

Разработка концептуальной структуры ядра медицинской информационной системы

А.А. Вакалюк
avakalyuk@yandex.ru

С.Н. Басманов
seregabasmanov@rambler.ru

Уральский государственный университет путей сообщения (Екатеринбург)

Аннотация

Разработана концептуальная структура ядра медицинской информационной системы для решения задач повышения производительности и масштабирования информационных систем. Предложен подход к разделению МИС на разделы, для каждой из которых были проанализированы основные требования. В ходе исследования была разработана структурная модель МИС с помощью CASE-средства AllFusion Erwin DataModeler, позволяющего автоматизировать множество трудоёмких задач, уменьшить сроки создания высококачественных и высокопроизводительных транзакционных баз данных и хранилищ данных.

Ключевые слова: CASE-средство AllFusion Erwin DataModeler; ядро МИС; БД; структурная модель.

Введение

Основным этапом разработки медицинской информационной системы является выбор СУБД. На данном этапе, в первую очередь, необходимо определиться с такими критериями, влияющими на эффективность использования БД как хранилища данных, как количество одновременно работающих в системе пользователей и объем вносимой в БД информации.

Современные ЛПУ накапливают большие объемы информации. В день количество человек, проходящих через них, может достигать десятков тысяч. Данная тенденция не связана с большим ростом объема базы данных пациентов, но способствует соразмерному увеличению БД исследований и результатов.

В связи с этим рациональным решением будет организация БД МИС из нескольких БД, каждая из которых будет отвечать за определенный раздел системы, имеющий свою собственную структуру и выполняющий определенные функции. За разделы БД принято использовать основополагающие разделы МИС: пациенты, результаты, исследования, нормативные и прочие документы.

Актуальность исследования обусловлена необходимостью повышения производительности медицинской информационной системы на основе совершенствования ее ядра. Целью исследования является разработка концептуальной структуры ядра медицинской информационной системы.

1 Анализ требований к разделу МИС пациенты

Раздел МИС пациенты является наиболее важным из всех разделов. Главной особенностью данного раздела является наличие большой БД с персональными данными, по которой идет непрерывный поиск. От

Copyright © by the paper's authors. Copying permitted for private and academic purposes.

In: G.A. Timofeeva, A.V. Martynenko (eds.): Proceedings of 3rd Russian Conference "Mathematical Modeling and Information Technologies" (ММИТ 2016), Yekaterinburg, Russia, 16-Nov-2016, published at <http://ceur-ws.org>

того, насколько правильно будет выбрана СУБД для данного раздела, будет зависеть полнота получаемой информации, скорость и корректность ее отображения. При этом чтение данных из базы должно быть мгновенным или занимать минимальное время. Данный показатель в первую очередь влияет на лояльность пользователя к системе.

Одним из основных критериев хранения данных в разделе пациенты является оптимизация, ярким примером которой является сжатие данных. Оно позволяет повысить скорость обращения к данным, уменьшить их размер на физическом диске, а также повысить производительность работы всей системы.

Таким образом, данные, хранящиеся в медицинской информационной системе, должны обладать полнотой информации, корректностью, высокой скоростью загрузки, чтение данных из базы должно происходить мгновенно или занимать минимальное количество времени, а хранение данных должно быть оптимизировано.

Опираясь на данные требования, было принято решение использовать для раздела МИС пациенты колоночную СУБД, одним из представителей которой является СУБД Infobright's Community Edition (ICE).

Отступление от традиционных, реляционных СУБД связано в первую очередь с тем, что СУБД, использующие колоночный принцип хранения данных, показывают лучшую производительность на аналитических нагрузках, таких, которые возникают в МИС. Повышение производительности в первую очередь влияет на скорость и время загрузки информации из базы данных.

Разница в производительности связана с особенной структурой колоночных СУБД. Физически данные в колоночном формате представляют собой совокупность колонок, каждая из которых представлена как отдельная таблица из одного поля. С точки зрения хранения данных на жестком диске, значения одного поля хранятся последовательно друг за другом. Это способствует тому, что при выполнении запросов на выборку данных, колоночные СУБД совершают меньше обменов с дисками, т.к. считываются значения только тех атрибутов, которые упоминаются в запросе.

В свою очередь, СУБД, хранящие записи в построчном виде предполагают физическое хранение всей строки таблицы в одной записи, поля которой идут последовательно друг за другом, а за последним полем идет первое поле следующей записи. Указатели на сами записи записываются в обратном порядке, начиная с конца. При этом каждая запись хранит все атрибуты таблицы. Данный метод хранения данных является классическим и используется с 1970-х годов [1].

2 Анализ требований к разделам МИС результаты, исследования, нормативные и прочие документы

Раздел МИС исследования представлен справочником, формируемым на начальном этапе разработки МИС, использующимся в основном только на чтение. В связи с этим, рациональным решением будет хранение его в колоночной СУБД ICE.

Раздел МИС результаты, в виду своего нечастого редактирования, большого объема и необходимости в мгновенной загрузке, в качестве основной СУБД будут так же использовать СУБД ICE. Единственным различием, по сравнению с разделом пациенты является тот фактор, что у них первоначально запись будет осуществляться во временную строчную, реляционную СУБД MySQL. Связано это в первую очередь с тем, чтобы уменьшить нагрузку на сервер с СУБД ICE, тем самым увеличив скорость работы всей системы и распараллелить потоки данных. Основной поток данных, связанных с разделом исследования и пациенты, будет происходить в момент простоя системы.

Выбор временной СУБД обусловлен необходимостью во временном хранилище для данных, поступающих с приборов. А выбор непосредственно СУБД MySQL выражен ее основными преимуществами, среди которых: простота в работе, богатый функционал, безопасность, масштабируемость, относительная скорость работы, бесплатное использование, открытый код и небольшой объем занимаемого места на физическом носителе.

Таким образом, разделы результаты для хранения данных используют две СУБД: основную и временную, организация которых позволит уменьшить нагрузку на МИС, повысить скорость работы системы и распараллелить потоки данных. Раздел нормативные и прочие документы предполагает хранение отчетных документов, шаблонов результатов заказов, шаблонов выборок. В качестве исходной БД рассматривались 3 принципа хранения информации:

1. Реляционная СУБД с возможностью хранения документов в БД в поле типа BLOB;
2. Хранение в реляционной БД ссылки на репозиторий с документами;

3. Документо-ориентированная СУБД MongoDB.

Принцип хранения документов в БД, используя первый принцип, имеет ряд существенных недостатков:

- запросы на чтение файлов требуют его загрузки в ОЗУ, что влияет на работу не только СУБД и сервера;
- для загрузки файла из БД необходимо использовать дополнительные скрипты, осложняющие работу системы;
- добавление документов в БД влечет увеличение размера самой БД.

Принцип хранения в реляционной БД ссылки на репозиторий с документами также имеет ряд недостатков:

- файлы не удаляются при удалении соответствующей записи БД;
- проблемы при одновременной попытке обновления файла;
- нарушение синхронизации между БД и файловой системой при откате транзакции;
- при резервном копировании и восстановлении информации в БД может возникнуть рассинхронизация с файловой системой;
- файлы не подчиняются ограничениям доступа, наложенным с помощью БД.

Что касается NoSQL БД MongoDB, то она так же имеет свои недостатки, но если рассматривать ее с точки зрения хранения отчетных документов, шаблонов результатов заказов, шаблонов выборок, то здесь она идеально подходит. При этом благодаря специальной технологии GridFS [2], MongoDB может хранить файлы абсолютно любых размеров, присваивая каждому загруженному объекту свои метаданные, ограничивающихся в основном физическим местом на жестких дисках сервера. А благодаря документо-ориентированной модели, MongoDB работает быстрее, обладает лучшей масштабируемостью и ее легче использовать по сравнению с другими СУБД.

Таким образом, раздел нормативные и прочие документы представлен NoSQL БД MongoDB, что позволит хранить файлы различных размеров, увеличит скорость работы системы и повысит ее масштабируемость.

Ключевой особенностью такой системы, состоящей из трех логически разных СУБД, является организация взаимодействия внутри системы. В целях распараллеливания данных, уменьшения нагрузки на сервер и ускорения работоспособности системы, было принято решение физически разделить ядро МИС на три сервера. На каждом из этих серверов будет установлена одна из трех СУБД, соответствующая основным разделам МИС.

Для лучшего взаимодействия между серверами и высокой скоростью доступа к данным необходимо их объединить в единую гигабитную сеть.

Таким образом, ядро МИС представлено тремя серверами: с колоночной СУБД ICE, с документо-ориентированной СУБД MongoDB и с реляционной СУБД MySQL, что позволит распараллелить потоки данных, уменьшить нагрузку на сервер и повысить скорость работы МИС в целом.

3 Разработка структурной модели ядра МИС

Структурная модель разрабатываемой МИС, использующей три сервера представлена на рисунке 1.

Структура ядра МИС формируется с помощью AllFusion Erwin DataModeler – CASE-средства для проектирования баз данных, которое позволяет создавать, документировать и сопровождать БД [3].

Внедрение данного CASE-средства позволит:

1. Сводить данные из нескольких платформ в единую модель с их графическим представлением;
2. Создать каталог массивов данных, который может храниться в центральной репозитории, что приведет к заданию стандартов и снижению избыточности;
3. Обеспечить интеграцию с другими приложениями, что приведет к автоматизации множества задач, таких как создание и настройка отчетов, импорт и экспорт моделей из других программ.

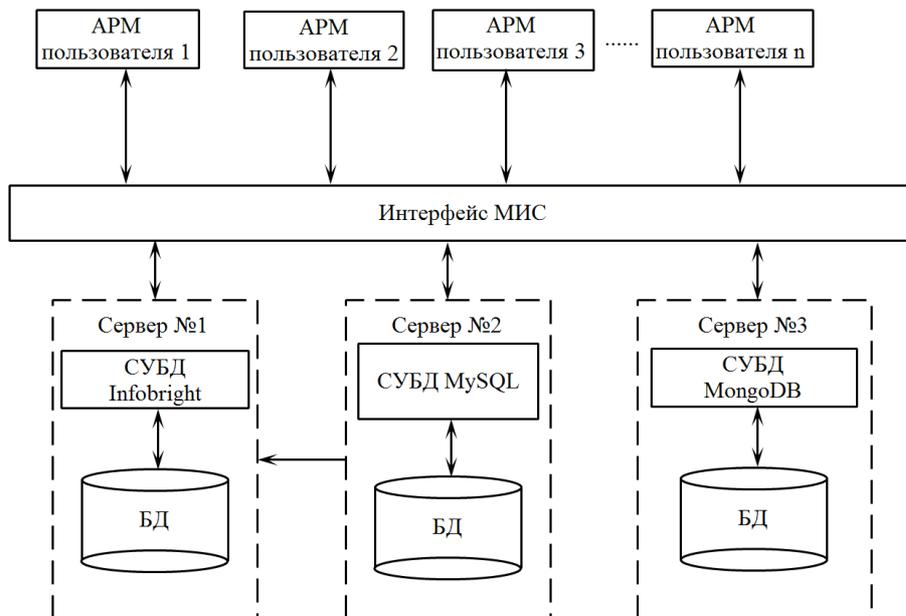


Рис. 1: Структурная модель медицинской информационной системы

Таким образом, организация структуры БД с помощью средства Data Modeler позволит автоматизировать множество трудоёмких задач, уменьшить сроки создания высококачественных и высокопроизводительных транзакционных баз данных и хранилищ данных [4]. Внедрение данного средства улучшит коммуникацию организации, обеспечит представление комплексных активов данных в удобном для понимания и обслуживания формате.

Основная деятельность МИС базируется на двух основных принципах работы с БД: запись и чтение. В зависимости от операций, производимых пользователями, происходит выполнение этих принципов или их совокупностью.

Выводы

1. Разработан подход к построению структуры ядра медицинской информационной системы на основе объединения колоночной и строчной СУБД, что позволяет повысить скорость выполнения процедуры чтения данных из СУБД, оптимизировать хранение данных в БД МИС, обеспечить корректность и полноту получаемой информации;
2. Рассмотрена организация структуры БД МИС с помощью средства Data Modeler, использование которой позволит автоматизировать множество трудоёмких задач, уменьшить сроки создания высококачественных и высокопроизводительных транзакционных баз данных и хранилищ данных. Внедрение данного средства улучшит коммуникацию организации, обеспечит представление комплексных активов данных в удобном для понимания и обслуживания формате;
3. Предложен подход к повышению масштабируемости МИС и хранению файлов различных размеров в СУБД, основанный на использовании документо-ориентированной СУБД MongoDB.

Список литературы

- [1] Ju.S. Izbachkov, V.N. Petrov, A.A. Vasil'ev, I.S. Telina. *Informacionnyye sistemy: uchebnyk dlja vuzov. 3-e izdanie* [Information systems: Tutorial for institutes of higher education. The 3-th edition]. SPb., Piter, 2011. (in Russian) = Ю.С. Избачков, В.Н. Петров, А.А. Васильев, И.С. Телина. *Информационные системы: учебник для вузов. 3-е изд.* СПб., Питер, 2011.
- [2] MongoDB. *docs.mongodb.com/manual/core/gridfs*. (in Russian) = Сайт MongoDB / Режим доступа URL: <https://docs.mongodb.com/manual/core/gridfs/> Дата обращения 10.11.2016.

- [3] V.I. Gorbachenko, G.F. Ubiennyh, G.V. Bobrysheva. *Proektirovanie informacionnyh sistem s CA ERwinModelingSuite 7.3: Ucheb. Posobie* [Designing information-analytical systems CA ERwinModelingSuite 7.3: Tutorial]. Penza, Publishing PGU, 2010. (in Russian) = В.И. Горбаченко, Г.Ф. Убиенных, Г.В. Бобрышева. *Проектирование информационных систем с CA ERwinModelingSuite 7.3 : учебное пособие*. Пенза, Изд-во ПГУ, 2012.
- [4] A.A. Vakalyuk, S.N. Basmanov, A.A. Basmanova. The conceptual structure development of the automatic control system for medical diagnostic enterprise. *Automation. Modern technologies*, 2, 2016. (in Russian) = А.А. Вакалюк, С.Н. Басманов, А.А. Басманова. Разработка концептуальной структуры автоматизированной системы управления медицинским диагностическим предприятием. *Автоматизация. Современные технологии*, 2, 2016.

Core conceptual structure development of medical information system

Andrey A. Vakalyuk

Ural State University of Railway Transport (Yekaterinburg, Russia)

Sergey N. Basmanov

Ural State University of Railway Transport (Yekaterinburg, Russia)

Abstract. A core conceptual structure of medical information system is developed for problem solving of information system performance and scaling. An approach is suggested for MIS decomposition on categories, for each of them basic requirements are analyzed. In the course of research a structural model of MIS is developed by AllFusion Erwin DataModeler CASE tool, who enables to automate greate number labor-intensive tasks, to decrease periods of high-quality and high-performance transactional database and data warehouse creation.

Keywords: AllFusion Erwin DataModeler CASE tool, core of MIS, database, structural model.