

Содержание

Лабораторная работа № 3 Фрагментация базы данных	2
Цель задания	2
Краткие теоретические сведения.....	2
Горизонтальная фрагментация	2
Вертикальная фрагментация.....	2
Фрагментация при распределенных запросах	3
Задание	5
Контрольные вопросы по лабораторной работе №3:	5
Список дополнительной литературы	6

Лабораторная работа № 3

Фрагментация базы данных

Цель задания

Изучение методов фрагментации базы данных, распределенной на нескольких компьютерах сети.

Краткие теоретические сведения

Таблицы базы данных могут быть фрагментированы, то есть, разбиты на части, и эти части могут храниться в разных узлах распределенной базы данных. Основанием для фрагментации является повышение производительности. Части таблицы хранятся в тех местах, где к ним происходят наиболее частые обращения - это уменьшает сетевой трафик и сокращает время доступа к данным .

Фрагментация может быть горизонтальной или/и вертикальной.

Горизонтальная фрагментация

Горизонтальная фрагментация - разбиение таблицы по строкам. В этом случае значение в некотором столбце (комбинация значений в столбцах) рассматривается как ключ фрагмента, однозначно определяющий фрагмент, в который входит строка таблицы. Так, например, если таблица PRODUCT представляет список товаров, то данные о каждом товаре могут храниться в узле того отдела, который занимается продажей этого товара. Очевидно, что ключом фрагмента в этом случае должен быть код отдела. Поскольку в базе данных узла фрагмент определяется как таблица, в определение таблицы может (и должно) вводиться ограничение на значение ключа фрагмента.

Вертикальная фрагментация

Вертикальная фрагментация - разбиение таблицы по столбцам. Так, характеристики сортности товара могут храниться в фрагменте на узле

отдела качества, столбцы, определяющие стоимость - в фрагменте на узле отдела продаж и т.д. Вертикальная фрагментация, по сути, представляет собой декомпозицию в схеме данных, когда одна таблица разбивается на две или более таблиц, связанных друг с другом отношением 1:1. Декомпозиция должна быть выполнена без потерь, то есть, набор значений первичных ключей во всех таблицах должен полностью совпадать.

Независимость от фрагментации состоит в том, что ни один из фрагментов не является производным от других фрагментов. Восстановление полной таблицы из фрагментов производится операциями объединения (горизонтальная фрагментация) или/и естественного соединения (вертикальная фрагментация) таблиц-фрагментов. Полная таблица может быть описана как представление, определяемое через указанные операции. При обновлении представления полной таблицы могут возникать проблемы, типичные для обновления представлений: необходимость определения того физического фрагмента, к которому относится обновление. Так, в приведенном выше примере горизонтальной фрагментации перевод товара из одной категории в другую (изменение значения в столбце кода категории) потребует перенесения записи из одного фрагмента в другой.

Фрагментация при распределенных запросах

В распределенных запросах оптимизация еще более важна, чем в локальных. Поскольку выполнение распределенных запросов включает в себя обращение к удаленным узлам и пересылку данных между узлами, минимизация числа таких обращений и объема пересылаемых данных может во много раз уменьшить время выполнения запроса. Выполнение распределенных запросов включает в себя этап глобальной оптимизации, на котором оптимизатор решает, какие данные и с какого узла на какой будут пересылаться, и локальной оптимизации на каждом участвующем в запросе узле.

Например, если на узле А имеется таблица ТА, содержащая 100 строк, а на узле В имеется таблица ТВ, содержащая 106 строк, и запрос требует соединения этих таблиц, то очевидно, что выгоднее переслать таблицу ТА на узел В и выполнить соединение на узле В, а не наоборот.

Другой пример. Выше мы рассмотрели пример таблицы PRODUCT (список товаров), фрагментированный горизонтально по продукции определенного вида. Если в описаниях ее фрагментов указано ограничение на код продукта для фрагмента, и оптимизатор "знает" об этом ограничении, то запрос:

```
SELECT * FROM product WHERE prod_cod = 'milk'
```

оптимизатор может локализовать на единственном узле - на том, для которого ограничение ключа фрагмента совпадает с условием, заданном в запросе.

Рассмотрим пример, иллюстрирующий оба типа фрагментации. Имеется таблица PROVIDER (prov_id, prov_name, prov_phone и т.д.), определенная в базе данных на узле WS1. Имеется точно такая же таблица, определенная в базе данных на узле в WS2. Обе таблицы хранят информацию о поставщиках товара. Кроме того, в базе данных на узле в WS3 определена таблица PRODUCT (prod_id, prod_price). Тогда запрос "получить информацию об организациях-поставщиках" может быть сформулирован так:

```
SELECT *  
FROM provider@ws1.salesment, provider@ws2.salesment,  
ORDER BY prov_id;
```

В то же время запрос "получить информацию о стоимости поставленных продуктов" будет выглядеть следующим образом:

```
SELECT product.prod_id, product.prod_price,  
       provider.prov_name  
FROM provider@ws1.salesment, provider@ws2.salesment,  
     product@ws3.salesment  
ORDER BY prod_id;
```

Задание

1. На основе горизонтальной фрагментации:
 - а) создать представление, в которое попадут все поставщики продукции всех видов;
 - б) создать представление с указанием всех атрибутов, в которое попадут все потребители продукции всех видов;
 - в) создать представление, содержащее список всех продуктов всех производителей с указанием кода (артикула), наименования и стоимости;
2. На основе вертикальной фрагментации:
 - а) выделить столбцы, отвечающие за характеристику товара (например, сортность) в отдельную таблицу и разместить ее на рабочей станции WS1;
 - б) выделить столбцы (код, наименование, стоимость) в отдельную таблицу и разместить ее на рабочей станции WS2;
 - в) с рабочей станции WS3 выполнить распределенный запрос и получить представление из таблиц, полученных ранее (см. пункты а) и б) задания 2), содержащее информацию о товарах 1 сорта, стоимость которых отличается от средней не более, чем на 10%;
3. Оформить отчет о выполнении лабораторной работы.

Контрольные вопросы по лабораторной работе №3:

1. Дайте определения следующим понятиям: горизонтальная, вертикальная фрагментация.

2. Какой из видов фрагментации является наиболее эффективным в данной базе данных?
3. Для Вашей базы данных приведите примеры возможной горизонтальной и вертикальной фрагментации.

Список дополнительной литературы

1. Технологии и средства консолидации информации: *Учебное пособие. Деревянко А.С., Солощук М.Н. - Харьков: НТУ "ХПИ", 2008. - 432с.*
2. Организация баз данных. 1 часть: *Курс лекций / Е.В. Сопченко, К.А. Кудрин. Самарский гос. аэрокосмический ун-т. Самара, 2000, 71 с.*
3. Сергей Кузнецов. Базы данных. Вводный курс. *www.cityforum.ru*
4. Сергей Кузнецов. Основы современных баз данных. *www.cityforum.ru*
5. Дейт К.Д. Введение в системы баз данных, 6-е издание. -М: Вильямс. 1999 г. -848 с.
6. Бобровски С. Oracle 8. Архитектура. – М: Издательство «Лори», 1998, 210 с.
7. Методические рекомендации к выполнению лабораторных работ по дисциплине «Серверные системы управления базами данных» для студентов специальности 230102 «Автоматизированные системы обработки информации и управления» всех форм обучения /Сост.: М.В. Додонов, А.Ю. Павлов. –Самара: СамГУПС, 2007. – 16 стр.