Математические вопросы кибернетики и вычислительной техники 18, 1997, 65-71.

Исследование и реализация некоторых методов объединения гипертекстовых систем вычислительных сетей и реляционных баз данных

Д. А. Мовсесян

Институт проблем информатики и автоматизации
НАН РА и ЕрГУ

В базах данных разного типа зачастую возникают задачи, требующие их совместного использования. Эти задачи по получению информации из нескольких баз данных должны решаться так, чтобы сохранялась целостность существующих баз данных и их СУБД не подвергались каким-либо изменениям. Поэтому необходима разработка объединенной системы, которая бы поддерживала распределенные задачи, работающие в нескольких системах разного типа. При этом необходимо учесть, что системы-компоненты первоначально не были разработаны, чтобы поддерживать это взаимодействие, и пока не существует какой-то общей модели для совместного использования таких изолированных информационных систем.

К СУБД-компонентам объединенной системы, которые получили наиболее широкое распространение в настоящее время, можно отнести, во-первых, реляционные СУБД организаций, в основном построенные по модели клиент-сервер, использующиеся для обработки больших объемов данных и позволяющие одновременно использовать их ресурсы большому числу пользователей организации. Значительно более широко, чем реляционные СУБД, в настоящее время используются гипертекстовые системы глобальных вычислительных сетей. Эти системы содержат огромный объем разнообразной информации, в том числе тексты, графику, аудио и видео информацию. К сожалению, гипертекстовые системы имеют слабые средства для организации взаимодействия с пользователями, защиты информации и низкие показатели производительности. Поэтому желательно иметь возможность из страницы Web получить доступ к данным из реляционной базы данных, что позволит пользователям глобальных сетей работать с данными, хранящимися в существующих базах данных, не создавая дополнительных программ.

Из имеющихся гипертекстовых систем по уровню распространенности первое место занимает гипертекстовая система глобальных вычислительных сетей - WWW (World Wide Web), которая объединила в себе возможности всех предшествующих систем в том числе и другого типа, таких как Telnet, FTP, Gopher и т.д. Как уже отмечалось выше, WWW система имеет средства расширения, так называемый CGI-интерфейс, позволяющий создавать шлюзы между ней и любыми внешними к ней системами. С помощью интерфейса CGI-BIN в системах WWW появляется возможность организовать диалог с пользователем.

Рассматриваемая объединенная информационная система осуществляет ввод информации и взаимодействие с пользователями реляционной СУБД и генерирует страницы формата HTML для WWW-сервера, и наоборот, запросы, введенные пользователем на HTML - странице передает реляционной СУБД. При этом большое значение приобретают вопросы безопасности таких систем, разграничение доступа пользователей к информации, защита от несанкционированных обращений и необходимость контролировать сессию пользователя с объединенной системой.

Преимуществом описываемой объединенной системы является также то, что для пользователей, предоставляющих информацию для WWW, зачастую

предпочтительнее использовать реляционную информационные систему (например типа клиент-сервер), поскольку в них имеется привычный пользовательский интерфейс, и они обладают мощными средствами контроля целостности информации и многопользовательского доступа.

Известно, что при реализации распределенной разнородной системы необходимо обеспечить взаимодействие между ее компонентами [3]. При объединении реляционных баз данных и гипертекстовых систем это взаимодействие может контролироваться интерфейсной программой, работающей на основе CGI-интерфейса сервера гипертекстовых услуг.

Кроме взаимодействия компонент важную роль играет также мобильность системы - возможность запускаться на различных компьютерных платформах. Использование коммуникационных возможностей сокетов TCP для взаимодействия между компонентами распределенной БД обеспечивает гибкость и надежность распределенной системы. Более того эти сокеты поддерживают модель взаимодействия клиент/сервер, которая пригодна для взаимодействия между динамическими и статическими компонентами системы. [3] В случае использования интерфейсной программы работающей на основе интерфейса CGI-BIN мобильность поддерживается в полной мере в части доступа к гипертекстовым системам. В части доступа к реляционной СУБД важную роль при этом играет поддержка стандартов доступа к различным СУБД, например ODBC.

Одной из трудностей в проведении оценки эффективности системы управления базами данных, состящей из нескольких компонент (МСУБД) является разработка адекватной имитационной модели, способной описывать взаимодействие между МСУБД и локальными системами, которые могут использовать любой тип механизма управления согласованием транзакций [4]. МСУБД обычно строится над локальными СУБД, которые работают автономно, что является обной из важнейших характеристик МСУБД, и что отличает ее от традиционных распределенных систем баз данных. Другое существенное отличие между ними - то , что в МСУБД имеется два типа транзакций - локальные и глобальные. Локальные транзакции исполняются локальными СУБД, независимо от МСУБД, в то время как глобальные транзакции выполняются по управлением МСУБД.

Эти характеристики приводят к значительным трудностям в разработке схем управления согласованием транзакций, которые обеспечивают глобальную координацию (сериализуемость)[4]. Алгоритмы управления согласованием транзакций для МСУБД делятся на две основные категории: алгоритмы, требующие некоторую управляющую информацию от локальных СУБД и алгоритмы не требующие такой информации. Моделирование этих двух алгоритмов [4] показывает, что в случае первого алгоритма параллельное исполнение глобальных транзакций всегда более эффективно, чем последовательное исполнение этих транзакций кроме случая полностью дублированной МБД. Первый алгоритм обычно вызывает большое число откатов глобальных транзакций, что делает его менее желательной схемой для МСУБД. В случае второго алгоритма число откатов глобальных транзакций относительно меньше, и производительность МСУБД почти такая же высокая как у распределенной системы. Т.о. МСУБД обеспечивает координацию выполнения глобальных транзакций по нескольким локальным системам, причем в отсутсвии управляющей информации от локальных СУБД. Управляющая программа глобальными транзакциями МСУБД и диспетчер глобальных транзакций должны рассчитывать на худшую из имеющихся ситуация с локальной СУБД, чтобы обеспечить целостность всей системы и избежать блокировок.

При рассмотрении вопросов защиты информации в объединенной системе необходимо учитывать также следующие особенности:

1. Развитие распределенных информационных структур компьютерных сетей идет в направлении расширения набора предоставляемых услуг и значительного увеличение числа обслуживаемых ими пользователей. При этом качество услуг определяется надежностью функционирования системы, устойчивостью к умышленному проникновению с целью несанкционированного доступа к информации, ее искажения и уничтожения в целом.

2. Наращивание системы новыми возможностями приводит к увеличению числа точек проникновения и делает ее еще более уязвимой и подверженной внешним и внутренним "атакам".

Отметим, что в условиях, когда в объединенной информационной системе источник запросов (клиент), и получатель (сервер) разнесены, повышается ее уязвимость, поскольку возникает опасность проведения боковой "атаки" системы. Это объясняется тем, что даже после реализации сколь угодно сложной схемы защиты вхождения в систему , все последующие обращения к ее ресурсам остаются незащищенными, т.к. все они идентифицируются уникальным образом и схема их адресации статическая. Как следствие этого, возможны злоумышленные обращения к ресурсам, если их адреса были перехвачены или в случае, когда схема формирования адреса тривиальна.

Для решения этих проблем может быть предложено ряд подходов, основывающихся на различных способах расширения WWW услуг. Однако мы рассмотрим для начала способы, не требующий внесения изменений на стороне клиента WWW-услуги. В настоящее время наиболее распространенный способ защиты информации использует промежуточный надежный сервер "Secure Server". При этом обмен секретной информацией между клиентом и сервером производится через промежуточный надежный сервер, который берет на себя функции обеспечения защиты от несанкционированного доступа. При этом надежный сервер использует различные методы: от принципа обратной связи (call back) и использования специальных вопросников до использования специальных методов кодирования информации с предварительным обменом ключами по надежному каналу.

Способ защиты информации с использованием встраиваемого контекста основывается на использовании динамически формируемых на стороне сервера последовательностей (динамических паспортов), идентифицирующих пользователей услуги. При этом каждый запрос пользователя будет определяться уникальным идентификатором, определяющим контекстом пользовательского запроса.

Отметим, что WWW-услуга является контекстно-независимой , т.е. не запоминает информацию о проводимых сессиях и эта характеристика является для нее основополагающей. В то же время для обеспечения безопасности требуется идентификация пользователя всякий раз, когда он просматривает последовательность ресурсов, что без сохранения контекста невозможно. В этой связи предлагается сохранять контекст пользователя непосредственно в теле гипертекстовых документов. Для того, чтобы контекст отсылался обратно на сторону сервера, при издании запросов, он должен вставляться в адреса ресурсов WWW (URL) [1] указанных в гиперссылках HTML-документов. Фактически в этом случае даже если при использовании какого-нибудь ресурса будет перехвачен контекст пользователя вместе с динамическим паспортом сервера, то он уже не будет актуальным при следующем доступе к серверу, поскольку при этом сервер уже использует новый динамический паспорт.

Однако, по-видимому, наиболее надежным способом защиты информации является шифрование информации на стороне клиента. При этом если ключи

передаются по надежному каналу, то система становится устойчивой к перехвату информации на канале между клиентом и сервером. Установленно также, что шифрование информации на стороне клиента лучше организовать, используя механизм Java, дающий возможность выполнения на стороне клиента. Java - программа передается WWW-броузеру клиента и выполняется на нем. В этом случае WWW - сервер предоставляет клиенту JAVA - программу шифрования, а клиент запрашивает у пользователя его уникальный идентификатор и ключ. Клиент передает идентификатор пользователю и по его идентификатору CGI скрипт выбирает ключ, которым он будет осуществлять шифрование предоставляемой клиенту информации. На стороне клиента информация будет расшифровываться переданным на компьютер пользователя Java - скриптом - дешифратором. Данный способ организации взаимодействия между клиентами и серверами гипертекстовых систем, являясь наиболее надежным из рассмотренных, уязвим лишь с точки зрения уязвимости операционной системы самого WWW сервера и выбранного алгоритма шифрования (см. рис. 1.).

Результаты исследований вопросов взаимодействия и управления в информационных системах различного типа, защита информации в объединенной информационной системе были использованы при создании объединенной информационной системы учета материалов международной конференции в Институте Проблем Информатики и Автоматизации.

В рамках созданного Web сервера информационных материалов конференции предоставлена возможность средствами клиента гипертекстовой системы Internet WWW работать с информацией из базы даннных по материалам конференции. В качестве системы управления реляционной базы данных использовался сервер СУБД Mini SQL в среде ОС UNIX. Данная система, размещаясь на коммуникационной ЭВМ, поддерживает известный стандарт запросов к БД - язык структурированных запросов SQL.

Реализованная система имеет следующую обобщенную структуру (см рис. 2.)

Взаимодействие между сервером WWW и СУБД осуществлятся через специально разработанную интерфейсную программу работающую по стандарту CGI-BIN этого WWW сервера.

Пользователь обращается к интрфейсному CGI-скрипту - программе на языке Perl, которая производит его аутентификацию и присваивает некоторый случайный контекст - динамический паспорт пользователя.

В дальнейшем все те запросы к скрипту, которые содержат в себе этот контекст транслируются на язык запросов к базе данных SQL и направляются к СУБД MSQL . Скрипт обрабатывает результаты запроса полученные от СУБД MSQL, превращает его в формат HTML и направляет соответствующему клиенту . Т.о. взаимодействие сервера с СУБД происходит прозрачно для клиента. Кроме обеспечения защиты информации интерфейсная программа также контролирует целостность объединенной базы данных и управляет транзакциями над ней,

Представим упрощенную схему интерфейсной программы для разработанной системы (рис. 3.)

В реализованной системе предусмотерены возможность просмотра, фильтрации, удаления, добавления и модификации информации из нескольких таблиц. В зависимости от приоритетов, на основе которых предоставляется контекст, пользователям доступны функции только по просмотру, либо как по просмотру, так и по модификации информации из базы данных.

В режиме модификации базы данных окно броузера клиента при работе с системой может иметь, например, следующий вид (рис. 4.)

В настоящее время разработанная система модифицируется с целью обеспечения возможности взаимодействия Web сервера с несколькими серверами БД расположенными на нескольких компьютерах. Т.е. система модифицируется в систему управления несколькими разнородными системами баз данных (МСУБД).

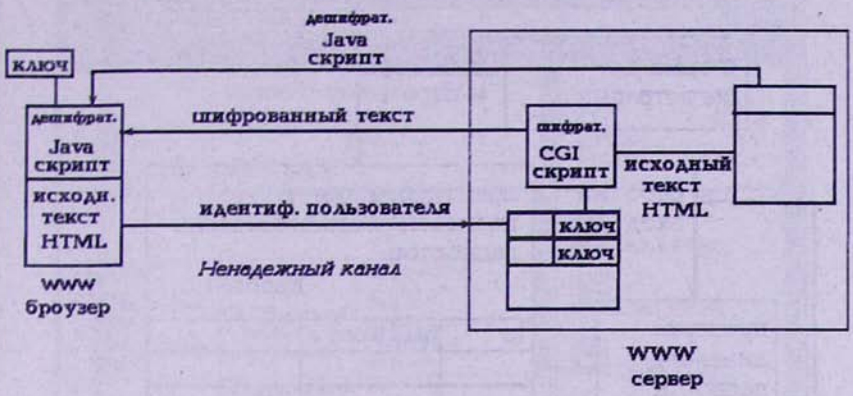

Рис. 1. Схема шифрования информации на стороне клиента

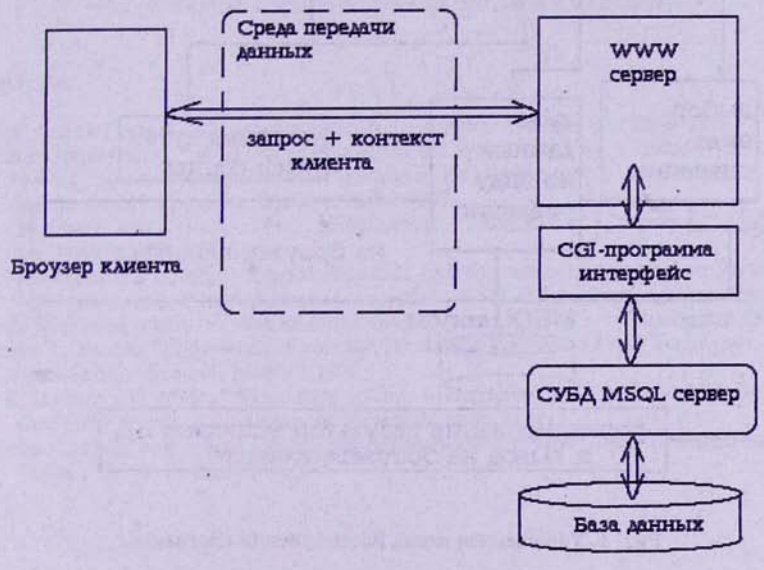

Рис. 2. Обобщенная структура объединенной информаационной системы.

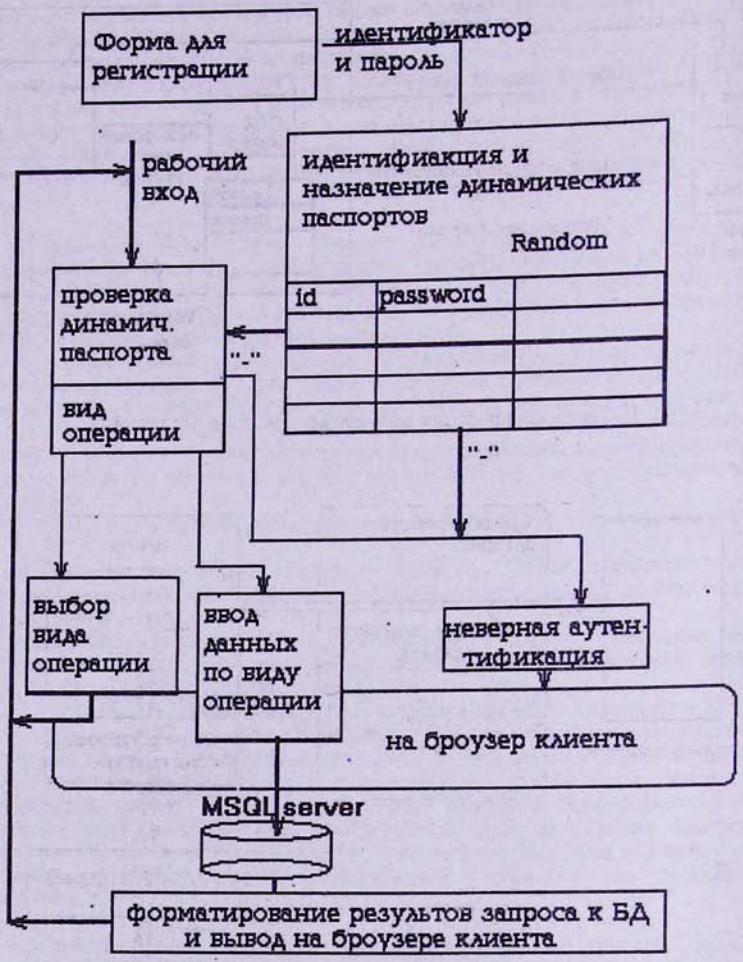

Рис. 3. Упрощенная схема интерфейсной программы

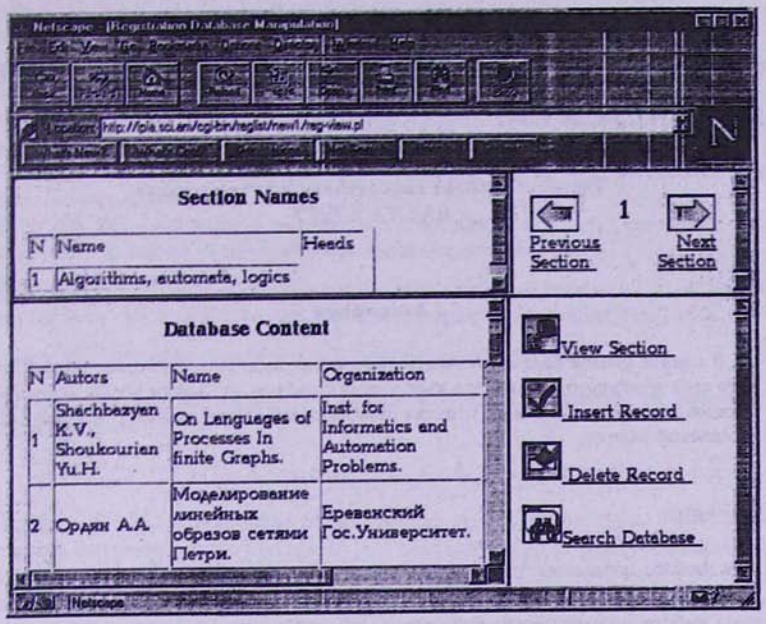

Рис. 4. Вид окна броузера клиента при работе с системой

Литература

1. Cricet Lui and others, " Managing Internet Inormation Services", O'Relly 1996
2. Shishir Gundavaram, " CGI-Programming on the Word wide Web", O'Relly 1997
3. J. Chen and others "Cooperative Mechanisms in Heterogenious Multidatabase Environment" Computer Systems, Vol 3 1993
4. Y. Breitbart and others, "Two Multidatabase Transaction Mangement Algoritms", Computer Systems, Vol 3 1993
5. D. Movsesyan and others, "Optimal allocation of Information in Information Systems of Networks with low loaded transfer channels"., CSIT-97 conference proceedings.
6. Owen Rees, and others "A Web of Distributed Objects" ANSA Workgroup 1996
7. Charles L. Brooks "Wide area Information Brownsing Assistance Final Technical Report " The Open Group Research Institute 1996.
8. D. B. Ingham and others, "Supporting Hilghly Managable Web Services " Dept. of Computer Science , University of Newcastle upon Tyne
9. Charles Cavaiani and others, "A Mutual Authenticating Protocol with Key Distribution in a Client / Server Environment."