Рассмотрение распределенных баз данных.

Основной причиной разработки систем, использующих базы данных, является стремление интегрировать все обрабатываемые в организации данные в единое Целое и обеспечить к ним контролируемый доступ. Хотя интеграция и предоставление контролируемого доступа могут способствовать централизации, последняя не является самоцелью. На практике создание компьютерных сетей приводит к децентрализации обработки данных. Децентрализованный подход отражает организационную структуру компании, логически состоящую из отдельных подразделений, отделов, проектных групп и тому подобного, которые физически распределены по разным офисам, отделениям, предприятиям или филиалам, причем каждая отдельная единица имеет дело с собственным набором обрабатываемых данных.

Разработка распределенных баз данных позволяет сделать данные, поддерживаемые каждым из существующих подразделений организации, общедоступными, обеспечив при этом их сохранение именно в тех местах, где они чаще всего используются. Подобный подход расширяет возможности совместного использования информации, одновременно повышая эффективность доступа к ней.

Распределенные системы призваны разрешить проблему островов информации. Базы данных иногда рассматривают как некие электронные острова, представляющие собой отдельные и, в общем случае, труднодоступные места, подобные удаленным друг от друга островам. Данное положение может являться следствием географической разобщенности, несовместимости используемой компьютерной архитектуры, несовместимости используемых коммутационных протоколов и т.д. Интеграция отдельных баз данных в одно логическое целое способна изменить подобное положение дел.

Распределенная база данных –набор логически связанных между собой разделяемых данных (и их описаний), которые физически распределены в некоторой компьютерной сети. Распределенная СУБД –программный комплекс, предназначенный для управления распределенными базами данных и позволяющий сделать распределение информации прозрачной для конечного пользователя.

Система управления распределенными базами данных (СУРБД) состоит из единой логической базы данных, разделенной на некоторое количество фрагментов. Каждый фрагмент базы данных сохраняется на одном или нескольких компьютерах, которые соединены между собой линиями связи и каждый из которых работает под управлением отдельной СУБД. Любой из сайтов способен независимо обрабатывать запросы пользователей, требующие доступа к локально сохраняемым данным (что создает определенную степень локальной автономии), а также способен обрабатывать данные, сохраняемые на других компьютерах сети.

Пользователи взаимодействуют с распределенной базой данных через приложения. Приложения могут быть классифицированы как те, которые не требуют доступа к данным на других сайтах (локальные приложения),и те, которые требуют подобного доступа(глобальные приложения).В распределенной СУБД должно существовать хотя бы одно глобальное приложение, поэтому любая СУРБД должна иметь следующие особенности:

1. Набор логически связанных разделяемых данных.
2. Сохраняемые данные разбиты на некоторое количество фрагментов.
3. Между фрагментами может быть организована репликация данных.
4. Фрагменты и их реплики распределены по различным сайтам.
5. Сайты связаны между собой сетевыми соединениями.
6. Работа с данными на каждом сайте управляется СУБД.
7. СУБД на каждом сайте способна поддерживать автономную работу локальных приложений.
8. СУБД каждого сайта поддерживает хотя бы одно глобальное приложение.

Из определения СУРБД следует, что для конечного пользователя распределение системы должна быть совершенно прозрачна(невидима). Другими словами, от пользователей должен быть полностью скрыт тот факт, что распределенная база данных состоит из нескольких фрагментов, которые могут размещаться на различных компьютерах и для которых, возможно, организована служба репликации данных.