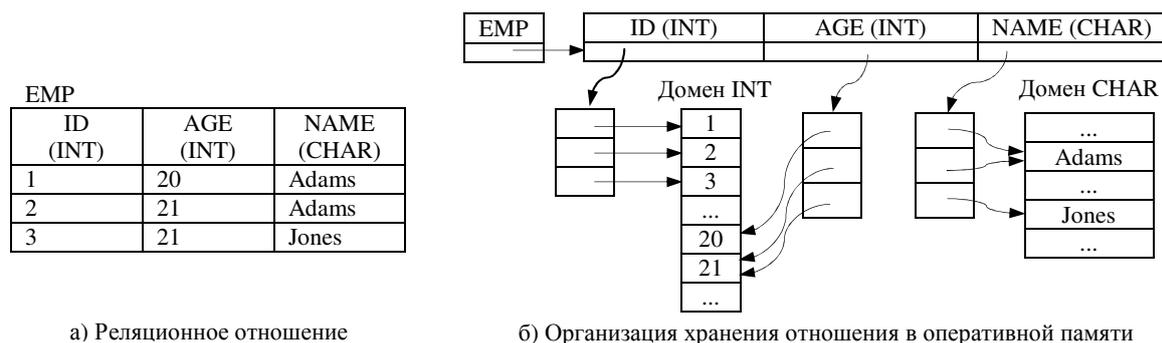


Рис. 1. Организация хранения данных в параллельной СУБД в оперативной памяти

Параллельная СУБД в оперативной памяти включает в себя Менеджер памяти, отсутствующий в дисковой СУБД Омега. Менеджер памяти реализует следующие основные функции: загрузка фрагментов отношений базы данных в оперативную память и превентивное периодическое сохранение данных на диск. Изменена реализация Менеджера файлов дисковой СУБД Омега, который обеспечивает функции чтения-записи кортежей.

*Работа выполнена при финансовой поддержке Российского фонда фундаментальных исследований (проект 09-07-00241-а).