

КОРПОРАТИВНАЯ ЭЛЕКТРОННАЯ ПОЧТА НА ОСНОВЕ СЕРВИСОРИЕНТИРОВАННОГО ПОДХОДА

Мараканов И.Н.

(Учреждение Российской академии наук
Институт проблем управления РАН, Москва)

Разработаны принципы построения и программное обеспечение корпоративной автоматизированной системы обмена информацией. Система предназначена для обеспечения надежного и оперативного обмена информацией в территориально распределенных государственных и коммерческих организациях, имеющих главный офис и подчиненные филиалы. Время доставки сообщений - от долей секунды до нескольких секунд. Исключаются дублирование и потери сообщений. Основными элементами системы являются центры обмена данными (ЦОД) и специализированные Web-службы, которые обеспечивают хранение и доставку информации. В статье рассмотрены основные функции ЦОД, дано описание их основных компонент и принципов построения программного и информационного обеспечения. Приведены результаты практического использования и основные характеристики системы.

Ключевые слова: информационные системы, корпоративная электронная почта, Web-службы, центры обмена данными, гарантированная доставка информации, многопользовательская обработка данных

Введение

Одной из важных функций современных территориально распределенных корпоративных информационных систем является обмен данными между пользователями и между приложениями. При этом от таких систем часто требуется обеспечить гарантированную доставку информации (данных, документов) в гетерогенной сетевой среде в сжатые сроки. В обычных средствах электронной почты (например, Outlook Express) такие возможности отсутствуют, и требуется разработка и использование специального программного обеспечения.

Одним из примеров систем, обеспечивающих интеграцию приложений, является использование Брокеров Сообщений и средств Messaging Middleware [1]. В системах интеграции корпоративных информационных систем одним из подходов для такой интеграции стал подход, основанный на использовании программного обеспечения гарантированной пересылки сообщений между приложениями. Примером такого ПО является Microsoft BizTalk Server

Базовая идея этого ПО следующая. Имеется несколько приложений, связанных некоторой коммуникационной средой, возможно не очень надежной. Одно приложение (например, система документооборота А) должно переслать информацию/документ другому приложению (системе документооборота В). Система А передает документ серверу пересылки сообщений и "забывает" о нем. Сервер пересылки сообщений обеспечивает гарантированную и однократную доставку информации до системы В. Брокер сообщений при этом выполняет не только маршрутизацию сообщений между информационными системами, но и выполняет также задачу трансформации этих сообщений для обеспечения совместимости информационных систем. Брокер сообщений интегрирует гетерогенные приложения и данные и предоставляет три типа служб: пересылка сообщений и перемещение данных обеспечивает физический транспорт доставки сообщений между приложениями; интеллектуальная маршрутизация, которая определяет для каждого сообщения то, к какому приложению оно должно попасть; трансформирование обеспечивает определение соответствия данных между различными семантиками одного приложения или разных приложений.

Системы корпоративной электронной почты позволяют передавать офисную электронную корреспонденцию между внутрикорпоративными абонентами, а также разнообразную технологическую информацию между автоматизированными рабочими местами и процессами. Такие системы имеют следующие основные возможности [2]:

- обеспечение приоритетов отправки сообщений,
- доставка сообщений группе пользователей,
- уведомления о доставке,
- ведение журнала входящей и исходящей корреспонденции,
- архивирование передаваемых сообщений,
- поиск сообщений в архивах,
- защита от дублирования сообщений,
- автоматическое сжатие почтовых сообщений,
- антивирусная проверка сообщений,
- ведение глобальной адресной книги,
- автопроцессинг почты с помощью скриптов или программных модулей,
- высокая надежность.

Цель создания рассматриваемой в настоящей статье системы заключается в обеспечении надежного и оперативного процесса обмена данными между корпоративными пользователями и приложениями с использованием специализированных Web-служб. Web-службы (web-services) представляют собой удаленные функции, расположенные на Web-сервере и позволяющие обращаться к ним удаленным клиентским приложениям, используя HTTP и SOAP-протоколы.

Web-службы могут реализовывать вызов удаленных методов, а могут просто принимать произвольные XML-сообщения для некоторой обработки. В первом случае они могут взаимодействовать с некоторым бизнес-объектом системы (определяемым полученным запросом), вызывая в свою очередь его методы для выполнения необходимой операции. В то же время они могут просто выполнять заданные функции в зависимости от параметров пришедшего XML-документа, что более характерно для модели обмена сообщениями. При этом клиент Web-службы, отправив запрос, может остановить свою работу и ждать ответа (синхронная модель), а может продолжать работу и принять ответ через некоторое время, когда тот придет от Web-службы (асинхронная модель). Web-службы позволяют не только передавать информацию между клиентом и сервером, но и выполнять обработку переданной информации, а также записывать ее в соответствующие базы данных, связанные с Web-сервером.

Таким образом, применение сервисориентированного подхода, основанного на использовании Web-служб; позволяет не только обеспечить передачу информации в режиме 'on-line', но и оперативно выполнять технологические операции, такие как автоматическая регистрация исходящей электронной почты у получателя и др.

В качестве основных элементов системы предлагается использовать Центры обмена данными (ЦОД) [3,4]. Создание ЦОД позволяет обеспечить надежный и контролируемый обмен данными на основе автопроцессинга (циклического процесса обмена файлами из специальных дисковых директорий), обеспечить автоматический процесс регистрации электронных документов, передаваемых в другие организации, исключить дублирование сообщений и др. Схема функционирования ЦОД показана на рис. 1.



Рис. 1. Схема функционирования ЦОД

Среди основных функций ЦОД можно выделить следующие: обеспечение обмена информацией между пользователями разных филиалов; ведение базы данных передаваемых сообщений; контроль входящих и исходящих сообщений; поиск переданных сообщений в базе данных; синхронизация справочников. Основными компонентами ЦОД являются специализированные Web-службы, обеспечивающие надежное хранение и оперативную доставку информации.

Программные средства, входящие в состав АРМ пользователей системы, должны обеспечивать такие функции как обмен данными между пользователями локальной сети филиала, обмен данными с ЦОД своего филиала (взаимодействие с Web-службами ЦОД), обеспечение совместной отправки исходящих и приема входящих сообщений несколькими пользователями, подготовка электронных документов на основе корпоративных шаблонов, ведение справочников и др.

1 Структура автоматизированной системы обмена информацией

При обмене информацией могут использоваться три основные схемы взаимодействия ЦОД: 'каждый с каждым', 'кольцо' и 'звезда'. Обмен информацией по схеме 'каждый с каждым' приводит к большому числу связей, а именно $N(N - 1)$, и усложняет систему. Использование схемы типа 'кольцо' вносит дополнительные задержки при доставке информации. Наиболее подходящей схемой взаимодействия для рассматриваемой системы является звездообразная схема. Данная схема требует использования дополнительного центрального сервера системы, но, с другой стороны, она обеспечивает эффективный процесс обмена информацией. Следует также отметить, что многие территориально распределенные корпоративные системы для решения своих задач используют аналогичную топологию аппаратных средств, поэтому на практике дополнительные затраты на техническое обеспечение рассматриваемой автоматизированной системы могут оказаться минимальными.

На рис. 2 показана структура разработанной системы.

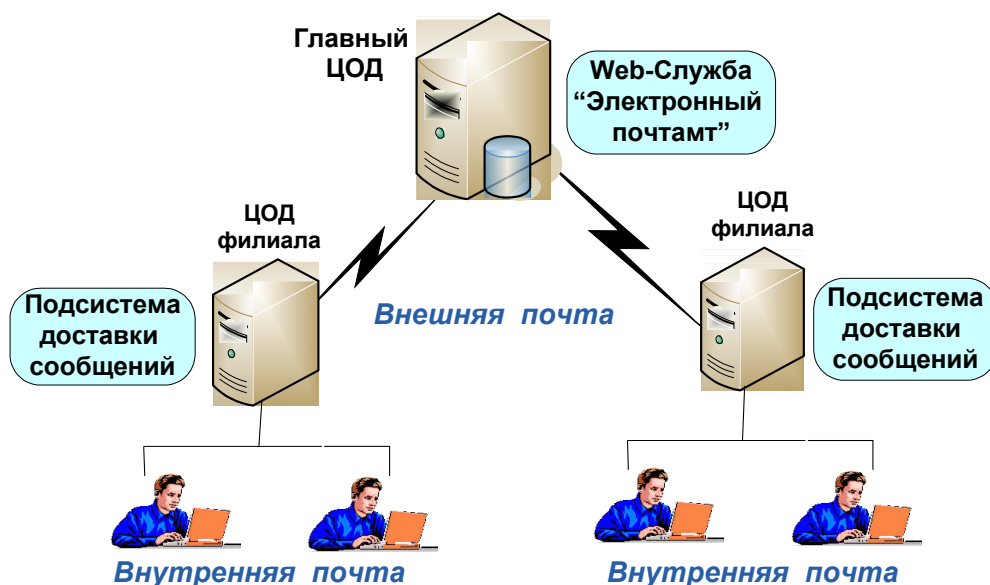


Рис. 2. Структура автоматизированной системы обмена корпоративной информацией

Сформированные пользователями сообщения (документы, файлы) передаются средствами “Внутренней почты” и сохраняются в специальной дисковой директории отправки ЦОД филиала. Подсистема доставки сообщений в заданном регламенте осуществляет циклический просмотр этой директории и передачу подготовленных сообщений в главный ЦОД системы, а также прием от главного ЦОД сообщений, предназначенных для пользователей данной филиала.

Принятые сообщения сохраняются в специальной директории приема ЦОД филиала. Просмотр и передача полученных сообщений соответствующим пользователям филиала может осуществляться как специальным сотрудником в интерактивном режиме средствами “Внутренней почты”, так и в автоматическом режиме с помощью специальной подсистемы диспетчеризации входящих сообщений [5].

В главном ЦОД имеется Web-служба, которая выполняет функции ‘Электронного почтамта’, т.е. принимает от ЦОД филиалов сообщения и размещает полученные файлы в соответствующих дисковых директориях, а также осуществляет отбор файлов, предназначенных для обратившегося филиала, и передачу их в ЦОД этого филиала.

Использование такой сервисориентированной архитектуры позволяет обеспечить оперативный обмен данными между указанной Web-службой главного ЦОД и обратившимся к ней клиентским приложением, а также реализовать доставку информации в режиме ‘on-line’.

2 Принципы организации информационного и программного обеспечения

Как уже отмечалось выше, обмен данными с главным ЦОД в рассматриваемой системе осуществляет подсистема доставки сообщений ЦОД филиала. В случае отправки какого-либо сообщения подсистема доставки сообщений просматривает дисковую директорию автопроцессинга (в которой хранятся отправляемые сообщения) и, если находит в ней готовое к отправке сообщение (в виде XML-файла с информацией о получателях и имеющихся приложениях), формирует соответствующий данному сообщению ZIP-файл, выполняет соединение с Web-службой главного ЦОД и передает ей указанный ZIP-файл. Web-служба главного ЦОД принимает ZIP-файл от ЦОД филиала и размещает его в специальной дисковой директории сервера.

При приеме сообщений от главного ЦОД подсистема доставки сообщений ЦОД филиала сначала обращается к Web-службе главного ЦОД с запросом об имеющихся в адрес обратившегося филиала сообщениях. В случае, если имеются такие сообщения в главном ЦОД, Web-служба передает обратившемуся клиентскому приложению перечень их номеров (каждый номер однозначно определяет сообщение). После этого подсистема доставки сообщений ЦОД филиала по очереди выполняет запросы на получение сообщений по каждому полученному номеру.

Для реализации рассмотренной схемы разработано специальное информационное и программное обеспечение главного ЦОД системы и ЦОД филиала. Рассмотрим его основные компоненты. К их числу относятся:

- специальная дисковая директория главного ЦОД, в которой хранятся поступающие от ЦОД филиалов сообщения,
- таблица WS_MAIL в составе базы данных главного ЦОД, обеспечивающая хранение и поиск доставляемых сообщений,
- Web-служба, обеспечивающая взаимодействие главного ЦОД с клиентскими приложениями ЦОД филиалов и доставку сообщений.

Таблица WS_MAIL содержит следующие основные поля:

- receiver_id - код получателя,
- sender_id - код отправителя
- doc_num – номер отправляемого документа,
- zip_name – полное имя доставляемого ZIP-файла,
- files_cnt – кол-во доставляемых файлов (в составе ZIP-файла),
- zip_size – размер ZIP-файла,
- doc_desc – краткое содержание отправляемого сообщения.

Таким образом, каждому получателю и номеру документа в таблице WS_MAIL соответствует одна запись. При этом в данной таблице хранятся записи только о недоставленных сообщениях. То есть после доставки сообщения получателю соответствующая запись удаляется из таблицы.

Таблица WS_MAIL позволит не только быстро искать и формировать необходимые для отправки данные, но просто и быстро получать статистику недоставленных сообщений за заданный оператором интервал времени. Эта статистика может формироваться в разрезе номеров документов или получателей. Если, например, подсистемы доставки сообщений ЦОД филиалов должны обращаться к главному ЦОД не реже, чем один раз в несколько часов, и для какого-то филиала имеются недоставленные сообщения более, чем за сутки, то это свидетельствует о ненормальной работе модуля приема сообщений в данном филиале (либо модуль работает со слишком большим интервалом и этот интервал необходимо уменьшить, либо модуль неисправен).

3 Оптимизация размера передаваемых фрагментов данных

Разработанное программное обеспечение системы позволяет передавать файлы заданными фрагментами и, в случае сбоев, осуществлять повторный обмен информацией, начиная с того фрагмента файла, на котором произошел сбой.

Выбор размера передаваемых фрагментов влияет на общее время доставки данных. Ниже рассмотрена задача выбора оптимального размера фрагмента файла, обеспечивающего минимизацию заданного критерия, например, среднего времени доставки информации. В общем случае аналитическая постановка данной задачи достаточно сложна, поэтому рассмотрим ее упрощенный вариант, который, тем не менее, позволяет получить необходимые для практического использования оценки [6].

Если предположить, что сбой во время доставки фрагмента файла может происходить не более одного раза, средний размер переданной до сбоя части фрагмента составляет половину фрагмента и скорость передачи данных постоянная, то среднее время доставки файла размера R можно записать в следующем виде:

$$T(r) = \frac{R}{r} \left\{ P(r) \left(\frac{r}{s} + \frac{r}{2s} + t + \tau \right) + (1 - P(r)) \left(\frac{r}{s} + \tau \right) \right\}.$$

Здесь введены следующие обозначения:

$T(r)$ – среднее время передачи данных (сек.),

r – размер фрагмента передаваемых данных (байт),

R – общий размер передаваемых данных (байт),

$P(r)$ – вероятность сбоя при передаче данных размера r ,

- s – скорость передачи данных (байт /сек.),
 t – время переналадки после сбоя (сек.),
 τ – время переналадки после нормального завершения передачи данных (сек.).

Задача заключается в отыскании такого значения r^* , для которого среднее время $T(r)$ минимально, т.е. справедливо равенство

$$T(r^*) = \min T(r), r \leq R,$$

где r^* – оптимальный размер фрагмента файла.

На практике зависимость $P(r)$ может быть задана в некоторых заданных точках и аппроксимирована кусочно-линейной функцией. Соответственно в этих же точках могут быть вычислены значения функции $T(r)$ и получена оценка значения r^* .

В разработанной системе реализован адаптивный алгоритм формирования функции $P(r)$ и выбора размера оптимального фрагмента передаваемых файлов. На начальной стадии функционирования системы используются значения $P(r)$, полученные экспертным путем. По мере функционирования системы зависимость $P(r)$ корректируется в зависимости от полученных результатов эксплуатации системы. Зависимость оптимального размера фрагмента r_{opt} от размера файла R показана на рис.3. Через R_{max} на рисунке обозначен максимальный размер передаваемого XML-файла.

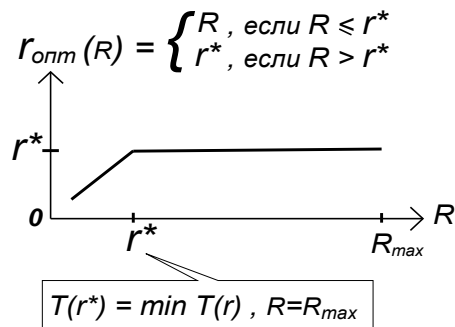


Рис.3. Вид зависимости оптимального размера фрагмента для файла размера R

Таким образом, вычислив по имеющимся данным значение r^* , формируется функция Γ_{opt} и для заданного значения R определяется оптимальный размер фрагмента файла.

4 Управление многопользовательской обработкой данных

В процессе обмена информацией могут возникать ситуации, требующие управления многопользовательской обработкой данных.

Например, несколько пользователей могут попытаться активизировать модуль доставки информации ЦОД филиала для отправки подготовленных сообщений получателям. Возникает необходимость управления доступом к общему ресурсу, которым в данном случае является модуль доставки сообщений. Общая схема доставки информации показана на рис. 4.

Задача управления многопользовательской отправкой данных заключается в следующем. В процессе работы пользователей сформированные ими файлы с передаваемыми данными сохраняются в общей дисковой директории ЦОД филиала. Каждый пользователь, если он обладает необходимыми правами, может вызвать на исполнение модуль доставки сообщений, который является клиентским приложением подсистемы доставки сообщений ЦОД филиала и обеспечивает обмен данными с главным ЦОД. Система должна управлять процессом обмена информацией, разрешая активизировать указанный модуль доставки только одному из пользователей, запрещая в это время работу с ним другим пользователям.



Рис. 4. Схема доставки информации

Таким образом, возникает задача управления блокировками общего ресурса, которым в данном случае является модуль доставки сообщений. В [7] предложен подход к решению такого рода задач. Блок-схема разработанной двухуровневой системы управления многопользовательской доставкой сообщений на примере выполнения запроса приложения на доступ к общему ресурсу показана на рис. 5.

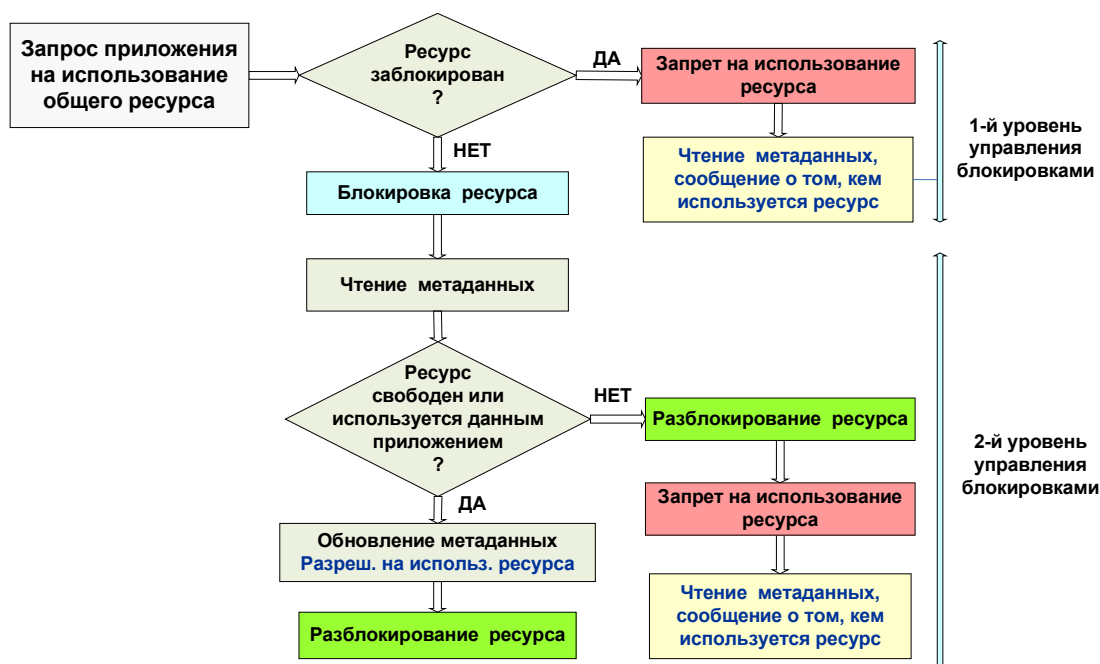


Рис.5. Блок-схема выполнения запроса на использование общего ресурса

На первом уровне обеспечивается доступ к общему ресурсу в каждый момент времени только одного из обратившихся пользователей. На втором уровне осуществляется чтение и анализ метаданных, и формирование разрешения или отказ на использование модуля доставки данных. Как можно заметить из приведенной блок-схемы, общий ресурс находится в заблокированном состоянии в течение небольшого отрезка времени, в течение которого выполняются операции обработки метаданных. Основные функции управления блокировками ресурса выполняются с помощью метаданных и специальной и программы-диспетчера, к которой

обращается модуль доставки сообщений. Программа-диспетчер обеспечивает чтение, анализ и формирование специальных метаданных, проверку состояния (занят/свободен) общего ресурса, получение информации о том пользователе, который в данный момент времени выполняет обмен данными, если ресурс занят, блокировку и разблокирование общего ресурса.

5 Результаты практического использования системы

На рис. 6 показана главная форма разработанной системы доставки сообщений. В нижней части формы располагается список готовых к отправке сообщений, а верхней части формы – список доставленных сообщений. В двух правых столбцах фиксируется входящий номер получателя сообщений, а также дата и время доставки. По мере доставки сообщений они переходят из нижнего списка в верхний. В случае сбоев или ошибок, возникающих в процессе доставки сообщений, они фиксируются в специальном файле и могут быть просмотрены. Оператор имеет возможность контролировать как сам процесс доставки, так и его результаты.

№	Получатель	Номер док-та	Дата	Файлов	Размер (кб)	Содержание сообщения	Входящий ГВЦ	Дата доставки
1	ГВЦ	154	27.05.2009	6	342	описание-1		03.09.2009 15:...
2	ГВЦ	160	02.09.2009	9	15745	КРАТКОЕ СОДЕРЖАНИЕ 123456789	239 Дубл...	03.09.2009 14:...
3	ГВЦ	154	27.05.2009	6	342	описание-1	237 Дубл...	03.09.2009 14:...
4	ГВЦ	160	02.09.2009	9	15745	КРАТКОЕ СОДЕРЖАНИЕ 123456789	239	03.09.2009 14:...
5	ГВЦ	900159	02.06.2009	2	1	КРАТКОЕ СОДЕРЖАНИЕ 123456789	238	03.09.2009 13:...

№	Получатель	Номер док-та	Дата	Файлов	Размер (кб)	Содержание сообщения	Входящий ГВЦ	Дата доставки
1	ГВЦ	154	27.05.2009	6	342	описание-1	237 Дубли...	03.09.2009 15:...
2	ГВЦ	160	02.09.2009	9	15745	КРАТКОЕ СОДЕРЖАНИЕ 123456789		
3	ГВЦ	9	02.06.2009	5	109	КРАТКОЕ СОДЕРЖАНИЕ 123456789		

Рис. 6. Главная форма системы доставки сообщений

Одной из важных задач, которая эффективно решается с помощью рассматриваемой системы, является регистрация исходящих из филиалов сообщений в главном офисе системы. Пользователи в филиалах в режиме реального времени получают информацию о факте регистрации в главном офисе и входящем регистрационном номере сообщения (электронного документа), направленного в адрес главного офиса. При этом система позволяет вести оперативный контроль и учет зарегистрированных сообщений и исключить ошибочные действия, связанные, например, с повторной регистрацией или потерей передаваемых сообщений. Аналогично решается и обратная задача регистрации исходящих из главного офиса сообщений, направленных в адрес филиалов.

Практическое использование данной системы подтвердило ее надежность и оперативность работы. Так, например, передача файлов, имеющих размеры в несколько килобайт, выполнялась в пределах одной секунды, а файлы в несколько мегабайт передавались в течение нескольких секунд.

Заключение

В статье рассмотрены принципы построения корпоративной электронной почты на основе использования сервисориентированной архитектуры и специализированных Web-служб. Система предназначена для обеспечения надежного и оперативного обмена информацией в территориально распределенных государственных и коммерческих организациях, имеющих главный офис и подчиненные филиалы. Время доставки сообщений - от долей секунды до нескольких секунд. Исключаются дублирование и потери сообщений.

Основные функции системы: шифрование информации, доставка сообщений группе пользователей, уведомление о доставке, ведение журнала входящих и исходящих сообщений, архивирование передаваемых сообщений, поиск информации в архивах, антивирусная защита, автопроцессинг отправляемых сообщений (циклический процесс обмена файлами из специальных дисковых директорий).

Предложенные подходы и методы решения рассмотренных задач использованы при разработке программного обеспечения системы. Рассмотрены результаты практического использования системы и ее основные характеристики. Использование системы позволяет обеспечить обмен и обработку данных в реальном масштабе времени.

Внедрение системы позволяет повысить оперативность и надежность работы корпоративных организаций. К их числу можно отнести организации МВД, МЧС, МО, а также коммерческие организации, осуществляющие оперативный обмен информацией.

Литература

1. <http://www.microsoft.com/Rus/Government/analytics/integration/overview.msp>
2. <http://www.freesoft.ru/?id=675314>
3. КОЗЛОВ А.Д., ЛЕБЕДЕВ В.Н., ЛИХОНИНСКИЙ В.С., МАРАКАНОВ И.Н., СТЕПАНОВ В.В. *Распределенная автоматизированная система документационного обеспечения управления на основе центров обмена документами*. // Проблемы экономики современных промышленных комплексов: Сборник статей 2-ой Всероссийской научно-практической конференции (2006г., Самара .Самарский государственный аэрокосмический университет им. Академика С.П. Королева, Институт проблем управления Российской академии наук им. В.А.Трапезникова).
4. МАРАКАНОВ И.Н. *Автоматизированная система подготовки и обмена электронными документами на основе XML-технологии*. // Проблемы управления и моделирования в сложных системах: Труды XI международной конференции (22-24 июня 2009г., Самара, Россия). – Самара: Самарский научный центр РАН, 2009. - С.582-587.
5. КОЗЛОВ А.Д., МАРАКАНОВ И.Н. *Решение задач управления электронным документооборотом на основе структурной идентификации XML-;документов*. Управление развитием крупномасштабных систем: Труды I международной конференции (Москва, 1-3 октября 2007г.). – Москва: ИПУ РАН, 2007. - С. 250-258..
6. КОЗЛОВ А.Д., МАРАКАНОВ И.Н. *Технология XML-документооборота*. // Управление развитием крупномасштабных систем: Труды III международной конференции (Москва, 3-5 октября 2009г.). – Москва: ИПУ РАН, 2009. - С.344-354.
7. ЛЕБЕДЕВ В.Н., МАРАКАНОВ И.Н., ОРЛОВ В.Л. *Метод двухуровневой блокировки для управления информационными ресурсами в разьединенном окружении*. Управление развитием крупномасштабных систем: Тезисы докладов I международной конференции (Москва, 1-3 октября 2007г.). – Москва: ИПУ РАН, 2007. - С.172-173.

CORPORATE E-MAIL ON THE BASIS OF SERVICE ORIENTED APPROACH

Marakanov Igor, Institute of Control Sciences of RAS, Moscow, Cand.Sc, (Moscow, Profsoyuznaya st., 65, +7(495)334-75-90). (inm@ipu.rssi.ru).

Abstract: Principles of construction and software of corporate automated system of information exchange are developed. The main elements of the system are data exchange center and specialized Web-services, which provide storage and delivery of information. The system is designed to ensure reliable and rapid exchange of information in a geographically distributed organizations with a main office and subordinate branches. Delivery time of messages - from a fraction of a second to several seconds, duplication and loss of messages are excluded. The basic functions of the data centers and description of their principles of design software and information are considered. The results of practical use and basic characteristics of the system are presented.

Keywords: corporate information system, e-mail, service oriented architecture, Web-services, data exchange center, express message delivery, multiuser data processing