



В рамках данной работы, учитывая особенности логического проектирования и физического хранения баз данных, планируется рассмотреть особенности реализации ограничений целостности данных и манипулирования данными в реляционных и объектных базах данных, реализуя различные по сложности запросы с оценкой времени их выполнения.

Литература

1. Коннолли, Т. Базы данных. Проектирование, реализация и сопровождение. Теория и практика. Третье издание. [Текст] / Т. Каннолли, К. Бегг. – Москва, Санкт-Петербург: изд-во Вильямс, 2003. – 1440 с.
2. InterSystems. Caché: Современная система управления базами данных [Электронный ресурс]. – <http://www.intersystems.com/ru/our-products/cache/cache-overview/>

Е.И. Чигарина, Ю.С. Чуркина

МЕТОДИКА СРАВНИТЕЛЬНОГО АНАЛИЗА СРЕДСТВ ГЕНЕРАЦИИ ОТЧЁТОВ В РЕЛЯЦИОННЫХ БАЗАХ ДАННЫХ

(Самарский национальный исследовательский университет
имени академика С.П. Королева)

Генераторы отчётов представляют собой специализированные средства, позволяющие представить информацию, хранящуюся в информационной системе, в виде документа или отчёта, который можно распечатать или сохранить в различных электронных форматах [1].

В настоящее время на рынке представлено большое количество систем генерации отчётов, и пользователю приходится тратить значительное время для того, чтобы понять какое из средств дает наибольшую функциональность и гибкость, соответствующую сложности и требованиям отчёта. К тому же непонятно как именно выбирать и что сравнивать. Можно, например, сравнить по размеру, который добавится к приложению при установке средства генерации. Но есть такие генераторы отчётов, которые к приложению ничего не добавляют, а представлены отдельным dll файлом, размеры которого гораздо больше самого приложения. Или же к исполняемому файлу добавляется немного, зато файлы сформированных отчётов в несколько раз больше самого приложения. Может проводить сравнение по удобству использования? Но в чем его измерить?

В связи с этим главной задачей данной работы является проведение сравнительного анализа использования средств генерации отчётов в реляционных базах данных. В работе предлагается следующая методика сравнения:

1. Выбор универсальных средств создания отчётов, обладающих максимальными функциональными возможностями для формирования и генерации отчётов. Среди функциональных возможностей были выделены следующие:
– допустимые виды отчётов;



- наличие мастера построения отчётов любой сложности (нестандартных отчётов);
- возможность предварительного просмотра печати готового отчёта;
- наличие средств работы с графикой (вставка изображений, объектов типа линия, квадрат и т.д.).
- поддержка различных видов СУБД;
- возможность вывода на печать;
- возможность перевода отчёта в различные форматы данных (Word, Excel, PDF, HTML).

В качестве средств генерации отчётов были выбраны Crystal Reports, Fast Reports и Reporting Services, так как они обладают схожими функциональными возможностями.

2. Выбор различных отчётов, наиболее часто встречающихся в системах генерации отчётов. В результате классификации были выделены простые отчёты, сложные отчёты с группировкой, отчёты с вычислениями, параметрические отчёты, перекрёстные отчёты, графические отчёты [2].

3. Реализация демонстрационной базы данных «Авиаперевозки по России» средствами СУБД реляционного типа (MS SQL Server, PostgreSQL) для последующего формирования и генерации отчётов на основе информации, содержащейся в базе данных [3].

4. Определение критериев времени формирования и времени генерации отчётов, а также параметров, влияющих на эффективность реализации отчётов.

5. Для каждого вида отчёта выбранным средством генерации с помощью выбранной СУБД определение значений критериев с различными значениями параметров.

В качестве параметров, влияющих на критерий времени генерации отчётов каким-либо средством реализации, были выбраны структура базы данных, а именно типы данных, а также объем данных, зависящий от количества записей в таблицах, участвующих в формировании отчёта, а также от количества полей, представленных в отчёте.

Чтобы выяснить время генерации отчётов в зависимости от типа данных были сформированы простые отчёты на основе тестовой базы данных. Сравнение проводилось для наиболее популярных типов данных: числовые (integer), символьные (varchar), временные (date, datetime). Типы varchar и integer сравнивались в двух вариантах с различной длиной строки. Тестовая база данных реализована в виде четырёх таблиц, каждая из которых состоит из двух полей: id и соответствующего типа данных. Таким образом, при реализации отчётов на демонстрационной базе данных, заполненной данными, приближенными к реальным, можно однозначно определить, почему идентичные отчёты формируются с разным временем (например, простой отчёт, вывод из одной таблицы, только разных типов данных). Данные в таблицах тестовой базы данных представлены в объеме 100, 1000, 10000, 100000 и 1000000 записей.

На рисунках 1 – 3 представлены результаты сравнения генерации отчётов в зависимости от типа данных средствами Crystal Report version for Visual



Studio (синяя линия на графике) и Fast Report 5 (красная линия на графике). По оси ординат на графиках обозначается время генерации отчёта в секундах, а по оси абсцисс – количество записей, участвующих в формировании отчёта.

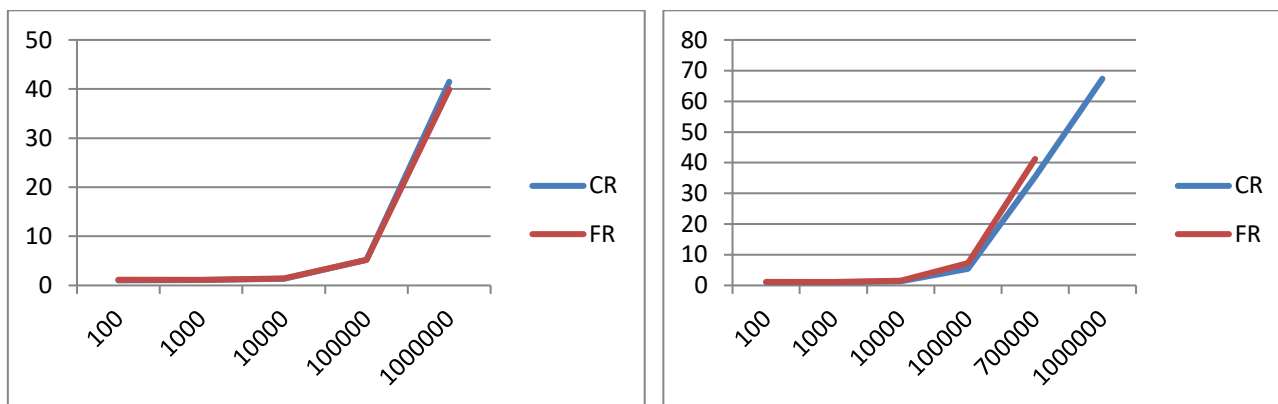


Рисунок 1 – Отчёты с типом данных varchar. Длина строки 60 и 300 символов

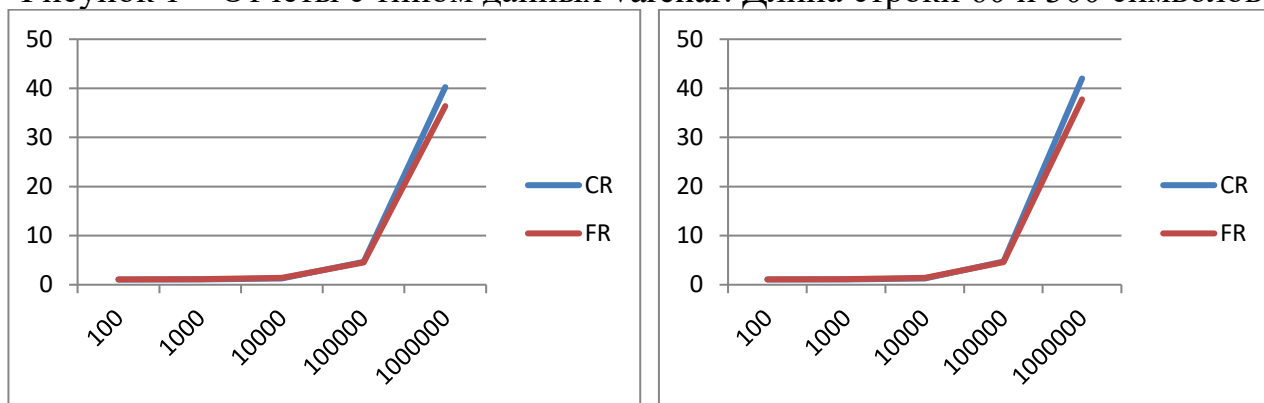


Рисунок 2 – Отчёты с типом данных integer. Длина строки 3 и 10 символов

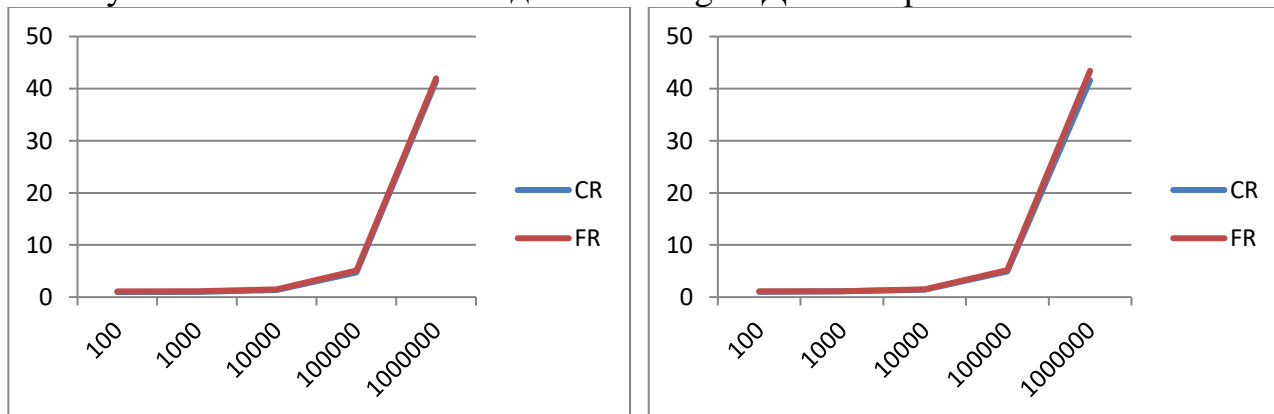


Рисунок 3 – Отчёты с типами данных date и datetime

Из полученных результатов были сделаны некоторые выводы.

Время формирования отчётов для любых типов данных пакетами CrystalReport и FastReport несущественно различаются при объеме записей менее 100000. С увеличением количества записей до 1000000 в каждой таблице, можно наблюдать возрастание скорости формирования отчёта на 10% с помощью FastReport для типа данных integer, и на 13% с помощью CrystalReport для типа данных varchar с длиной строки 300 символов.



Длина строки существенно влияет на время формирования отчёта только для типа данных `varchar` при использовании генератора `CrystalReport`. Если формировать отчёт с помощью `FastReport` различия во времени будут незначительны. Так же незначительны будут различия во времени для типа `integer` при формировании отчёта обоими пакетами.

Отчёты, сформированные с использованием таких данных как дата и дата со временем, практически не различаются по времени формирования ни в пакете `CrystalReport`, ни в пакете `FastReport`. Однако отчёт с типом данных `datetime` формируется быстрее с помощью `FastReport`.

Также в ходе исследования было выяснено, что в генераторе отчётов `CrystalReport` есть ограничение на длину текстового поля в размере 255 символов. В `FastReport` такого ограничения нет. Однако в `FastReport` выявилось ограничение на объем данных выборки. Не удалось сформировать отчёт с типом `varchar(300 символов)` для 1000000 записей с помощью `FastReport`. Максимально допустимое количество записей оказалось равным 700000. При увеличении количества записей система выдает сообщение об ошибке о недостаточном объеме памяти.

На основе информации, содержащейся в демонстрационной базе данных «Авиаперевозки по России», сформируются основные типы отчётов. Для каждого вида отчёта выбранным средством генерации с выбранным средством реализации СУБД определяется время $t_{i,j,k,l}$, затраченное на формирование и генерацию отчётов, где $i = 1..7$ – тип отчёта (1 – простой отчёт из одной таблицы, 2 – простой отчёт из двух таблиц, 3 – простой отчёт на основе вида, 4 – сложный отчёт с группировкой, 5 – сложный отчёт с вычислениями, 6 – графический отчёт, 7 – перекрестный отчёт), $j = 1..3$ – средства генерации отчёта (1 – `Crystal Reports`, 2 – `Fast Reports`, 3 – `Reporting Services`), $k = 1..2$ – СУБД (1 – `PostgreSQL 9.5`, 2 – `MS SQL Server 2014`), $l = 1..5$ – количество записей в отчёте (1 – 100, 2 – 1000, 3 – 10000, 4 – 100000, 5 – 1000000). Для каждого типа отчёта производится по 10 измерений и рассчитывается среднее время формирования документа каждым из генераторов.

После произведенных расчетов составляются графики зависимостей величины $t_{i,j,k,l}$ от числа записей в таблице для каждого типа отчёта различными средствами генерации. На составленных графиках визуально видно, какой из генераторов отчётов является наиболее эффективным при заданных параметрах. На основе проведенных исследований можно сформировать некоторые рекомендации по использованию генераторов отчётов в системах реляционных баз данных.

Литература

1. Определение генератора отчётов [Электронный ресурс]. – <http://www.ngpedia.ru/id639519p1.html>
2. Классификация отчётов [Электронный ресурс]. – <http://compress.ru/article.aspx?id=10013&iid=418>
3. Демонстрационная база данных [Электронный ресурс]. – <https://postgrespro.ru/education/demodb>