БУ ВО

«СУРГУТСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ

Ханты-Мансийского автономного округа – Югры»

Кафедра автоматизированных систем обработки информации и управления (АСОИУ)

ДОКЛАД

по теме: «Формы распределения данных и их различия»

Выполнил студент гр. 606-71м

Боровков Дмитрий Евгеньевич

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_(подпись)

Преподаватель:

Горбунов Дмитрий Владимирович

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_(подпись)

Сургут, 2018

ВВЕДЕНИЕ

Распределение данных – это процесс, где данные распределяются по различным носителям. Любое производство содержит огромное количество информации из самых разных областей, как правило, это базы данных о сотрудниках, заказчиках, посредниках при торговле, поставщиках сырья и так далее. Кроме баз данных имеются сведения о производимых товарах, финансовые отчеты, бухгалтерия и так далее. Все эти сведения являются самой главной ценностью компании, для их сохранения предприниматели распределяют информацию по различным компьютерам, как правило, не объединенных в сеть, или записывают ее на различные информационные носители. Это сберегает информацию от потери. Распределение данных может быть характерно не только для предпринимательских компаний, но и любых иных общественных объединений, научных институтов [1].

1. Потребность в распределении данных

Самой ценной вещью для компании является информация, её в первую очередь интересует сохранность этой информации – защита их от возможного повреждения. Вероятность, что, например, информационный вирус приведет в негодность все компьютеры предприятия, гораздо ниже, чем вероятность, что он сотрет данные с какого-то одного компьютера. Поэтому выгодно распределять информацию и распространять ее копии.

Распределение данных иногда совершается с целью облегчения работы различных отделов предприятия. Например, отделу статистики, вероятнее всего, не нужны будут базы данных со сведениями о сотрудниках, зато им понадобятся отчеты компании об уровне производства, среднем значении цен и так далее. Распределив данные, можно разместить в каждом отделе только самые необходимые информационные материалы, что ускорит обработку данных и снизит нагрузку на сервера.

Третье значение распределения данных заключается в том, что после распределения появляется возможность быстрого изъятия и предоставления необходимых сведений. Например, если владельцу компании срочно понадобится проанализировать статистические материалы, он запрашивает данные в отделе статистики и получает их немедленно. Если бы данные не были распределены и находились на каком-то одном сервере, то поиск необходимой информации занимал бы больше времени: сказывалось бы и отсутствие структурированности и информационная загруженность, замедляющая работу компьютера.

1. Формы распределения данных
   1. Централизованная организация данных

Централизованная организация данных является самой простой для реализации.

На одном сервере находится единственная копия базы данных. Все операции с базой данных обеспечиваются этим сервером. Доступ к данным выполняется с помощью удалённого запроса или удалённой транзакции.

Достоинством такого способа является лёгкая поддержка базы данных в актуальном состоянии, а недостатком – то, что размер базы ограничен размером внешней памяти: все запросы направляются к единственному серверу с соответствующими затратами на стоимость связи и временную задержку. Отсюда ограничение на параллельную обработку. База может быть недоступной для удалённых пользователей при появлении ошибок связи и полностью выходит из строя при отказе центрального сервера.

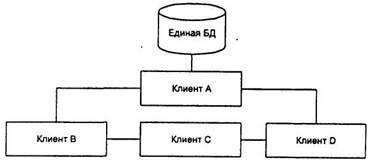


Рис.

* 1. Децентрализованная организация данных

Децентрализованная организация данных предполагас1 разбиение информационной базы на несколько физически распределенных Каждый клиент пользуется своей базой данных, которая может быть либо частью общей информационной базы (рис 2.6), либо копией информационной базы в целом (рис 2.7), что приводит к ее дублированию для каждого клиента.

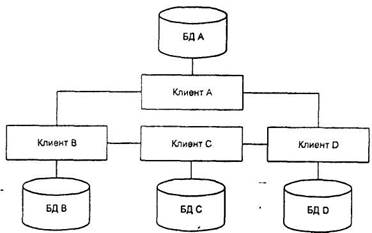


Рис 2.6. Децентрализованная организация данных способом распределения

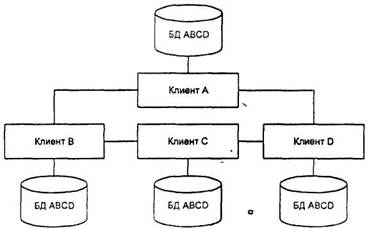


Рис 2.7. Децентрализованная организация данных способом дублирования

При распределении данных на основе разбиения база данных размещается на нескольких серверах. Существование копий отдельных частей недопустимо. Достоинства этого метода: большинство запросов удовлетворяются локальными базами, что сокращает время ответа; увеличиваются доступность данных и надежность их хранения; стоимость запросов на выборку и обновление снижается по сравнению с централизованным распределением; система останется частично работоспособной, если выйдет из строя один сервер. Имеются и недостатки: часть удаленных запросов или транзакций может потребовать доступ ко всем серверам, что увеличивает время ожидания и цену обслуживания; необходимо иметь сведения о размещении данных в различных БД. Однако доступность и надежность увеличатся. Расчлененные базы данных наиболее подходят к случаю совместного использования локальных и глобальных сетей ЭВМ.

Способ дублирования заключается в том, что в каждом сервере сети ЭВМ размещается полная база данных Это обеспечивает наибольшую надежность хранения данных. Недостатки способа: повышенные требования к объему внешней памяти, усложнение корректировки баз, так как требуется синхронизация в целях согласования копий. Достоинства - все запросы выполняются локально, что обеспечивает быстрый доступ. Данный способ используется, когда фактор надежности является критическим, база небольшая, интенсивность обновления невелика.

* 1. Смешанная организация данных

Смешанная организация хранения данных объединяет два способа распределения: разбиение и дублирование, приобретая при этом и преимущества и недостатки обоих способов.

Появляется необходимость хранить информацию о том, где находятся данные в сети. При этом достигается компромисс между объемом памяти под базу в целом и под базу в каждом сервере, чтобы обеспечить надёжность и эффективность её работы: легко реализуется параллельная обработка, то есть обслуживание распределённого запроса, или транзакции. Несмотря на гибкость смешанного способа организации данных остаётся проблема взаимозависимости факторов, влияющих на производительность системы, проблема её надёжности и выполнения требований к памяти. Смешанный способ организации данных можно использовать лишь при наличии сетевой СУБД.

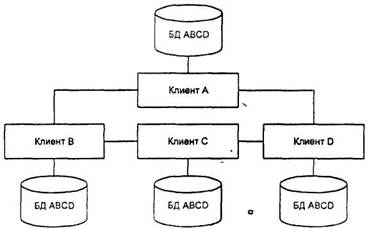


Рис.

В базах данных коллективного пользования центральным технологическим звеном становятся серверы баз данных. Программные средства серверов баз данных обеспечивают реализацию многопользовательских приложений, централизованное хранение, целостность и безопасность данных. Производительность серверов баз данных на порядок выше по сравнению с файл-серверами, которые используются в локальных сетях. Локальные вычислительные сети создавались для совместного использования дорогостоящего периферийного оборудования. Использование сервера баз данных обеспечило доступ многих пользователей к одним и тем же файлам. Это и стало предпосылкой создания сетевых СУБД.

Мощность сетевых СУБД, основанных на файл-сервере, в настоящее время недостаточна. В нагруженной сети неизбежно падает производительность, нарушаются безопасность и целостность данных. Проблема производительности возникла не потому , что процессоры 386 не обладают достаточной мощностью, а потому что файл-серверы реализуют принцип «все или ничего». Полные копии файлов базы перемещаются взад-вперед по сети. Проблемы с безопасностью, целостностью возникли из-за того, что с самого начала файл-серверы не были сконструированы с учетом сохранения целостности данных и их восстановления в случае аварии.

Технология клиент-сервер, как более мощная, заменила технологию файл-сервер. Она позволила совместить достоинства однопользовательских систем (высокий уровень диалоговой поддержки, дружественный интерфейс, низкая цена) с достоинствами более крупных компьютерных систем (поддержка целостности, защита данных, многозадачность)

В классическом понимании СУБД представляет собой набор программ, позволяющих создавать и поддерживать базу данных в актуальном состоянии. Функционально СУБД состоит из трех частей: ядра (базы данных), языка и инструментальных средств программирования.

Инструментальные средства программирования относятся к интерфейсу клиента, или внешнему интерфейсу. Они могут включать процессор обработки данных на языке запросов. Язык - это совокупность процедурных и непроцедурных команд, поддерживаемых СУБД. Наиболее употребительными языками являются SQL и QBE. Ядро выполняет все остальные функции , которые включены в понятие «обработка базы данных»

Основная идея клиент-сервера заключается в юн, чтобы серверы расположить на мощных машинах, а приложения клиентов, использующих язык, - на менее мощных машинах Тем самым будут задействованы ресурсы более мощного сервера и менее мощных машин клиентов. Ввод-вывод к базе основан не на физическом дроблении данных, а на логическом, т.е. сервер отправляет клиентам не полную копию базы, а только логически необходимые порции, тем самым сокращая трафик сети. Трафик сети - это поток сообщений сети. В технологии клиент-сервер программы клиента и его запросы хранятся отдельно от СУБД. Сервер обрабатывает запросы клиентов, выбирает необходимые данные из базы данных, посылает их клиентам по сети, производит обновление информации, обеспечивает целостность и сохранность данных.

1. Технологии распределенной обработки данных.

Технология клиент-сервер, ориентированная на автономный компьютер, т.е. и клиент, и сервер размещены на одной ЭВМ. По функциональным возможностям такая система аналогична централизованной СУБД. Ни распределенная обработка, ни распределенная СУБД не поддерживаются.

Технология клиент-сервер, ориентированная на централизованное распределение. При использовании этой технологии клиент получает доступ к данным одиночного удаленного сервера, данное могут только считываться, динамический доступ к данным реализуется- посредством удаленных транзакций и запросов, их число должно быть невелико, чтобы не снизилась производительность системы.

Технология клиент-сервер, ориентированная на локальную вычислительную сеть. Эта технология характеризуется следующими особенностями: единственный сервер обеспечивает доступ к базе; клиент формирует процесс, отвечающий за содержательную обработку данных, их представление и логический доступ к базе; доступ к базе данных замедлен, так как клиент и сервер связаны через локальную сеть.

Технология клиент-сервер, ориентированная на изменения данных в одном месте В случае применения этой технологии реализуется обработка распределенной транзакции; удаленные серверы не связаны между собой сетью ЭВМ, т.е. отсутствует сервер-координатор; клиент может изменять1 данные только в своей локальной базе; возникает опасность «смертельных объятий», т.е. ситуация, когда задача А ждет записи, заблокированные задачей В, а задача В ждет записи, заблокированные задачей А. Поэтому распределенная СУБД должна иметь средство контроля совпадений противоречивых запросов. Распределение данных реализует метод расчленения.

Технология клиент-сервер, ориентированная на изменение данных в нескольких местах. В отличие от предыдущей технологии здесь имеется сервер-координатор, поддерживающий протокол передачи данных между различными серверами Возможна обработка распределенных транзакций в разных удаленных серверах. Это создаст предпосылки разработки распределенной СУБД. Реализуется стратегия смешанного распределения путем передачи копий с помощью СУБД.

Технология клиент-сервер, ориентированная на распределенную СУБД. Она обеспечивает стратегию разбиения и дублирования, позволяет получить более быстрый доступ к данным. Распределенная СУБД обеспечивает независимость клиента от места размещения сервера, глобальную оптимизацию, распределенное административное управление.

Во всех технологиях существуют два способа связи прикладных программ клиента и. сервера баз данных: прямой и непрямой. При прямом соединении прикладная программа клиента связывается непосредственно с сервером базы данных, а при непрямом- доступ к удаленному серверу обеспечивается средствами локальной базы. Возможно объединение обоих способов.

Использование технологии клиент-сервер позволяет перенести часть работы с сервера на ЭВМ клиента, оснащенную инструментальными средствами для выполнения его профессиональных обязанностей. Тем самым данная технология позволяет независимо наращивать возможности сервера баз данных и совершенствовать инструментальные средства клиента. Недостаток технологии клиент-сервер заключается в повышении требований к производительности ЭВМ- сервера, в усложнении управления вычислительной сетью, а при отсутствии сетевой СУБД - в сложности организации распределенной обработки.

Под операционной средой сервера баз данных понимают возможности ОС компьютера и сетевой ОС. Каждый сервер баз данных может работать на определенном типе компьютера и сетевой ОС. Операционные системы серверов - это DOS-версии 5.0, XENIX, UNIX, Windows NT, OS/2 и др В настоящее время наиболее употребительными являются около десяти серверов, в частности SQL-server, SQLBASE-scrvcr, ORACLE-servcr и др. По экспертным оценкам, серверам баз данных принадлежит будущее

Серверы баз данных рассчитаны на поддержку большою числа различных типов приложении Для реализации интерфейса с сервером базы данных можно использовать объектно-ориентированные средства, электронные таблицы, текстовые процессоры, графические пакеты, настольные издательства и другие информационные технологии.

ЗАКЛЮЧЕНИЕ

До сегодняшнего дня не решен вопрос об эталонной форме распределения данных, скорее всего, и не будет решен. Каждая компания уникальна и их приоритеты и цели тоже различны. Выбор формы распределения данных целиком зависит от входных параметров, таких как, уровень безопасности, географическая распределённость предприятия, бюджет и так далее.

СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ

<https://utmagazine.ru/posts/14172-raspredelenie-dannyh>

<https://life-prog.ru/1_27914_sposobi-raspredeleniya-dannih.html>

https://xstud.ru/135630/informatsionnye\_tehnologii\_v\_ekonomike/raspredelennye\_tehnologii\_obrabotki\_hraneniya\_dannyh