

Колокольцева О.М.

УПРАВЛЕНИЕ КОНТЕНТОМ ОБРАЗОВАТЕЛЬНОГО ПОРТАЛА РАСПРЕДЕЛЕННОГО УНИВЕРСИТЕТА

space-olga@rambler.ru

*Южно-Российский государственный технический университет
г. Новочеркасск*

В статье проведен анализ подходов к хранению образовательных информационных ресурсов порталов. Обоснована целесообразность применения концепции комбинированного подхода для портала распределенного университета с большим объемом мультимедийного контента.

The analysis of approaches for educational information resources storage was done in the article. The application expediency of combined approach for distributed university portal with great volume of multimedia content was proved.

Введение. При создании и накоплении информационных ресурсов в ВУЗах необходимо обеспечить интеграцию образовательного контента, созданного различными кафедрами и другими подразделениями, например, сотрудниками Центра новых информационных технологий или филиалов. Поскольку кафедры и подразделения ВУЗа, как правило, являются территориально распределенными, учебный материал может существовать в различных версиях и периодически корректироваться. При таких условиях необходим удобный, быстрый доступ к отдельным информационным ресурсам и транспортировка получателям нужной информации.

Для повышения качества обучения необходимо наполнять курсы мультимедийной информацией, которая значительно улучшает воспринимаемость теоретического материала. Изучение многих дисциплин невозможно без лабораторных работ. Со временем оборудование устаревает, и в большинстве случаев учебное заведение не способно создать установки для их проведения по финансовым или иным причинам. При наличии большого количества мультимедийных учебных материалов возникает проблема их хранения и доставки обучающимся [1, 2]. Качество и скорость передачи учебных материалов по сети зависит от пропускной способности каналов. Необходима разработка стратегий для управления хранением и использования такого контента.

Подходы к хранению контента образовательных порталов. Образовательный контент большинства порталов размещается в *централизованном хранилище*, которое содержит множество образовательных информационных ресурсов (ОИР) $R = \{r_i, i = \overline{1, n}\}$ и их метаданные, а также данные для функционирования интерактивных сервисов.

Организация централизованного хранения образовательного контента реализуется на основе двухзвенной и трехзвенной архитектуры клиент-сервер. *Двухзвенная архитектура клиент-сервер* позволяет создать портал в условиях ограниченного бюджета проекта при относительно небольшом количестве пользователей портала, но не подходит для порталов с большим объемом мультимедийных образовательных ресурсов [5, 6]. *Трехзвенная архитектура кли-*

ент-сервер используется при большом количестве одновременно работающих клиентов, интенсивно генерирующих запросы к базе данных, и больших объемах образовательных ресурсов, предоставляемых пользователям.

При построении порталов на основе двухзвенной и трехзвенной архитектур клиент-сервер объединение образовательных информационных ресурсов в единое хранилище данных может быть затруднительно по следующим причинам:

- на процесс интеграции влияет территориальная распределенность подразделений, поскольку многие ВУЗы являются многофилиальными структурами с множеством кафедр, ведущих подготовку по широкому набору специальностей;
- ограниченность ресурсов жесткого диска сервера баз данных. Так как мультимедийные курсы и видео-файлы имеют большой объем, то не всегда возможно представить все образовательные ресурсы в рамках единого хранилища портала.

Решением вышеуказанных проблем является построение образовательного портала на основе концепции распределенного хранения информации или, в случае использования в учебном процессе мультимедийных учебных курсов, *применение специальных систем хранения*, например, RAID-массивов, SAN и NAS.

Использование систем хранения RAID, SAN и NAS решает проблему ограниченности ресурсов жесткого диска сервера баз данных, но не избавляет от сложностей интеграции информационных ресурсов в единое хранилище портала. Кроме того, системы хранения данных имеют достаточно высокую стоимость.

Концепция распределенного хранения образовательного контента реализуется при построении образовательного портала на основе распределенного хранилища или на основе концепции GRID [2, 3, 7]. Распределенный подход позволяет решить и проблему обновления ОИР, и проблемы их интеграции. Но при таком подходе предъявляются требования к качеству, загруженности каналов передачи данных и доступности серверов, на которых размещены ОИР.

Для устранения недостатков централизованного и распределенного подходов необходимо исследовать *комбинированный подход к хранению образовательного контента*. Для реализации комбинированного подхода может быть предложена следующая концепция: в базе данных на web-сервере размещаются метаданные ОИР и данные, необходимые для работы служб портала, часть информационных ресурсов реплицируется на web-сервер или сервер баз данных с других серверов или рабочих станций. Образовательный контент хранится в распределенной базе данных на файл-серверах кафедр и подразделений образовательного учреждения, ответственных за их создание.

Управление образовательным контентом при комбинированном подходе к его хранению. Задача управления контентом образовательного портала при комбинированном подходе к хранению учебных материалов может быть

сформулирована следующим образом. Множество пользователей образовательного портала $A = \{A_s, s = \overline{1, s_0}\}$ формируют запросы в соответствии с интенсивностями $\lambda = (\lambda_1, \dots, \lambda_s, \dots, \lambda_{s_0})$. Время формирования запроса s -м пользователем является случайной величиной, распределенной по экспоненциальному закону с плотностью $f_s(t) = \lambda_s \cdot e^{-\lambda_s t}$, $s = \overline{1, s_0}$. Образовательные информационные ресурсы представлены множеством $R = \{r_i, i = \overline{1, n}\}$. ОИР $r_i \in R$ характеризуется подмножеством $X = \{K, V, TP\}$, где $K = \{k_d, d = \overline{1, K_0}\}$ - множество разделяемых единиц контента, связанных в рамках определенной модели, K_0 - количество разделяемых единиц контента, составляющих ОИР, $V = \{v_d, d = \overline{1, K_0}\}$, где v_d - объем в байтах каждой разделяемой единицы контента, $TP = \{tp_d, d = \overline{1, K_0}\}$, где tp_d - тип разделяемой единицы контента (текст, изображение, аудио фрагмент, видео фрагмент).

Полное множество запросов состоит из двух непересекающихся подмножеств Q^F - множество запросов на чтение, Q^W - множество запросов на обновление ОИР. Связь между множеством запросов $Q = Q^F \cup Q^W$ и множеством ОИР задается матрицами $B^F = \|b_{vd}^F\|$ и $B^W = \|b_{kd}^W\|$. Элементы b_{vd}^F матрицы B^F , идентифицирующие взаимосвязи между запросами на чтение и образовательными информационными ресурсами R , определяются следующим образом:

$$b_{vd}^F = \begin{cases} 1, \text{ если ОИР используется при выполнении запроса } q_v \in Q^F; \\ 0 - \text{ в противном случае} \end{cases}$$

Элементы b_{kd}^W матрицы B^W формируются так:

$$b_{kd}^W = \begin{cases} 1, \text{ если } q_p \in Q^W \text{ модифицирует ОИР } r_i; \\ 0 - \text{ в противном случае} \end{cases}$$

Все ОИР размещены на множестве узлов сети $U = \{u_m, m = \overline{1, M}\}$. Параметры, определяющие характеристики сети заданы тройкой вида: $P = \langle WR, DA, SP \rangle$. $WR = (WR_1, \dots, WR_m, \dots, WR_M)$, где WR_m - среднее время поиска и считывания данных в узле сети $u_m \in U$; $DA = (DA_1, \dots, DA_m, \dots, DA_M)$, где DA_m - среднее время доступа к данным узла сети; $SP = \|SP_{mh}\|$, $m, h = \overline{1, M}$, где SP_{mh} - среднее время передачи данных по каналам сети между узлами $u_m \in U$ и $u_h \in U$.

Распределение ОИР по узлам сети задается в виде бинарной матрицы $G = \|g_{mi}\|$, $m = \overline{1, M}$, $i = \overline{1, n}$, элементы которой принимают значения $g_{mi} = 1$, если ОИР r_i хранится на узле u_m , в противном случае $g_{mi} = 0$.

Предполагается, что выбор пользователем запроса носит вероятностный характер, значения вероятностей определяются элементами матрицы $F = \|f_{sy}\|$,

$s = \overline{1, S_0}$, $y = \overline{1, Q_0}$, f_{sy} - вероятность формирования пользователем A_s запроса $q_y \in Q^F \cup Q^W$.

Необходимо построить математическую модель, которая позволяла бы найти такое подмножество ОИР $R^{res} = \{r_j^{res}, j = \overline{1, n}\}$ множества R , которые необходимо реплицировать на web-сервер. При этом среднее время реакции системы на запросы пользователей T и количество невыполненных запросов пользователя Q^{lost} из-за загруженности web-сервера должно быть минимальным.

На параметры модели накладываются следующие ограничения:

1. ОИР могут быть размещены либо на одной из рабочих станций сети, либо одновременно и на узле сети и на web-сервере, т.е. $r_i \in R : \sum_{i=1}^n g_{mi} \leq 2$,
2. Если ОИР копируется на web-сервер, то происходит репликация всех разделяемых единиц контента k_d , из которых формируется этот ОИР.
3. Суммарный объем разделяемых единиц контента всех ОИР, реплицируемых на web-сервер, не должен превышать объема, выделенного для их хранения на жестком диске web-сервера:

$$\sum_{i=1}^n \sum_{d=1}^{K_0} v_{id} \leq VD,$$

где VD – количество байт на жестком диске web-сервера, выделенных для хранения ОИР.

1. При оценке времени реакции системы на запросы пользователей будут учитываться только запросы на чтение ОИР $q_{vd} \in Q^F$, поскольку редактировать ОИР, как правило, могут пользователи, имеющие соответствующие права: администраторы портала, авторы курсов, и основная нагрузка на портал состоит в выполнении запросов на предоставление и чтение образовательной информации.

Таким образом, задача нахождения множества R^{res} будет иметь вид: найти R^{res} такое, что $T \rightarrow \min$, $Q^{lost} \rightarrow \min$, при условии, что $r_i \in R : \sum_{i=1}^n g_{mi} \leq 2$,

$$\sum_{i=1}^n \sum_{d=1}^{K_0} v_{id} \leq VD.$$

Заключение. При рассмотрении подходов к хранению контента образовательных порталов можно сделать следующие выводы:

1. При хранении в единой базе данных образовательных информационных ресурсов обеспечивается высокая реактивность запросов пользователей портала, а также возможность централизованного резервного копирования и восстановления базы данных. Недостаток такой концепции – трудность управления обновлением образовательных информационных ресурсов.

2. Распределенный подход позволяет реализовать портал, в котором необходимо работать с большим объемом мультимедийной информации, а также портал распределенного университета, когда сложно следить за обновлением информации и сложно интегрировать все информационные ресурсы в единое хранилище данных. Однако при таком подходе предъявляются высокие требования к пропускной способности и загрузке каналов передачи данных.
3. Для построения портала ВУЗа с большим объемом мультимедийной информации целесообразным является использование комбинированного подхода к хранению контента, при котором часть образовательного контента реплицируется на web-сервер или сервер баз данных с других серверов или порталов. Реализация такого подхода к хранению образовательного контента позволяет решить проблемы интеграции ОИР, низкой реактивности и пропускной способности каналов передачи данных.
4. При комбинированном подходе к хранению образовательного контента необходимо определить множество информационных ресурсов R_{res} , к которым будет осуществляться централизованный доступ. В качестве критериев эффективности предполагается использовать время реакции системы на запросы пользователей T и число невыполненных запросов пользователей Q^{lost} .

СПИСОК ЛИТЕРАТУРЫ:

1. Анищенко Н.Г., Васильев П.М., Кореньков В.В., Крюков Ю.А. Хранение и доставка учебных видеоматериалов при дистанционной форме обучения \ Материалы конференции GRID'2006, 26 – 30 июня, <http://grid2006.jinr.ru/rus/programme30.asp>
2. Васильев П.М. Методы и технологии доступа к видеоинформации системы дистанционного обучения с использованием распределенной базы данных: Дис. ... канд. техн. наук. – Дубна, 2006. – 123 с.
3. Жижимов О.Л., Мазов Н.А., Фролов А.С. Доступ к базам данных ISIS из Internet и построение распределенной информационной системы \ Вычислительные технологии – 1997. – Т.2. – №3. – с. 45-50.
4. Ибрагимов И.М. Информационные технологии и средства дистанционного обучения: Учеб. пособие для студ. высш. учеб. заведений / Под ред. А.Н.Ковшова. – М.: Издательский центр «Академия», 2005. – 336 с.
5. Лясин А.С. Как создать портал в Internet: Основы использования web-технологий. – М.: Познавательная книга Пресс, 2003. – 288 с.
6. Позднеев Б.М., Буханов А.Н. Анализ систем управления информационным содержанием образовательных порталов \ http://magazine.stankin.ru/arch/n_20/3/index.html, 2002
7. Шокин Ю.И., Ламин В.А., Федотов А.М., Барахнин В.Б., Жижимов О.Л., Мазов Н.А., Пищик Б.Н., Покровский Н.Н. Распределенная информационная система «Виртуальный музей науки и техники СО РАН» // Труды 5-ой Всеросс. науч. конф. RCDL'2003 – 2003. – с. 112-116