

УДК 004.272.3

Иванкевич А.В., канд. техн. наук

ИСПОЛЬЗОВАНИЕ СПЕЦИАЛИЗИРОВАННЫХ ПРОГРАММНЫХ КОМПЛЕКСОВ НА БАЗЕ РАСПРЕДЕЛЕННЫХ ХРАНИЛИЩ ДАННЫХ АВИАПРЕДПРИЯТИЙ УКРАИНЫ

Институт компьютерных технологий
Национального авиационного университета

Предложены пути внедрения балансировки распределения информационных файлов по узлам компьютерной сети при обработке больших объемов данных в информационных системах, которые используются на авиапредприятиях Украины. Исследованы особенности работы систем распределенных систем авиапредприятий и перспективы повышения производительности компьютерных сетей, на базе которых работают системы

Введение

В настоящее время большинство авиапредприятий мира, в том числе и аэропорты Украины, используют большое количество сложных информационных систем. В системах хранятся и обрабатываются большие объемы данных, которые с каждым днем продолжают возрастать. Данные в системах имеют реальную ценность лишь в том случае, если эти данные получены в приемлемые сроки. Это значит, что для таких систем необходимо обеспечить как можно меньший промежуток между запросом пользователя и получением им данных.

Поэтому на большинстве авиапредприятий не только наращивают емкость носителей, которые хранят базы данных (БД) систем, но и увеличивают количество центров хранения и обработки данных, а также повышают скорость доступа к ним по сети.

Для ускорения обработки больших БД часто используют распределение файлов БД по нескольким серверам, соединенным между собой локальной сетью [1-3]. Системы управления БД (СУБД), использующие распределенную обработку таблиц БД, позволяют оптимизировать работу с БД за счет применения индексных таблиц, кэширования, упреждающего чтения данных и т.д. [4]. Однако такая оптимизация эффективна при работе с БД небольших объемов, а ее эффект уменьшается по мере увеличения объема данных.

Одним из способов дополнительного повышения производительности существующих информационных систем на базе СУБД является распараллеливание алгоритмов [5]. Обработка запроса в СУБД требует обращения к нескольким неделимым операциям, разложение и последовательность которых зависит от конкретного запроса, а также плана выполнения, выбранного оптимизатором запроса. Обычно операции БД выполняются в строгой последовательности, причем выход одной операции служит входом следующей. Для поддержки параллельной обработки запросов применительно к нескольким серверам распределенной БД (РБД) выбирают неделимые операции, которые могут быть скопированы и далее одновременно производят обработку фрагментов данных и операции, результаты которых могут быть представлены так, как будто единая цепочка вычислений выполнила операцию. С этой целью запросы разбивают и обрабатывают независимо от способов коммуникации. Параллельная обработка различных участков БД на серверах позволяют СУБД эффективно использовать многомерные таблицы и разбиение таблиц. При правильном распределении параллельной обработки между серверами в сети существенно возрастает эффективность работы с БД и управления ресурсами [1-3].

Целью работы являлась доработка существующей в аэропорту Борисполь и других аэропортах Украины, системы,

позволяющей распределять потоки данных информационных систем аэропорта по нескольким серверам и увеличивать производительность обработки запросов в этих системах.

Структура информационных систем аэропорта

Аэропорт является сложным комплексом, который включает в себя комплексы, оперирующие оперативной, статистической и финансовой информацией. В настоящее время развитие информационных систем аэропорта сводится к созданию единой открытой системы, которая основана на использовании единого хранилища данных различными информационными комплексами, основанными на различных платформах и операционных системах. Единое хранилище построено на основе системы серверов, соединенных между собой высокоскоростной сетью.

Для доступа к единому хранилищу все комплексы должны быть соединены между собой локальной сетью, которая имеет выходы в глобальные сети *SITA*, *AFTN*, *Internet* и т.д. В целях обеспечения безопасности, системы, работающие с глобальными сетями, не имеют прямого доступа к единому хранилищу [6].

Для таких систем созданы дополнительные хранилища, которые обеспечивают полную независимость единой системы от внешних факторов и позволяет вносить модификации данных, полученных из различных внешних источников. Например, Управлением воздушным движением Украины, в БД системы управления взлётом и посадкой может быть внесена корректировка, которая через определенное время автоматически вносится во все остальные системы аэропорта, например в систему информирования пассажиров. К подобным системам также можно отнести систему информирования о погодных условиях, системы *SITA* и *AFTN*.

К системам, которые работают в аэропорту и используют единое хранилище данных [6], можно отнести:

- информационную систему индикации расписания полетов в аэропорту (*AFIDS*);

- информационную систему информирования пассажиров о взлете и посадке (*ALDIS*);

- систему администрации контроля авиаперевозок (*ATCAS*);

- систему составления полетных планов, работающая как подсистема Управления воздушным движением Украины (*ATCBill*);

- систему выставления счетов;

- план полетов;

- систему агента по обработке грузов;

- систему бухгалтерского учета;

- службу сообщений;

- информационную систему службы авиационной безопасности

- систему информирования пассажиров через внутреннюю сеть аэропорта (*RAMP*);

- систему работы с агентами авиакомпаний (*HAS*).

Распределение данных между серверами компьютерной сети аэропорта

Для реализации алгоритмов оптимизации размещения информационных файлов по узлам реальной компьютерной сети аэропорта, прежде всего, необходимы достоверные исходные данные [7]. Для определения интенсивности обращений к различным файлам БД используется резидентная программа "Анализатор запросов", написанная на языке Ассемблер. Этот язык программирования обеспечивает компактность и гибкость при написании резидентных программ и не вносит погрешность в измеряемые процессы. Резидентная программа загружается на всех узлах сети. Время компьютеров синхронизируется. Информация накапливается в течение одного рабочего дня.

Программа анализирует все обращения к файлам и протоколирует в журнал обращений дату, время, имя таблицы, длину запроса, время до ответа, длину ответа. Протоколирование производится во внутренний буфер и записывается на диск только во время простоя компьютера, когда осуществляются ввод с клавиатуры или другие операции, требующие ожидания ответа.

С помощью нерезидентной части программы проведен анализ обрабатываемой информации и выявлены временные характеристики обращений к файлам. Загрузка сети характеризуется неравномерностью – полным отсутствием обращений или одновременным обращением со всех рабочих станций.

Для перераспределения поисковых запросов к информационно-справочным файлам между узлами компьютерной сети предназначена программа "Оптимизатор запросов в локальной сети", существенно уменьшающая вероятность образования очередей.

Работа данной программы предусматривает подготовку справочника местонахождения файлов *inquire.txt*. Для обеспечения возможности изменять по желанию перечень файлов-справочников и их адреса в локальной сети при работе с программой "Оптимизатор запросов в локальной сети" предусмотрено получение информации о конфигурации сети из файла *inquire.txt*. В файле хранится информация обо всех файлах-справочниках и их местонахождении. Разработанный формат записи в файле *inquire.txt* выглядит так: № сети, № узла, устройство \[путь\]имя файла. Изменения в файле *inquire.txt* требуют перезагрузки программы.

Программа "Оптимизатор запросов в локальной сети" состоит из трех частей: инсталляции, перехвата запросов, обработки запросов. В процессе инсталляции программа может выдавать определенные сообщения.

Поскольку предстоит обрабатывать запросы на открытие-закрытие, а также чтение-записи файла, то резидентной части программы требуется обрабатывать функции открытия файла; закрытия файла; чтения из файла; записи в файл.

Рассмотрим алгоритм работы программы перехвата запросов.

1. Перехват вектора прерывания 21H.

2. Если номер функции равен FFH и номер подфункции равен 66H, то в регистр AX заносится кодовое значение 6677H и выход из прерывания.

3. Если номер функции не равен 3DH или 3FH, или 3EH, или 40H, то прерывание отдается старому обработчику. Иначе обработчик заполняет пакет протокола IPX с запросом обработки процесса на каком-либо из узлов сети, посылает сообщение по сети и включает режим ожидания ответа на свой запрос. После получения ответа программа заполняет регистры флагов и AX значениями, прошедшими в пакете, и передает управление вызвавшей прерывание программе.

Оптимизация обработки данных для информационной системы агентов авиакомпаний

В аэропортах Украины часто используется система управления данными агентов авиакомпаний. Система работает с БД большого объема и установлена на сегменте локальной сети, состоящем из 8 узлов.

Система осуществляет автоматическую обработку информации, вводимой авиакомпаниями и агентами по обработке грузов. Информация вводится в состав системы и в последующем обрабатывается в соответствии со стандартами IATA. При просмотре и редактировании данных агенту авиакомпании предоставляются такие возможности:

- возможность добавления, удаления и приостановки выполнения рейса, ввод времени вылета и прибытия для пассажиров;

- ввод и редактирование информации по пассажирам и грузам (включающее возможность заказа определенного бортипитания);

- ввод запросов на выделение мест регистрации пассажиров, время открытия и закрытия регистрации на рейс;

- ввод запросов на выделение гейтов для прохода пассажиров на воздушное судно (ВС), время открытия и закрытия гейтов;

- ввод деталей полета (которые не регламентируются Управлением воздушным движением Украины);

- ввод оценочного времени взлета и посадки ВС во время прибытия и вылета (которое не регламентируются Управлением воздушным движением Украины);

- ввод сезонных изменений в расписании.

Ввод оценочного времени взлета и посадки используется системой выставления счетов авиакомпаниям, чтобы вычислить стоянку времени. Оценочное время также используется для сбора статистики по времени рулежки воздушного судна.

Информация по пассажирам и грузам информация используется для вычисления нагрузок на терминалы аэропорта, для сбора статистической информации и вычисления коэффициента загрузки для терминалов в каждый момент времени, а также для анализа прибыли аэропорта.

Информация о времени открытия и закрытия регистрации пассажиров, а также информация о времени открытия и закрытия гейтов для прохода пассажиров, используется в базе данных системы индикации информации о полетах и системой выставления счетов за услуги аэропорта авиакомпаниям.

При функционировании системы возможны следующие действия пользователя:

- использование универсального фильтра, т.е. любая выборка;
- изменение порядка отображения записей (сортировка);
- печать текущей отфильтрованной таблицы с заданием определенного порядка показа печатаемых колонок, с изменением их ширины, с подсчетом сумм по числовым столбцам;
- удаление или замена отфильтрованных данных;
- превращение таблицы в карточку для более наглядного просмотра одной записи таблицы.

Кроме того, система предоставляет пользователю дополнительные возможности:

- 1) работу с бланками заказов, не выходя из программы, просмотр и печать поступивших и хранящихся в специальной базе изображений заказов, выдачу отчета по загрузке различных рейсов авиакомпании;
- 2) контекстно-зависимую помощь;

3) контроль за загруженностью мест регистрации;

4) окончательное расписание;

5) печать на бланках IATA, формирование описей, реестров и т.д.;

6) проверку целостности информации и автоматическое устранение нарушений;

7) реконфигурацию системы по желанию пользователя или при возникновении изменений в законодательстве;

8) гибкую систему паролей для ограничения доступа, крипто-графическое закрытие информации.

Большие объемы обрабатываемой информации могут значительно снизить скорость работы программного пакета при нерациональном распределении информационных файлов.

Для распределения информационных файлов в системе управления данными агентов авиакомпаний использована программа *opt1*. В качестве критерия оптимальности принято общее время, необходимое для обслуживания всех запросов, поступивших в систему за единицу времени.

Пользователи сети работают с данными, хранящимися в 8 информационных файлах БД. При формировании массивов исходных данных учитывают оценку интенсивностей корректирующих и поисковых запросов λ_i , λ'_i и среднего времени поиска в файлах t_{ij} , полученных с помощью резидентной программы "Анализатор запросов"[7].

Интенсивность запросов на корректировку и поиск в различных информационных файлах РБД системы управления данными авиакомпаний и среднего времени поиска были записаны и из этих данных был составлен файл исходных данных *opt1.dat* для программы оптимизации *opt1* в такой последовательности:

- интенсивность поисков и корректировок в каждом файле РБД;
- длина записей в файлах и среднее время обработки поисковых запросов к информационным файлам.

Если в качестве критерия оптимальности принято общее время, необходимое

для обслуживания всех запросов, поступивших в систему в течение единицы времени, то будет составлена матрица рационального распределения файлов, в которой первые четыре файла БД, интенсивности корректировок в которых достаточно велики, хранятся на одном из более производительных узлов. Такое размещение копий информационных файлов даст возможность избежать лишней избыточности информации и трудностей по ее согласованию.

Выводы

Предложена практическая реализация метода балансировки загрузки узлов сети при обработке данных больших и сверхбольших объемов. Результаты работы системы управления данными агентов авиакомпаний, которая работает с использованием предложенного метода, свидетельствуют о возможности значительного увеличения скорости обработки данных в РБД большого объема за счет использования механизмов оптимизации загрузки узлов сети, используемых для обработки РБД. Применение предложенного метода позволяет повысить производительность и снизить время реакции информационной системы, работающей с РБД.

Список литературы

1. Жуков И. А., Зайченко Ю. П., Печурин Н. К. Модель распределения информационных ресурсов в компьютерных сетях // Проблемы інформатизації та управління. – К.: НАУ, 2005. – Вип. 3 (14). – С. 9 – 14.
2. Аббасов А. М. Оптимизация размещения информационных баз с копиями в сети ЭВМ // Автоматика и вычислительная техника. – 1988. – №4. – С. 71 – 75.
3. Янбых Г. Ф., Бобер В. И., Бокоев Т. И. Оптимизация размещения файлов и каналов передачи данных в сети ЭВМ // Автоматика и вычислительная техника. – 1984. – №4. – С. 25 – 29.
4. Озкарахан Э. Машины баз данных и управление базами данных. – М.: Мир, 1989. – 696 с.

5. Системы параллельной обработки Пер. с англ. / Под ред. Д. Ивенса. – М.: Мир, 1985. – 416 с.

6. Airport Information Management and Display System (<http://www.iecbd.net/Products/AIMDS/tabid/74/Default.aspx>)

7. Жуков И. А., Иванкевич А. В., Салим Аль Шибани, Аль-Сурики Ибрагим Метод и программные средства оптимизации распределения информационных файлов в компьютерных сетях // Проблеми інформатизації та управління. – К.: НАУ, 2005. – Вип. 4 (15). – С. 78 – 8.