



УДК 004.65:004.051

© 2006 г. **Ю.А. Григорьев**, д-р техн. наук,  
**С.П. Остриков**,

(Московский государственный технический университет им. Н.Э. Баумана)

**А.Д. Плутенко**, д-р техн. наук

(Амурский государственный университет, Благовещенск)

## **МОДЕЛИРОВАНИЕ ИНФОРМАЦИОННОЙ СИСТЕМЫ «КОНТИНГЕНТ»**

В статье приведены результаты моделирования модуля «Контингент» автоматизированной системы МГТУ им. Н.Э. Баумана на этапе его проектирования. С помощью экспертной системы КСАМ выполнен анализ нагрузочной способности модуля и влияния наполнения базы данных. По результатам моделирования выработаны рекомендации по устранению потенциальных «узких мест» проектируемого модуля.

### **Введение**

Система "Контингент" является модулем автоматизированной системы управления вузом и предназначена для ввода, хранения и обработки информации о студентах, обучающихся в МГТУ им. Н.Э. Баумана, для сопровождения структуры учебных подразделений вуза и учебных групп, а также для описания направлений подготовки и специальностей высшего профессионального образования, отраслей науки и специальностей послевузовского профессионального образования, по которым осуществляется обучение в университете. Система позволяет решать следующие задачи: поддержка организационной структуры учебных подразделений университета; поиск информации о студентах по различным критериям; формирование и предоставление списков студентов; реализация механизмов и процедур изменения сведений о студентах на основании приказов; изменение справочных данных о студентах посредством прямого редактирования карточек; накопление и предоставление доступа к приказам по контингенту; формирование и вывод на печать типовых отчетов.

Целью моделирования ИС «Контингент» с помощью КСАМ [1] является определение работоспособности системы при пиковых нагрузках, возможности расширения системы с использованием существующих аппаратных средств, выявление узких мест системы и поиск способа их устранения с учетом запросов к базе данных.

## Структура системы «Контингент»

При проектировании системы была проанализирована трехзвенная архитектура: «Web-клиент – Web-сервер+сервер приложений – сервер базы данных». Структура системы «Контингент» представлена на рис. 1.

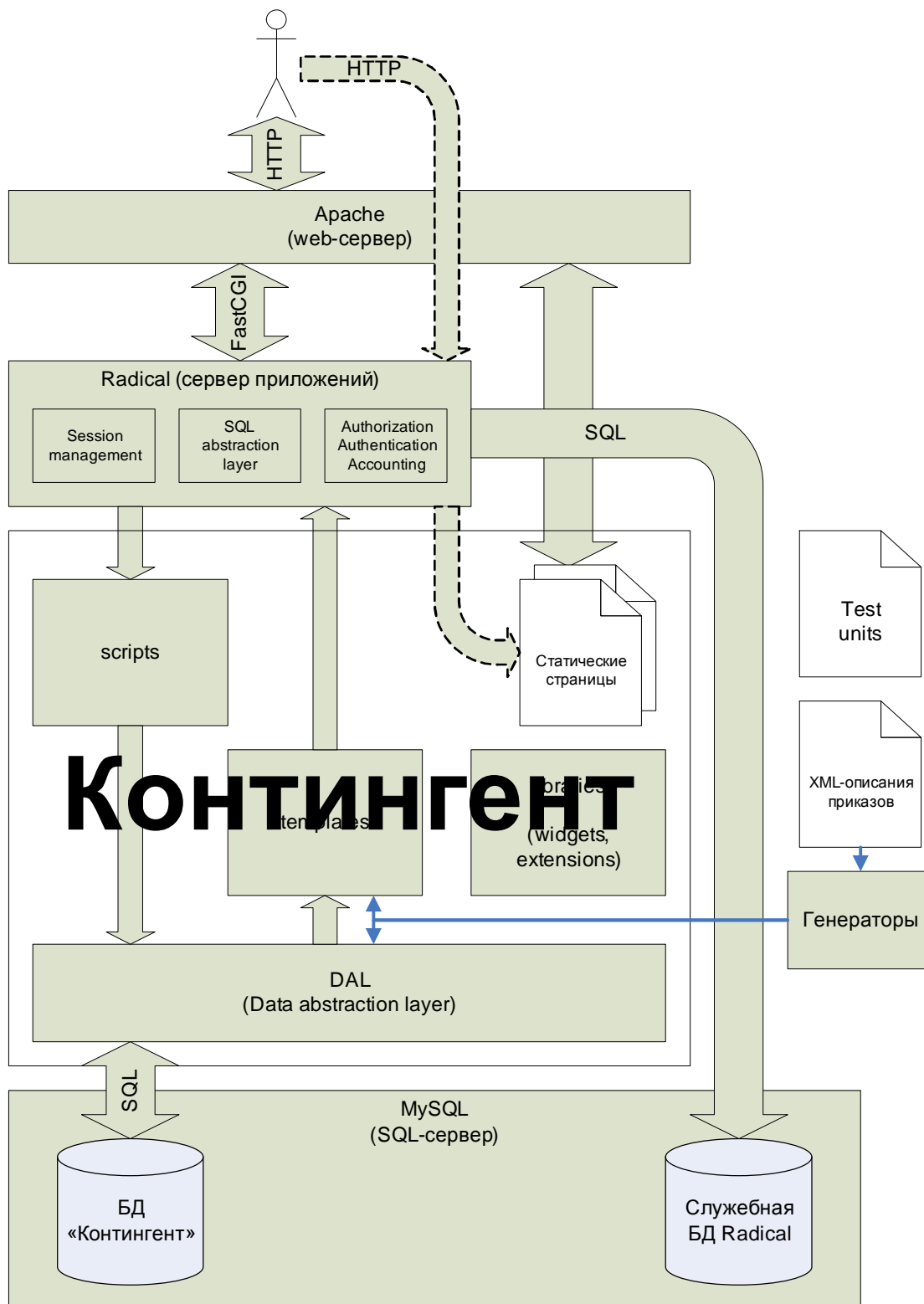


Рис. 1. Структура системы «Контингент».

Пользователь формирует запрос и отправляет его Web-серверу, который реализован на основе Apache и функционирует под управлением ОС Linux Red Hat. Если это запрос к HTML-документу, то Apache выдает требуемый статический документ, в противном случае запрос передается серверу приложений Radical по протоколу FastCGI.

Тот запускает прикладную программу (приложение), написанную на скриптовом языке Ruby. При разработке приложения программист может использовать абстрактные интерфейсы DAL (Data Abstraction Layer). Эти классы позволяют определять сущности-объекты бизнес-процессов – такие как «студент» (Student), «приказ» (абстрактный класс Order и множество его конкретных реализаций – по одной на каждый тип приказа) и т.п. DAL, в свою очередь, занимается только представлением этих объектов в терминах реляционной базы данных – он работает уже с конкретными SQL-запросами, отправляя их на SQL-сервер MySQL 4.0 и интерпретируя полученные ответы.

Стоит отметить также еще несколько важных компонентов, не входящих напрямую в состав Web-приложения в рамках Radical. Для унифицированного описания приказов был создан специальный формат описания приказа на XML-языке. Такой XML-файл «пропускается» через отдельное приложение-«генератор», в результате чего из одного целостного описания приказа автоматически создаются несколько файлов для DAL и шаблонов.

На рис. 2 приведена схема базы данных «Контингент» и описание ее компонентов.

С помощью KСAM были промоделированы следующие SQL-запросы, поддерживающие выполнение задач системы «Контингент», для следующих групп запросов:

реализующих функциональность «Поддержка структуры университета»: запрос на получение списка факультетов, запрос на получение списка курсов внутри выбранного факультета, запрос на получение списка кафедр внутри выбранного курса данного факультета, запрос на получение списка групп данной кафедры данного курса данного факультета, запрос на формирование списка студентов;

реализующих функциональность «Формирование проектов приказов»: запрос на получение списка студентов, включенных в данный приказ, запрос на добавление студента в приказ, запрос на перемещение студента между разделами проекта приказа, запрос на определение атрибутов сущности, принадлежащих приказу, запрос на установку атрибутов сущности принадлежности к приказу;

реализующих процедуру ввода в действие проекта приказа: «заморозка» приказа, «разморозка» приказа, ввод проекта приказа в действие;

реализующих процедуру накопления и предоставления прав доступа к приказам в системе «Контингент»: запрос на получение списка приказов;

реализующих процедуру сопровождения личных карточек студен-

тов: запрос на получение вкладки личной карточки студента.

Большинство из перечисленных запросов являются сложными, они включают функции агрегирования и соединение нескольких таблиц (до 5), на основании которых формируются подзапросы.

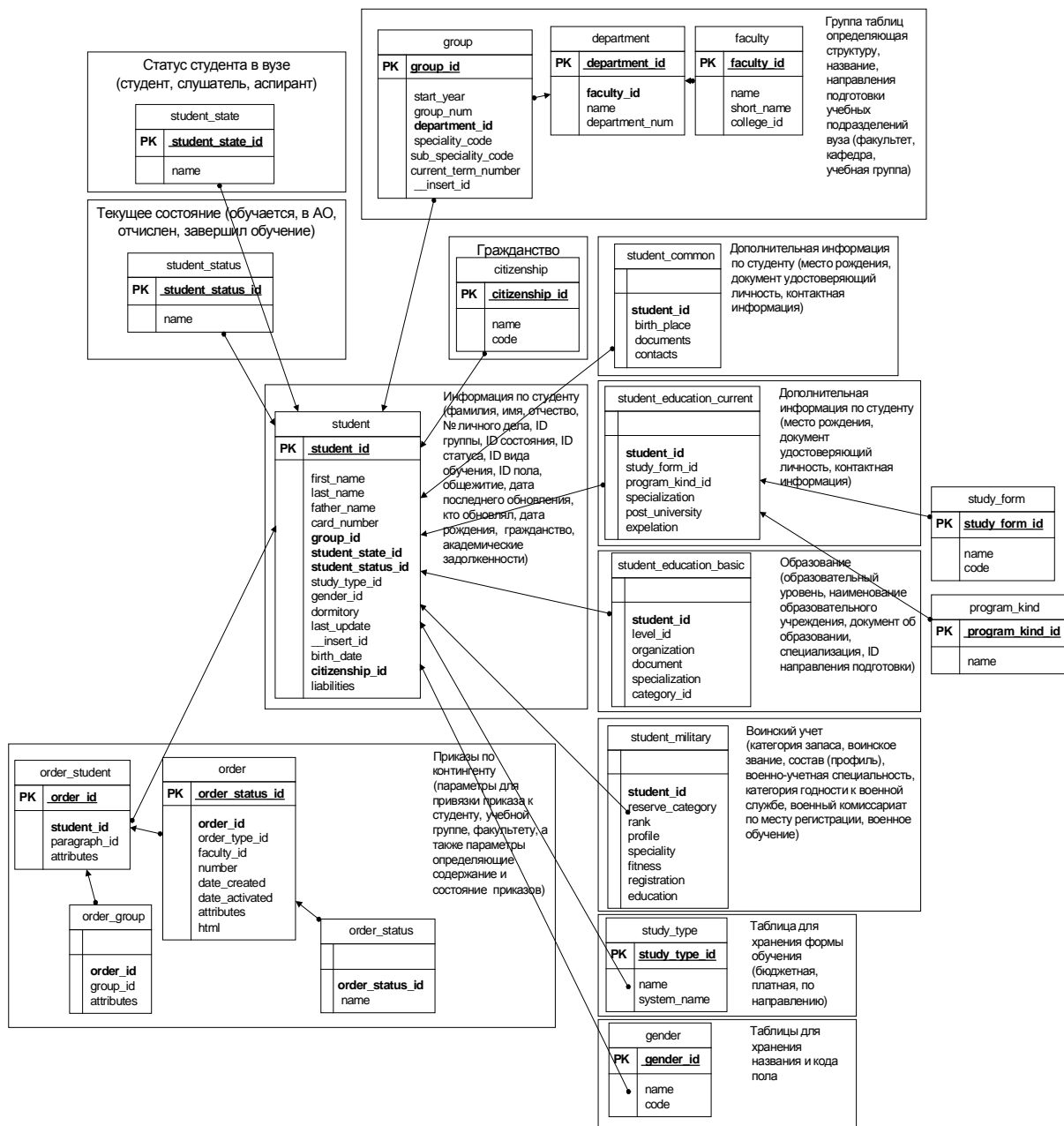


Рис. 2. Схема базы данных «Контингент».

Ниже дано описание аппаратно-программной платформы системы «Контингент». Сервер приложений и сервер БД физически располагаются на одном компьютере серверного класса Proliant ML530G2T 2P, характеристики которого приведены в табл. 1

Таблица 1

Процессор	2 x 2.8 Ghz Pentium III Xeon 512k Cache
Оперативная память	8Gb DDR = 4x2Gb
Контроллеры жестких дисков	Integrated ULTRA-3 SCSI Controller SMART 5302 Array Controller 5 x SMART 5304 Array Controller
Жесткие диски	21x18.2Gb SCSI Drive 8x36.4Gb SCSI Drive
Общий объем жестких дисков	3891.4 Gb
Операционная система	Linux Red Hat
Web-сервер	Appach
Сервер приложений, язык серверных скриптов	Radical, Ruby
СУБД	MySQL 4.0

Компьютеры, используемые в качестве рабочих станций, имеют характеристики, которые приведены в табл. 2.

Таблица 2

Процессор	700Mhz Pentium III 128k Cache
Оперативная память	256Mb
Контроллеры жестких дисков	Контроллеры IDE
Жесткие диски	40Gb SCSI Drive
Операционная система	Windows XP SP2
Web-браузер	Internet Explorer 6.0

### Результаты моделирования системы «Контингент»

ИС «Контингент» представлена сервером СУБД, Web-сервером и сервером приложений, а также рабочими станциями (терминалами). Физически сервер СУБД, Web-сервер и сервер приложений расположены на одной станции. Запросы с терминалов передаются по сети на сервер приложений, обрабатываются, а их результаты передаются обратно на терминалы. При необходимости сервер приложений обращается к СУБД.

Для описания параметров ИС «Контингент» в среде КСАМ были выполнены следующие действия: описана схема базы данных (см. рис. 2), описаны наиболее важные запросы и транзакции, описана архитектура сети, к которой подключен модуль «Контингент», распределены транзакции и таблицы по узлам сети, определены транзакции, к которым выполняется обращение из узлов.

*Анализ нагрузочной способности ИС «Контингент».* При моделировании число рабочих станций было увеличено с 1000 до 3000. Таким образом, моделировалась ситуация, когда к информационной системе могут обращаться преподаватели (особенно при вводе результатов сессии) и студенты.

На рис. 3 представлены графики изменения загрузки «Сервера БД» (подсистемы ввода/вывода с учетом наличия общесистемного буфера), Канала 1 и Канала 2 от числа активных рабочих станций.

Канал 1 моделирует сеть между узлами «Сервер приложений» и «Терминал». Тип сетевого канала – порт коммутатора на 100 Мбит/с, к которому подключен сервер модуля «Контингент». Канал 2 (1000 Мбит/с) моделирует передачу данных между задачами «Сервер приложений» и «Сервер БД» с учетом времени задержки на переключение задач.

Как видно из рисунка, «узким местом» является «Сервер БД».

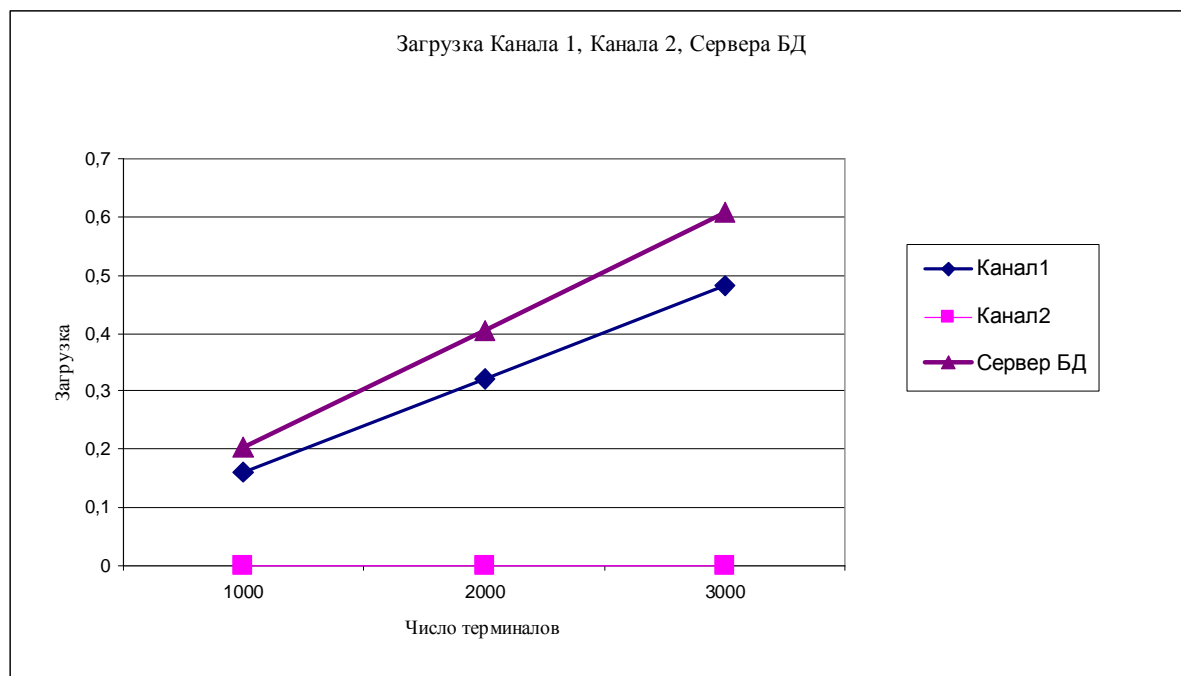


Рис. 3. Изменение загрузки «Сервера БД», Канала 1 и Канала 2 от числа активных рабочих станций.

Причем число 3000 рабочих станций является предельной величиной. При увеличении количества рабочих станций загрузка «Сервера БД» превысит значение 0.6, что приведет к резкому увеличению времени реакции системы [1].

На рис. 4 показано распределение загрузки «Сервера БД» по транзакциям. Можно заметить, что загрузка «Сервера БД» почти равномерно распределена по 5 транзакциям, которые предназначены для получения различных списков. Относительно большая загрузка «Сервера БД» при выполнении каждой из этих транзакций объясняется тем, что при реализации соответствующего запроса происходит соединение 4 таблиц. При этом из базы читаются, как показывает моделирование, около 400 блоков. Если агрегированные данные (число студентов) не определять каждый раз, а хранить их в записях таблиц Faculty (факультет), Department (кафедра), Group (группа), Student (студент), то загрузку «Сервера БД» можно уменьшить почти в 200 раз. Обновление агрегированных данных в каждой

из перечисленных таблиц легко реализовать с помощью триггеров СУБД. Освободившиеся ресурсы сервера можно использовать для увеличения его нагрузочной способности или для расширения функциональных возможностей.

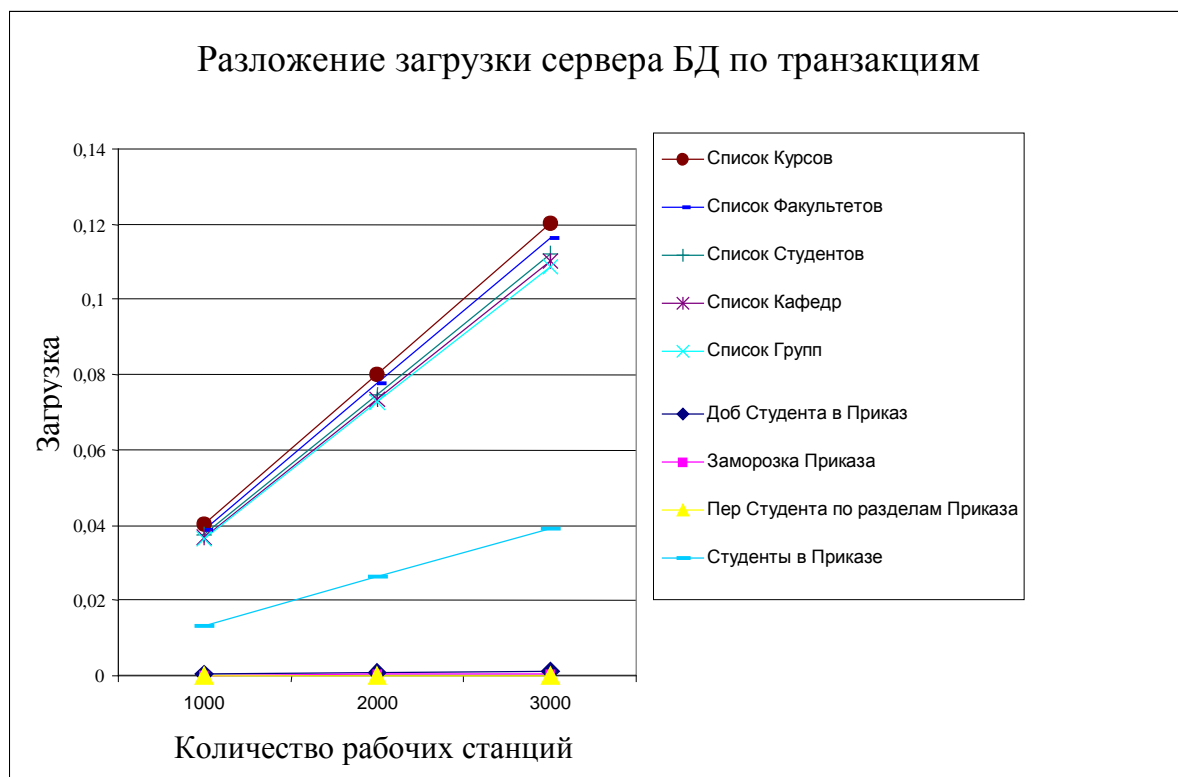


Рис. 4. Распределение загрузки «Сервера БД» по транзакциям.

Загрузка Канала 1 (порта коммутатора, к которому подключен сервер) увеличивается от 0.16 до 0.48 (см. рис. 3). Применение коммутатора здесь оправдано. Реализация входной сети на концентраторах привела бы к снижению максимального числа рабочих станций до 2000, т. к. предельная нагрузка сети Ethernet с общей шиной равна примерно 0.32 [1]. В этом случае потенциальным «узким местом» становится шина сети (см. рис. 3).

Распределение загрузки Канала 1 по транзакциям приведено на рис. 5. Большая доля загрузки Канала 1 приходится на транзакцию «Студенты в приказе». Анализ соответствующего запроса показывает, что при его выполнении из базы данных читаются и передаются на рабочую станцию значения атрибутов 5 записей таблиц, включенных в запрос: Order\_student (приказ, связанный со студентом), Faculty (факультет), Department (кафедра), Group (группа), Student (студент). Но в таблице Order\_student имеется поле attributes длиной 200 Кбайтов, передача которого и вызывает большую нагрузку Канала 1. Уменьшение длины этого поля до 20 Кбайтов позволяет снизить нагрузку Канала 1 в 10 раз.

На рис. 6 представлены графики зависимостей времени выполнения транзакций от числа активных рабочих станций. Здесь приведено время реакции системы на различные запросы без учета фоновой загрузки сетей

и сервера.



Рис. 5. Распределение загрузки Канала 1 по транзакциям.

На время выполнения транзакции «Студенты в приказе» влияние оказывает загрузка Канала 1 (см. рис. 3 и 5). Следует также отметить, что время реакции системы резко возрастает при превышении нагрузки своего порогового значения (в данном случае, если количество рабочих станций будет больше 3000).

*Анализ влияния наполнения базы данных системы «Контигент».* Целью данной серии модельных экспериментов являлась проверка временных характеристик системы при увеличении объема базы данных, связанного с сохранением информации о студентах, которые завершили обучение в университете. Это потребовало задания и изменения большого числа параметров модели. Было определено количество записей в таблицах для случаев, когда информация о студентах не хранится после их обучения (×1), хранится 6 лет (×2), 24 года (×5) и 54 года (×10). Были увеличены мощности первичных и внешних ключей таблиц (кроме таблиц Faculty и Department), а также мощности других полей, число значений которых зависит от количества записей. Для наиболее ресурсоемких запросов были учтены операции группирования и агрегирования. Таким образом, фактор «Число записей» является качественным и включает в себя много параметров. Условно значения этого фактора равны 1, 2, 3, 4. При моделировании число рабочих станций равнялось 1000.

На рис. 7 приведены загрузки «Сервера БД», Канала 1 и Канала 2 при увеличении наполнения базы данных. Как видно из рисунка, наиболее



загруженным ресурсом является «Сервер БД».

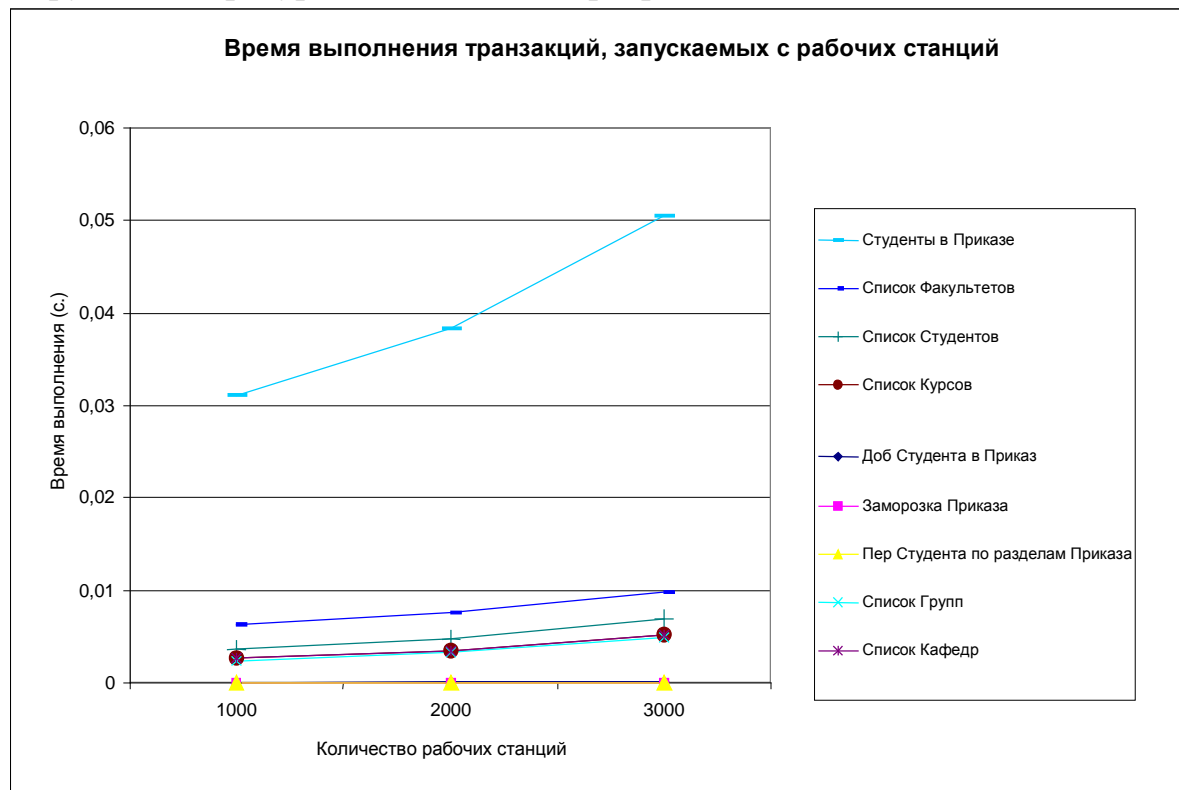


Рис. 6. Время выполнения транзакций, инициируемых с рабочих станций.

На рис. 8 представлено распределение загрузки «Сервера БД» по транзакциям. Она равномерно распределена по 5 транзакциям, которые выполняют чтение различных списков. При реализации каждого из этих запросов выполняется соединение 4 таблиц. Хотя число блоков, обработанных в процессе соединения, возросло с 400 до 4000 (т.е. в 10 раз), загрузка сервера увеличилась с 0.2 до 0.4 (т.е. только в 2 раза). Это объясняется тем, что число дисков, на которых хранится база данных, возросло в 5 раз (процент заполнения диска зафиксирован на уровне 30%). Распределение данных по разным дискам особенно важно для таблицы Student, так как основная доля соединяемых записей читается из этой таблицы. Нагрузку на внешнюю память можно уменьшить при чтении списков (и в результате сократить число дисков), если агрегированные данные хранить в записях таблиц (см. анализ нагрузочной способности ИС).

Загрузка Канала 1 невелика (см. рис. 7). Это объясняется наличием в запросах операций группирования и пересылкой на рабочие станции, в основном, агрегированных данных. Большая доля загрузки Канала 1 приходится на транзакцию «Студенты в приказе» (см. анализ нагрузочной способности системы, приведенный выше).

На рис. 9 приведена зависимость времени выполнения транзакций, инициируемых с рабочих станций, от наполнения базы данных. Выбранная современная конфигурация технических средств, невысокая загрузка сервера и сети определяют небольшое время реакции системы. Относительно

большое время выполнения транзакции «Студенты в приказе» объясняется большим объемом выходных данных, передаваемых по Каналу 1.

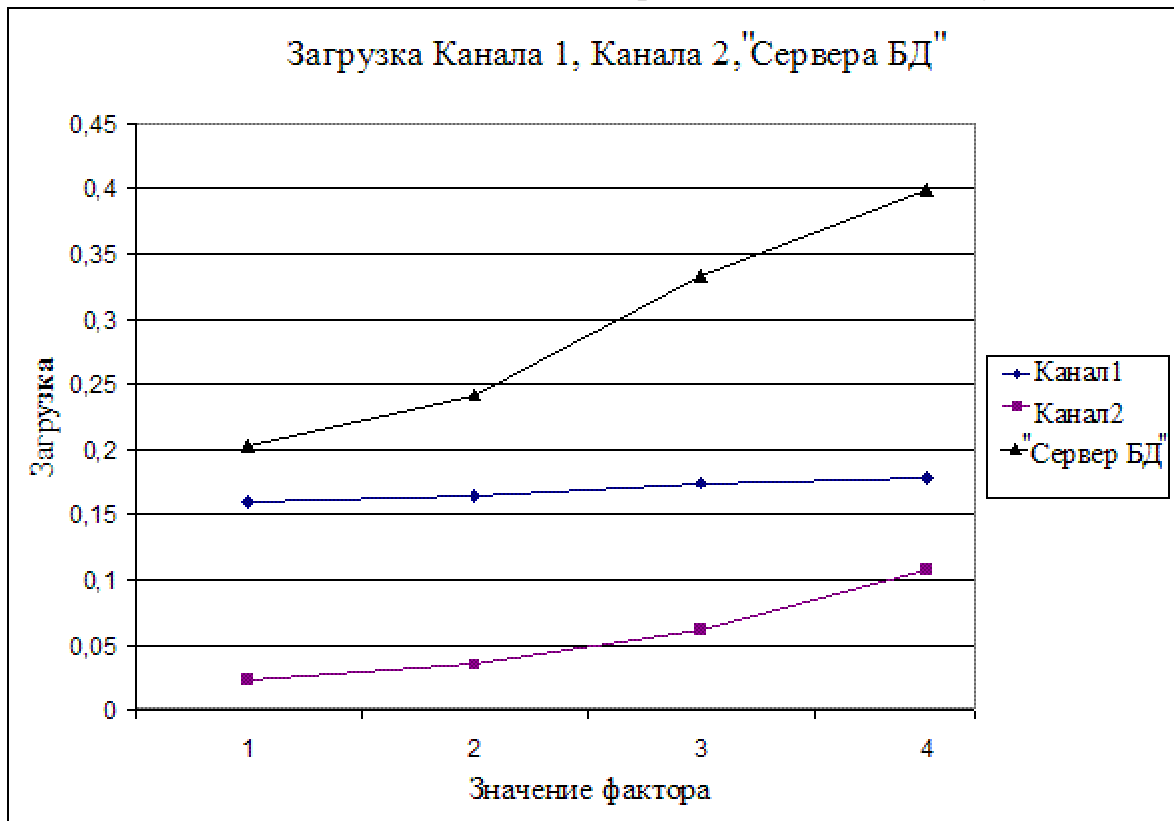


Рис. 7. Загрузка «Сервера БД», Канала 1 и Канала 2 при увеличении наполнения базы данных.

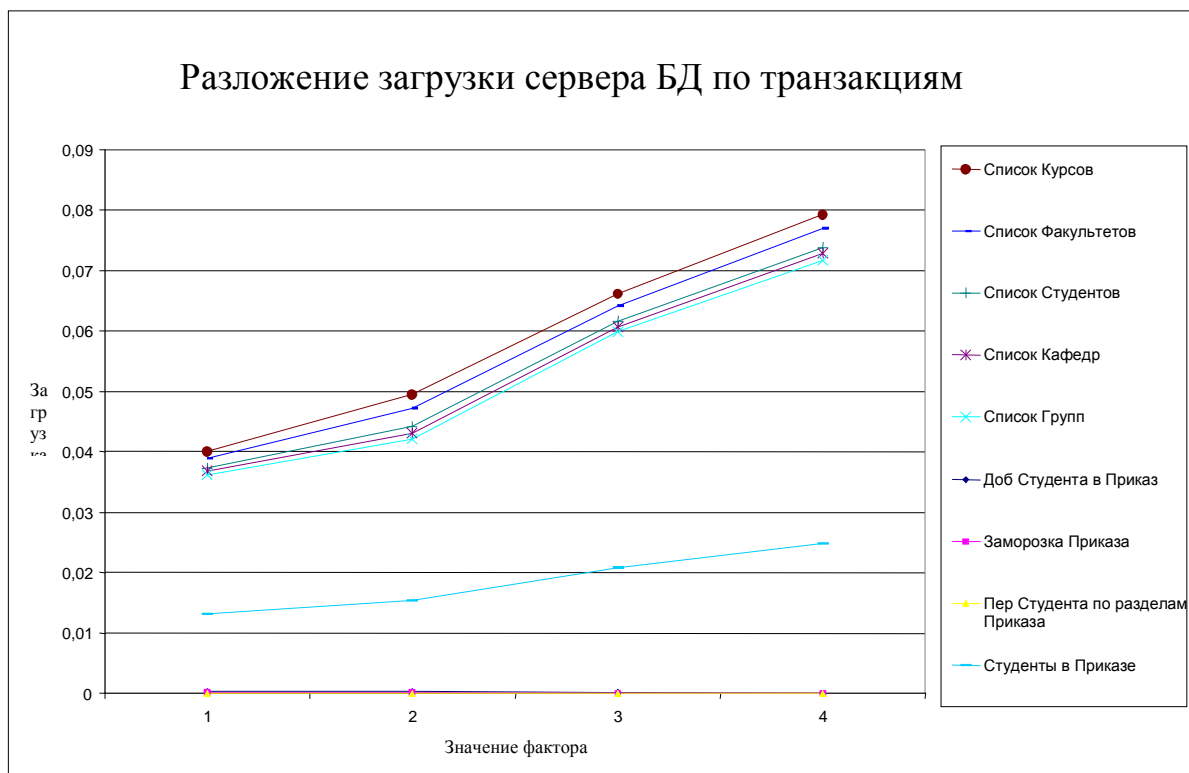


Рис. 8. Распределение загрузки «Сервера БД» по транзакциям при увеличении наполнения базы данных.

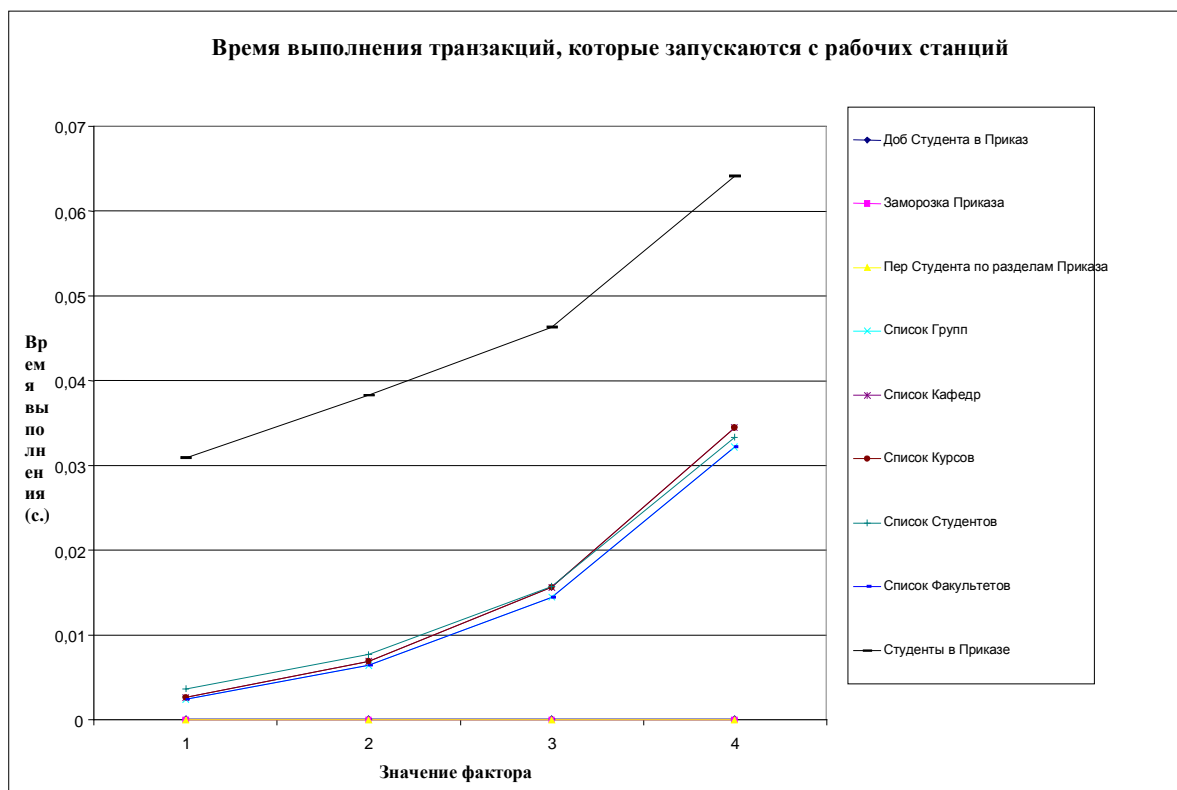


Рис. 9. Зависимость времени выполнения транзакций от наполнения базы данных.

## Заключение

1. С помощью КСАМ выполнен анализ нагрузочной способности системы «Контингент» при увеличении числа рабочих станций с 1000 до 3000. Здесь «узким местом» является «Сервер БД», его загрузка увеличивается с 0.2 до 0.6. Если агрегированные данные (число студентов) не определять каждый раз, а хранить их в записях таблиц и использовать триггеры СУБД, то загрузку «Сервера БД» можно уменьшить почти в 200 раз, что позволяет увеличить нагрузочную способность сервера или расширить его функциональные возможности. Использование коммутатора здесь оправдано, т.к. применение входной сети на концентраторах привело бы к снижению максимального числа рабочих станций с 3000 до 2000. Уменьшение длины поля attributes (200 Кбайтов) таблицы Order\_student до 20 Кбайтов позволяет снизить загрузку Канала 1 в 10 раз.

2. В среде КСАМ выполнен анализ влияния наполнения базы данных системы «Контингент» (число записей таблиц было увеличено в 10 раз). При реализации каждого из 5 запросов выполняется соединение 4 таблиц. Хотя число блоков, обработанных в процессе соединения, возросло с 400 до 4000 (т.е. в 10 раз), загрузка сервера увеличилась с 0.2 до 0.4 (т.е. только в 2 раза). Это объясняется тем, что число дисков, на которых хранится база данных, возросло в 5 раз (процент заполнения диска зафиксирован на уровне 30%). Распределение данных по разным дискам особенно важно

для таблицы Student, так как основная доля соединяемых записей читается из этой таблицы.

#### ЛИТЕРАТУРА

1. *Григорьев Ю.А., Плутенко А.Д.* Теоретические основы анализа процессов доступа к распределенным базам данных. – Новосибирск: Наука, 2002.