

## ЕДИНАЯ РЕГИСТРАЦИОННАЯ БАЗА ДАННЫХ В ВИРТУАЛЬНОМ ОБЛАКЕ ИНФОРМАЦИОННОЙ СИСТЕМЫ УЧЕБНОГО ЗАВЕДЕНИЯ

Донецкий национальный университет

Рассматривается проблема сбора и поддержки в актуальном состоянии регистрационных данных участников образовательного процесса. Предлагается для информационных систем учебных заведений организовывать доступ к регистрационным данным контингента в виде виртуального вычислительного облака, использующего единую регистрационную базу данных.

### *Введение*

Каждое учебное заведение, от школ до ВУЗов, в своей деятельности оперирует большим количеством ежегодно пополняемых и изменяемых регистрационных данных об учениках, студентах, педагогах и сотрудниках. Это — паспортные данные, документы об образовании, о состоянии здоровья, о составе семьи и т.п. Если образовательная организация автоматизирует свои бизнес-процессы, то регистрационные данные заносятся в локальную (корпоративную) регистрационную базу данных (ЛРБД) и эта база становится основой информационной системы организации (ИСО). ИСО представляет собой программно-аппаратную систему, включающую базы данных, системы управления базами данных и прикладные программные средства для автоматизации бизнес-процессов организации.

Каждый абитуриент, студент может неоднократно обращаться в различные учебные заведения, центры тестирования и другие организации для подачи в них документов, передачи контактной информации о себе. При обращении в каждую новую организацию он всё это делает заново. При этом для всех ИСО мы получаем многократный ручной ввод (особенно в период вступительных компаний) и множественное хранение одних и тех же регистрационных данных, причём не всегда актуальных, в первую очередь, контактных.

Создание единой регистрационной базы данных (ЕРБД) в Internet-пространстве позволит заинтересованным организациям хранить, модифицировать и потреблять в реальном времени готовые актуальные регистрационные данные [1, 2]. При этом сами граждане смогут контролировать и дополнять свои данные через Internet, т.е. участвовать в актуализации ЕРБД.

ИСО может и должна работать с этими базами. При этом доступ к локальным (ЛРБД) и обновляемым удалённым данным (ЕРБД) может реализовываться в виде обращения к некоторому виртуальному вычислительному облаку ИСО, включающему и ЛРБД, и ЕРБД. Именно из этого облака ИСО должна получать регистрационные данные своего контингента с учётом актуальности и скорости доступа к ним. Такое облако может существенно снизить затраты организаций на поддержку ЛРБД. Рассмотрим подробнее принципы организации такого облака и особенности его функционирования.

### *ЕРБД в виртуальном облаке ИСО*

На рис. 1 представлено виртуальное вычислительное облако ИСО для доступа к регистрационным данным контингента. Облако имеет локальную и удалённую части. К локальной части облака относится ЛРБД и программно-аппаратные средства on-line маршрутизации запроса к регистрационным дан-

ным организации. К удаленной части облака относятся ЕРБД и Web-сервисы доступа к данным ЕРБД. Удалённая часть облака располагается в сети Internet.

Доступ ИСО к регистрационным данным контингента реализуется по идентификаторам личностей — ID. ИСО на каком-либо этапе своего функционирования получает ID личности, например, из ЛРБД (см. рис. — 0) или из других своих документов. Далее ИСО на основании ID и сопутствующих параметров, формирующих конкретный запрос, обращается к маршрутизатору облака (1), который возвращает запрошенные данные из локальной части облака (2) или обращается к посреднику Web-сервиса (3). В последнем случае,

если канал авторизованного сеанса с Web-сервисом ЕРБД ранее не был открыт, то посредник вызывает метод создания авторизованного сеанса (4), получает подтверждение авторизации (5) и вызывает необходимый метод Web-сервиса ЕРБД (6) для реализации запроса, получает запрошенные данные (7). Если канал авторизованного сеанса был открыт, то посредник сразу вызывает необходимый метод (6). После получения данных, сеанс с Web-сервисом ЕРБД может быть закрыт, полученные данные из удалённой части облака возвращаются в ИСО (8). Безопасность обмена данными с удалённой частью облака обеспечивается поддержкой сетевого протокола SSL.

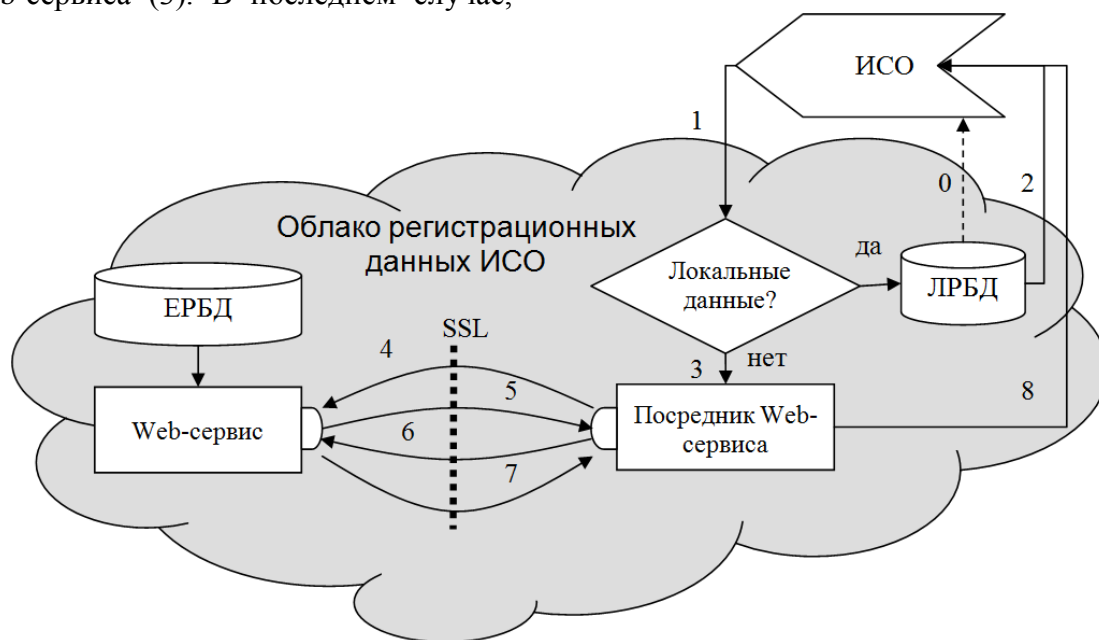


Рис. 1. Виртуальное облако ИСО для доступа к регистрационным данным своего контингента в ЛРБД или в ЕРБД

### ***ЛРБД в виртуальном облаке ИСО***

Совместное использование ЕРБД с ЛРБД, кроме повышения актуальности данных, позволяет снизить объем ЛРБД и отказаться от сложных процедур синхронизации ЛРБД с ЕРБД, а также – от процедур ручного обновления данных контингента организации.

Обсудим возможные модели формирования ЛРБД в рассматриваемом

виртуальном облаке и оценим соответствующие снижения объемов ЛРБД за счёт on-line использования некоторых данных ЕРБД. Пусть вуз имеет 1000 студентов, а система управления базами данных (СУБД) для ЛРБД вуза – это MS SQL. При этом строковые данные – это varchar(15) – 15 байт, целочисленные int – 4 байта, даты date – 4 байта.

Можно выделить четыре основные модели формирования ЛРБД в зависи-

мости от данных, которые в ней хранятся:

- 1) Фамилия, имя, отчество личности (ФИО – три строки), уникальный идентификатор (УИ – строка) и дата рождения (ДР – тип дата);
- 2) ФИО, УИ, ДР, а так же контактные данные личности (E-mail, номера телефонов, адрес проживания и т.д.)
- 3) ФИО, УИ, ДР, контактные данные и основные документы об образовании (сертификаты сдачи единого государственного экзамена (ЕГЭ), аттестат, диплом);
- 4) ЛРБД хранит в себе всю информацию о личности, имеющуюся в ЕРБД.

В первой модели ЛРБД объем данных одной личности будет определяться как

$$V_1 = V_{\text{ФИО}} + V_{\text{УИ}} + V_{\text{ДР}} = 64 \text{ байт} \quad (1)$$

где  $V_{\text{ФИО}} = 45$  байт – объем ФИО личности,  $V_{\text{УИ}} = 15$  байт – объем данных, занимаемый уникальным индексом строкового типа,  $V_{\text{ДР}} = 4$  байта – объем информации о дате рождения.

Во втором способе объем данных

$$V_2 = V_1 + V_{\text{КД}} = 184 \text{ байта} \quad (2)$$

где  $V_{\text{КД}} = 120$  байт – объем контактных данных, представленных восемью полями строкового типа.

Таблица 1. Объем ЛРБД в зависимости от её модели

Модель ЛРБД, $i$	Приблизительный объем ЛРБД, байт	$V_4 / V_i$
1	80 000	6250
2	230 000	2174
3	3 605 000	139
4	500 000 000	1

Как видно из таблицы, использование ЛРБД, построенной по моделям 1-3, позволяет существенно сэкономить

Третий способ

$$V_3 = V_2 + V_{\text{ДОК}}^* = 2884 \text{ байт} \quad (3)$$

где  $V_{\text{ДОК}}^* = 2700$  байт – объем 6 документов (диплом, 4 сертификата ЕГЭ, аттестат). Каждый документ состоит из четырех строк, одной даты и текста содержимого, в результате чего в среднем, объем одного документа составляет  $V_{\text{ДОК}} = 450$  байт.

Четвертый способ предполагает хранение всех данных личности из ЕРБД, т.е.

$$V_4 = V_3 + V_{\text{ДД}} \approx 500000 \text{ байт} \quad (4)$$

здесь  $V_{\text{ДД}} \approx 500$  Кбайт – средний объем всех остальных данных и документов личности, в том числе тех, которые редко используются (информация о составе семьи, медицинские справки и карта, другие сертификаты, наградные документы и т.д.).

В таблице 1 представлены полученные приблизительные объемы ЛРБД, которые рассчитывались по формуле

$$V_i = nV_1 + V_{\text{СД}}, \quad i = 1, 2, 3, 4, \quad (5)$$

где  $n = 1000$  – количество студентов,  $V_{\text{СД}}$  – объем служебных данных, зависит от конкретной СУДБ и ее версии, для MSSQL в среднем составляет 25% от объема базы [3].

на её объёме, не говоря уже об экономии на поддержке регистрационных данных ИСО в актуальном состоянии.

### **Выводы**

Таким образом, мы видим, что облако регистрационных данных, в общем, содержит относительно небольшую ЛРБД и может обеспечить доставку актуальных данных о контингенте организации из ЕРБД без их дублирования в ЛРБД.

Разработчики подсистем ИСО с описанным облаком могут создавать программное обеспечение без жесткой связи с физическими адресами тех или иных данных контингента, что способствует созданию приложений, которые легко модернизируются и масштабируются. Для доступа к данным личности необходимо вызвать соответствующие функции маршрутизатора, вместо прямого SQL-запроса к конкретной базе. Доступ к необходимой базе реализует маршрутизатор через свои функции.

В тоже время, если в ЛРБД полностью копировать всю необходимую информацию из ЕРБД, то у администратора ИСО появится возможность управлять местом выборки данных, оптимизировать скорость доступа и степень актуальности данных посредством настройки маршрутизатора облака. Например, при высокой скорости доступа к ЕРБД и низкой стоимости каналов доступа целесообразно получать

актуальные регистрационные данные непосредственно из ЕРБД. В противном случае следует частично (например, только для контактных данных) или полностью переключиться на ЛРБД, пожертвовав при этом актуальностью некоторых данных.

### **Список использованных источников**

1. Толстых В.К. О единой регистрационной базе граждан и структуре данных личности в этой базе / В.К. Толстых, Л.Н. Киселева // Реєстрація, зберігання і обробка даних. — 2012, Т. 13, № 3. С. 81—91.
2. Пат. 60435 Украина, МПК G06Q 90/00, G06F 17/40. Спосіб контролю та споживання даних єдиної реєстраційної бази даних громадян / Толстих В.К., Киселева Л.Н.; заявитель и правообладатель Донецкий национальный университет, Толстих В.К., Киселева Л.Н. — № u201011883; заявл. 07.10.2010; опубл. 25.06.2011, Бюл. № 12/2011
3. Nielsen P. Microsoft SQL Server 2008 Bible / Nielsen P.; Parui U.; Wiley Pbl. — 2009, 1680p.