Модели управления доступом к распределенным информационным ресурсам♣

© Жижимов О.Л., Федотов А.М.

Институт вычислительных технологий СО РАН
zhizhim@sbras.ru, fedotov@sbras.ru

Аннотация

На основе анализа типовых сценариев работы информационных серверов (WWW, FTP, Z39.50 и т.п.) сформулированы задачи, которые должны решаться при организации системы контроля доступа к распределенным информационным ресурсам. Рассмотрены возможности технологии LDAP как наиболее подходящей для построения подобной системы. В рамках этой технологии обсуждаются три модели управления доступом, отличающиеся степенью интеграции функций информационных серверов с технологией LDAP.

Создание и поддержка распределенных информационных систем и электронных библиотек, интегрирующих разнородные информационные ресурсы и функционирующих в различных программно-аппаратных средах, требует специальных подходов к управлению этими системами [1-2]. Если управление собственно самими ресурсами или данными может осуществляться в локальном режиме даже для распределенных информационных систем [2], то задача управления доступом к распределенным ресурсам не может быть решена в рамках локального администрирования. Обоснование последнего тезиса можно увидеть при рассмотрении типичных сценариев работы информационного сервера.

1 Типичные сценарии работы информационного сервера приложений

Несмотря на разнообразие приложений, предоставляющих доступ к информационным ресурсам, их функционирование происходит по однотипным сценариям. Типичный сценарий акта извлечения информации может быть представлен в следующем виде

1. Клиент посылает серверу запрос на просмотр информационного ресурса.
2. Сервер принимает запрос и выделяет идентификационные параметры клиента (адрес, имя, пароль, сертификат и т.п.).
3. Сервер проверяет подлинность клиента по предъявленным идентификационным параметрам (аутентификация).
4. Сервер проверяет доступность данному клиенту запрошенного им информационного ресурса (авторизация).
5. Сервер извлекает ресурс из хранилища и передает его клиенту.

Здесь продемонстрирована последовательность операций с положительным исходом каждой. В реальной ситуации эта последовательность может быть прервана с фиксацией состояния ошибки или включать дополнительные процедуры для уточнения тех или иных параметров запроса клиента.

Таким образом, при обработке запроса на извлечение информационного ресурса из некоторого хранилища сервер вынужден обратиться к трем разнотипным базам данных:

1. Тип 1 - база данных пользователей, включающая список пользователей и их идентификационные параметры.
2. Тип 2 - база данных прав пользователей на доступ к информационным ресурсам.
3. Тип 3 - собственно хранилище информационных ресурсов.

При этом зачастую тип 3 содержит два совершенно различных хранилища:

1. Хранилище описаний информационных ресурсов – метаданные, каталог ресурсов и т.п.
2. Хранилище собственно информационных ресурсов, например, полные тексты, бинарные данные и т.п.

Заметим, что доступ к этим двум подсистемам хранилища информационных ресурсов может осуществляться по различным технологиям, т.е. через различные серверы. Например, доступ к метаданным – по протоколу Z39.50, а доступ к соответствующим бинарным данным – по протоколу FTP или HTTP. При этом такое различие технологий доступа к ресурсам не должно приводить к различию баз данных типов 1 и 2 для различных сервисов обслуживания пользователей.

Ситуация еще более усугубляется при организации распределенных информационных систем, в которых целая совокупность разнородных информационных серверов должна использовать единую интегрированную для всех информацию типа 1 и 2. Так возникает естественное желание иметь единую точку доступа к данным типа 1 и 2 для всех сервисов распределенной информационной системы. Это желание не может быть удовлетворено в рамках локальных информационных систем.

Труды 9ой Всероссийской научной конференции «Электронные библиотеки: перспективные методы и технологии, электронные коллекции» - RCDL'2007, Переславль-Залесский, Россия, 2007.

2 LDAP – технологическая основа системы управления доступом к ресурсам

Как следует из вышесказанного, для решения задачи управления доступом к распределенным ресурсам необходимо внедрение технологий, изначально базирующихся на парадигмах «распределенности», с одной стороны, и «итеропрабельности», с другой стороны. Для успешного использования различными серверами построение технологии должно быть основано на международных стандартах открытых систем и хорошо поддерживаемы производителями программных продуктов.

В качестве подобной технологии для корпоративного применения может быть использована технология LDAP (Lightweight Directory Access Protocol — облегченный протокол доступа к каталогам) при наличии развитой системы корпоративных каталогов, объединенных в единую корпоративную распределенную справочную систему (КРСС) (см., например, [3-4]).

Однако для достижения цели, т.е. создания системы управления доступом к распределенным информационным ресурсам, необходимо решения в рамках технологий LDAP ряда дополнительных задач:

- Создание логической надстройки над КРСС, включающей определения дополнительных схем данных и основных процедур контроля доступа к ресурсам корпоративной распределенной информационной системе (КРИС).
- Создание информационной составляющей системы управления доступом к распределенным информационным ресурсам (СУДРИР) – расширения КРСС, допускающие хранение и обработку исходных данных по контролю доступа к ресурсам.
- Адаптация серверного программного обеспечения, предоставляющего доступ ресурсам (Z39.50, WWW, FTP и т.д.), к возможности работы в соответствии с правилами СУДРИР.
- Создание интерфейсов для управления СУДРИР.

Единая корпоративная распределенная справочная система (КРСС) как базовый элемент СУДРИР должна удовлетворять определенным требованиям:

- Организация информации в КРСС должна обеспечивать ее сегментацию на отдельные административные сегменты.
- Актуальность информации в КРСС должна поддерживаться множеством администраторов независимо в каждом сегменте системы.
- Доступ к КРСС должен быть основан на открытых стандартах.
- КРСС должна обеспечивать возможность хранения информации различного типа с возможностью ее поиска по различным атрибутам.

При использовании технологий LDAP для создания КРСС перечисленные требования могут быть удовлетворены [5]. Наличие КРСС является необходимым условием для успешного построения СУДРИР. Основные требования, которые можно предъявить к системе управления доступом к распределенным информационным ресурсам (СУДРИР) можно сформулировать следующим образом

- СУДРИР должна быть интегрирована с КРСС.
- Технология СУДРИР должна быть основана на международных стандартах и протоколах.
- СУДРИР должна допускать масштабирование и быть многоплатформенной.
- СУДРИР должна включать демократичные пользовательские интерфейсы.

СУДРИР должна обеспечивать решение следующих задач управления:

- Контроль доступа к каждому информационному ресурсу КРИС в соответствии с установленными политиками.
- Поддержка подсистемы формирования политик контроля доступа для различных клиентов КРИС. При этом клиентами КРИС могут быть
  - Физические лица – пользователи, идентифицируемые именем или цифровым ключом.
  - Компьютеры – рабочие места пользователей, идентифицируемые по IP-адресу или цифровому ключу.
  - Сетевые сервисы – программы, запрашивающие ресурсы из КРИС и идентифицируемые по сетевым протоколам, портам и цифровым ключам.
  - Группы, объединяющие вышеперечисленные субъекты в любой комбинации.

3 Модели функционирования системы управления доступом к ресурсам

Выбор технологии LDAP для построения СУДРИР оставляет открытыми вопросы реализации механизмов контроля управления доступом к распределенным информационным ресурсам. Эта реализация зависит от выбора модели СУДРИР.

Если выделить основные функциональные элементы СУДРИР:

1. функция идентификации клиента КРИС (аутентификация),
2. функция задания правил доступа к ресурсам для различных категорий клиентов,
3. функция определения прав конкретного клиента КРИС (авторизация),
4. функция обеспечения соответствия прав клиента КРИС уровню предлагаемого сервиса КРИС,
5. функция учета используемых ресурсов (биллинга);

то только элемент 1 (аутентификация клиента) может быть реализован в технологиях LDAP без каких-либо дополнительных построений над КРСС. Реализация других элементов зависит от выбранной модели контроля доступа к распределенным информационным ресурсам. В зависимости от степени

«распределенности» перечисленных выше элементов можно выделить следующие модели.

1. Простая модель, в которой КРСС используется только для аутентификации клиента встроенными средствами LDAP-серверов, другие элементы СУДРИР реализованы локально для каждого сервиса и ресурса КРИС. Положительные качества модели – простота реализации единого пространства имен и паролей для клиентов КРИС, недостаток – отсутствие возможности поддержки единых политик доступа к распределенным ресурсам. Однако даже в этой простой модели на основе LDAP решается задача ведения единого реестра пользователей КРИС, их паролей и цифровых ключей на основе поддержки КРСС. Пример реализации этой модели продемонстрирован на Рис.1 для двух серверов – WEB-сервера Apache и FTP-сервера ProFTPD.

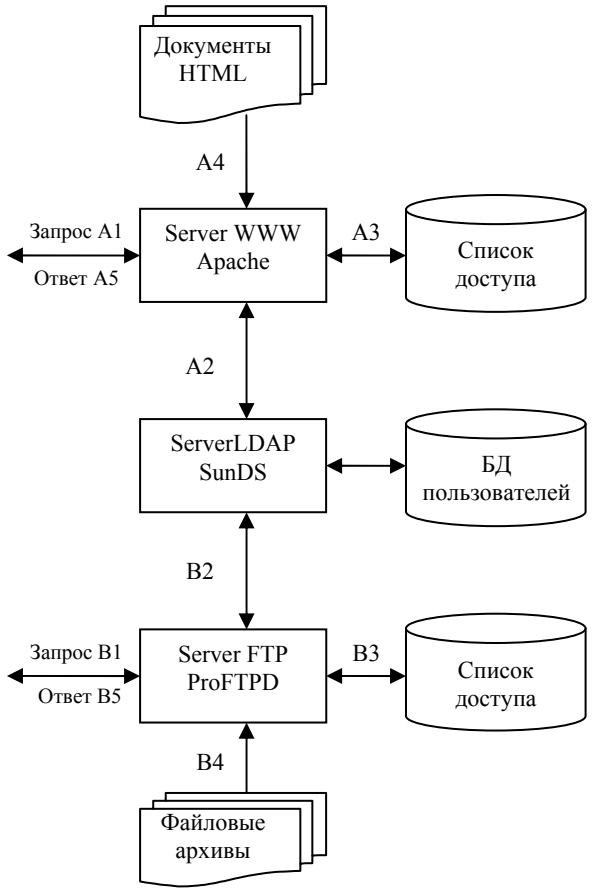

Рис.1 Иллюстрация модели 1:
A2, B2 – запросы на аутентификацию;
A3, B3 – запросы на авторизацию;
A4, B4 – извлечение данных.

2. Модель, в которой формулирование, проверка и реализация прав клиента происходит на основе технологий LDAP, т.е. в КРСС, допускает различные вариации, вплоть до выдачи сертификатов в модели X.509. Поскольку сегодня принято осуществлять контроль доступа к различным информационным объектам, основываясь на списках доступа (ACL – Access Control List), то эти вариации могут отличаться как способом хранения ACL, так и способом привязки информационных объектов к ACL.

a. В наиболее простом варианте ACL формулируются на основе встроенных механизмов LDAP-серверов как наборы штатных серверных инструкций (ACI - Access Control Instructions) по управлению доступом к элементам дерева КРСС. Положительное качество этого способа – простота реализации при условии наличия в каталоге КРСС элемента описания соответствующего информационного объекта. Проверка прав клиента на доступ к информационному объекту при этом сводится к проверке этих прав на доступ к соответствующему описанию объекта в каталоге КРСС. В ситуации, когда необходим различный уровень доступа к первичным и вторичным (описаниям) информационным объектам, этот способ не может быть использован.

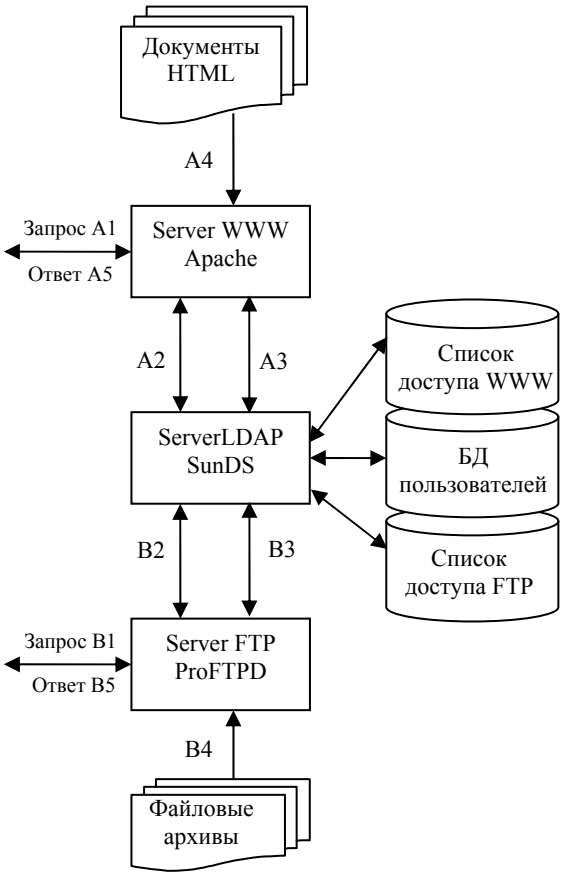

Рис.2 Иллюстрация модели 2:
A2, B2 – запросы на аутентификацию;
A3, B3 – запросы на авторизацию;
A4, B4 – извлечение данных.

b. Более сложным представляется вариант, когда ACL формулируются на основе специальной схемы данных – набора объектов и атрибутов каталога КРСС. При этом описания информационных объектов КРИС должны обязательно присутствовать в КРСС и содержать атрибуты, определяющие правила доступа к собственно объектам, а не к их описаниям, доступ к которым определяется ACI (см. выше). Каждый информационный сервер при этом должен проверять права доступа клиентов к ресур-

сам, обращаясь к серверу КРСС и анализируя соответствующие атрибуты описания запрошенного ресурса. Этот способ более сложный и затратный, чем предыдущий, но позволяет реализовать полный контроль над доступом к информационным ресурсам в соответствии с определенными выше требованиями.

Оба варианта модели 2 требуют, чтобы, с одной стороны, в каталоге КРСС (корпоративном LDAP-каталоге) существовали объекты определенного класса – описания информационных ресурсов, интегрированных в КРИС, а с другой – чтобы информационные серверы КРИС (WWW, FTP, Z39.50 и т.п.) при предоставлении доступа к ресурсу всегда обращались к соответствующим описаниям. На основе анализа кодов возврата (вариант a) или значения некоторых атрибутов (вариант b) информационный сервер должен принять решение о соответствия прав клиента КРИС уровню предлагаемого сервиса КРИС.

Пример реализации модели 2 продемонстрирован на Рис.2 для двух серверов – WEB-сервера Apache и FTP-сервера ProFTPD.

Эффективность применения той или иной модели контроля доступа к распределенным информационным ресурсам может быть определена только для определенной информационной системы с конкретной топологией и информационными ресурсами. В Сибирском отделении РАН в рамках целевой программы «Информационно-телекоммуникационные ресурсы СО РАН» [6-7] сегодня создается распределенная корпоративная информационная система, в основу которой закладываются механизмы, реализующие различные модели. Работы в этом направлении поддержаны РФФИ и интеграционными проектами СО РАН для создания нового поколения средств управления доступом к информационным ресурсам различного типа, анализа их свойств и технологий реализации в целях повышения эффективности использования информационных ресурсов российского академического сообщества.

Литература

[1] Жижимов О.Л., Федотов А.М., Чубаров Л.Б., Шокин Ю.И. Технология создания распределенных информационно-вычислительных ресурсов СО РАН // Труды I междунар. конф. САИТ-2005. Системный анализ и информационные технологии. 12-16 сентября 2005 г., г. Переславль-Залесский, т. 2., Переславль-Залесский, 2005, 161-165.

[2] Жижимов О.Л., Мазов Н.А. Принципы построения распределенных информационных систем на основе протокола Z39.50. - ОИГГМ СО РАН, Новосибирск: ИВТ СО РАН. - 2004. - ISBN 5-9554-0017-6. - 361 с.

[3] Созыкин А.В., Масич Г.Ф., Масич А.Г., Бездушный А. Н. Вопросы интеграции информационных и сетевых служб. Варианты использования LDAP каталогов // Труды 6-ой Всероссийской научной конференции «Электронные библиотеки: перспективные методы и технологии, электронные коллекции» - RCDL2004,Пущино, Россия, 2004.

[4] Бездушный А. А., Нестеренко А. К., Сысоев Т.М., Бездушный А. Н., Серебряков В. А. Возможности технологий ИСИР в поддержке Единого Научного Информационного Пространства РАН // Электронные библиотеки. – 7 (6). - 2004.

[5] Жижимов О.Л. Турпанов А.А. Федотов А.М. Корпоративный каталог СО РАН // Электронные библиотеки: перспективные методы и технологии, электронные коллекции: Труды Восьмой Всероссийской научной конференции (RCDL'2006) Суздаль, 17-19 окт., 2006 г. - Ярославль, 2006. - С. 226-230.

[6] Шокин Ю.И., Федотов А.М., Жижимов О.Л., Мазов Н.А. Интегрированная распределенная информационная система (ИРИС) Сибирского отделения РАН // Выездное заседание координационного научного совета СО РАН по целевой программе: Информационно-телекоммуникационные ресурсы СО РАН, г. Иркутск, 29-30 июля, 2002 г.: Материалы заседания. - Иркутск, СО РАН. Иркутский научный центр. - 2003. - С.139-149.

[7] Жижимов О.Л., Мазов Н.А., Федотов А.М. Центр доступа к электронным информационным ресурсам СО РАН // Библиотеки и информационные ресурсы в современном мире науки, культуры, образования и бизнеса: 13-я междунар. конф. "Крым 2006" (10-18 июня 2006 г., г. Судак): Труды конф., М., Изд-во ГПНТБ России, 2006.

Models of management of access to the distributed information resources

Zhizhimov O.L., Fedotov A.M.

Institute of Computational Technologies SB RAS
zhizhim@sbras.ru, fedotov@sbras.ru

On the basis of the analysis of typical scenarios of work of information servers (WWW, FTP, Z39.50, etc.) problems which should dare at the organization of the monitoring system of access to the distributed information resources are formulated. Possibilities of technology LDAP as similar system most suitable to construction are considered. Within the limits of this technology by the access, differing degree of integration of functions of information servers three models of management are discussed with technology LDAP.

*Работа выполнена при частичной поддержке РФФИ: проекты 06-07-89060, 06-07-89038, 07-07-00271, президентской программы «Ведущие научные школы РФ» (грант № НШ-9886.2006.9) и интеграционных проектов СО РАН.