ПРОБЛЕМЫ ЗАЩИТЫ ИНФОРМАЦИИ В ГЛОБАЛЬНО-РАСПРЕДЕЛЕННЫХ ИНФОРМАЦИОННО-ВЫЧИСЛИТЕЛЬНЫХ СИСТЕМАХ

Золотницкий И.Ю., Шеин А.В.
ФГУП "ЦНИИАТОМИНФОРМ" (127434, Москва, Дмитровское шоссе, 2)
socium@comtv.ru, shein@ainf.ru

PROBLEMS OF INFORMATION SECURITY IN THE GLOBAL-DISTRIBUTED INFORMATION-COMPUTING SYSTEMS

Zolotnitski I.Yu, Shein A.V.
FSUE “TSNIIATOMINFORM” (127434, Moscow, Dmitrovskoe shosse, 2)
socium@comtv.ru, shein@ainf.ru

It is considered separate aspects of a problematics of information security in existing and new approaches to construction of the distributed information-computing systems in the global computer environment.

В контексте глобальной компьютерной среды проблематика электронных библиотек, как сложных интегрированных информационных систем, направлена на разработку и интеграцию компьютерных, методологических, правовых и других решений, связанных с формированием единого, высокоорганизованного и защищенного информационного пространства. На уровне компьютерного представления информации это пространство должно в перспективе объединить в единое целое существующие библиотечные и архивные фонды, упорядочить разнообразные аспекты издания и распространения печатных материалов, дать качественные сервисы по поиску, регламентированному доступу как для широкой публики, так и в условиях строгих ограничений по доступу при работе с данными [1, 2]. В перечисленных вопросах проблемы защиты информации больших и сверхбольших хранилищ информации занимают высокие приоритеты.

Подавляющая часть информационных систем в глобальной компьютерной среде построена в сетевой архитектуре "клиент-сервер" – простейшей из возможных. Браузеры и веб-серверы посредством гипертекстовых технологий, работающих на базе сетевых протоколов и управляемые при помощи графических интерфейсов, создали глобальный информационный феномен под названием Интернет.

Портальные технологии, став вершиной глобальных гипертекстовых достижений, не дали полномасштабного решения проблем электронных библиотек. Построены и успешно используются большое количество библиотечных порталов, но, как все понимают, это лишь первые практиче-

ские шаги. Разрозненность и отсутствие единого стиля построения и пользования этих порталов – весьма серьезная проблема для потребителей.

С точки зрения проблематики электронных библиотек, слабым звеном архитектуры "клиент-сервер" становится явная перегруженность специалистов, обеспечивающих сервер информацией. В случае электронных библиотек, объемы данных и концентрация сервисов в серверной части становятся главным ограничителями в развитии. Понятно, что вряд ли в обозримом будущем можно ожидать появления суперсервера, который соберет в одном месте все информационные ресурсы и обеспечит регулярный сервис по доступу к ним.

Решение лежит, скорее, в направлении *распределенного* накопления и упорядочения информации в глобально-распределенной вычислительной среде с вовлечением сколь угодно широкого круга заинтересованных участников. При этом компьютеры должны быть организованы в единую, целостную информационно-вычислительную среду, с возможностями быстрых и сложных поисков, автоматической реструктуризации, способную к самоорганизации во взаимодействии со специалистами и потребителями.

Для движения в этом направлении нужны функционально более богатые сетевые архитектуры. Это распределенные сетевые архитектуры, называемые Peer-to-Peer (P2P). Различные их реализации составляют платформы для универсального метакомпьютинга. В перспективе метакомпьютинг призван обеспечить универсально программируемые вычисления в глобальной компьютерной среде, подобно тому, как современные технологии программирования обеспечивают программирование внутренних ресурсов отдельного компьютера.

С середины 90-х технологии универсального метакомпьютинга развиваются под названием GRID–технологий. В основном это университетские проекты. Строятся они, как правило, на основе сетевых реализаций технологий объектно-ориентированного программирования. На их основе разворачиваются системы для решения крупномасштабных задач в области физики элементарных частиц, геофизики, биологии [3]. Одной из наиболее продвинутых GRID-платформ является система GLOBUS

Рассмотрим подробнее концепцию защиты, применяемой в GRID технологиях на данный момент [4].

Изначально организация PACI (Partnerships for Advanced Computational Infrastructure), создававшая распределенную информационно-вычислительную систему на основе GRID платформ, встретилась с проблемой объединения институтов-участников в единое целое. Институты-участники имели собственные развитые компьютерные центры и в каждом из них существовали хорошо продуманные правила и процедуры для различных аспектов работы таких центров, в том числе для защиты информации и, в частности, аутентификации. На одних узлах использовались

различные версии Kerberos и DCE (распределенной вычислительной среды), на других Secure shell (SSH).

Инфраструктура Kerberos, используемая автономно или в рамках распределенной вычислительной среды, выполняет аутентификацию пользователей посредством защищенных транзакций с поддерживаемым централизованным образом сервером ключей. Kerberos обеспечивает межкорпоративную аутентификацию за счет создания в организациях доверительных серверов ключей. Kerberos соответствует многим из базовых требований для аутентификации в виртуальной организации, но порождает две следующие проблемы.

- Использование Kerberos для межузловых аутентификаций также означает его обязательное использование для аутентификации в пределах узла, что для части организаций невозможно из-за высокой стоимости оборудования и необходимости привлечения дорогостоящих специалистов.
- Узлы должны заключать множество межузловых соглашений об аутентификации, и многих это не устраивает из-за слишком большого внешнего контроля над тем, что происходит локально.

Распределенная вычислительная среда (DCE), обеспечивает сервисы, построенные преимущественно на тех же принципах, что и Kerberos. DCE обеспечивает дополнительные сервисы авторизации, а также специальную файловую систему и прочие сервисы. Но, даже обеспечивая значительно более эффективные сервисы авторизации, чем у простого Kerberos, DCE сохраняет его существенные недостатки. Например, нет использования криптографии с открытым распределением ключей для P2P аутентификации.

Также есть существенные недостатки и у Secure shell. Технология SSH требует, чтобы пользователи сами управляли отношениями межузловой аутентификации за счет копирования открытых ключей (или отслеживания паролей) для всех узлов, к которым они обращаются. Данная задача может оказаться крайне обременительной, если пользователи обращаются ко многим узлам. Более того, SSH не позволяет узлам управлять аутентификацией, поэтому нельзя, к примеру, отказать в доступе конкретному пользователю без нарушения его конфиденциальности. Также SSH поддерживает только ограниченные возможности - rshell и передачу файлов, но не другие средства, требующие аутентификации, такие как среды совместной работы и Web-браузеры.

В самом начале формирования PACI предпринимались активные попытки убедить большинство участников использовать Kerberos, чтобы превратить эту систему в предпочтительное, всеобщее средство аутентификации. Однако оказалось, что добиться этого невозможно в силу раз-

личных технических, финансовых и организационных причин (в том числе указанных выше).

В результате стало ясно, что члены сообщества в ближайшем будущем по-прежнему будут применять и Kerberos, и иные решения. Механизмы поддержки совместного использования ресурсов, применяемые в PACI, должны сосуществовать с разнообразными локальными механизмами.

GSI (GRID Security Infrastructure) являет собой альтернативный подход к организации защиты GRID систем. Его разработка была начата в рамках проекта GLOBUS. GSI имеет дело с внутренними операциями, предоставляя локальные решения защиты для узлов входящих в GRID и обеспечивает реализацию политики безопасности GRID[].

Ключевыми пунктами, определяющими политику безопасности GRID, являются [5]:

1. Вычислительная среда GRID состоит из многих *доменов доверия*.

По сути, эти структуры осуществляют принцип необходимости объединения разнородных наборов локально управляемых пользователей и ресурсов. Среда GRID никакого влияния на локальную политику безопасности оказывать не должна. Следовательно, защита среды GRID заключается в управлении междоменными взаимодействиями и своевременном отображении междоменных операций на локальную политику безопасности.

2. Ограничения, введенные в отдельном домене доверия, - субъекты лишь локальной политики безопасности

То есть не происходит никаких дополнительных действий и не создается дополнительных защитных сервисов со стороны среды GRID над сервисами, существующими в локальной политике безопасности домена доверия. Реализация локальной политики безопасности происходит, как правило, реализацией множества методов, включая брандмауэры, Kerberos, SSH.

3. Существует два типа субъектов: глобальные и локальные. Существует частичное отображение из глобальных субъектов в локальные для каждого домена доверия.

В действительности, каждый пользователь ресурса будет иметь два имени: глобальное имя и, возможно отличающееся от него, локальное имя для каждого ресурса. Отображение глобального имени в локальное – каждый раз индивидуально. Например, сайт может отобразить глобальное имя пользователя в предопределенное локальное имя, динамически резервируемое локальное имя или даже имя определенной группы. Существование же глобального субъекта позволяет обеспечивать однократную авторизацию при работе в среде GRID.

4. Операции между объектами расположенными в различных доменах доверия требуют взаимной аутентификации

5. Заверенный глобальный субъект, отображенный в локальный субъект, принимается как локальный субъект домена.

Другими словами, внутри домена доверия комбинация политики аутентификации GRID и локального отображения удовлетворяет требованиям безопасности хоста домена.

6. Всё разграничение доступа осуществляется локально, на уровне локальных субъектов.

То есть разграничение доступа осуществляется локальными системными администраторами

7. Программа или процесс, который действует от имени пользователя, обладает подмножеством прав этого пользователя.

Данный элемент политики безопасности необходим для обеспечения выполнения долгодействующих программ, которые могут запрашивать ресурсы динамически без вмешательства и взаимодействия с пользователем. Также это необходимо для создания процессами других процессов.

8. Процессы, запущенные от имени того же субъекта внутри того же домена доверия, могут делить одиночные наборы мандатов доверия (так называемые *credentials*)

GRID вычисления могут включать в себя сотни процессов, выполняющиеся на одном ресурсе. Данный компонент политики безопасности допускает масштабируемость архитектуры безопасности для её использования в крупномасштабных параллельных приложениях и позволяет избежать необходимости создания уникального мандата для каждого процесса.

GRID Security Infrastructure реализует вышеописанную политику безопасности и обладает следующими важными особенностями [4].

- Полномочия, использующие сертификаты стандарта X.509v3 в качестве частных ключей, представляют “личность”, или средства идентификации, каждого объекта - пользователя, ресурса или программы, указывая имя объекта и дополнительную информацию, скажем, открытый ключ. Уполномоченный по выдаче сертификатов (certification authority - CA), некая пользующаяся доверием независимая организация, подписывая сертификат, связывает средства идентификации объекта с открытым ключом.
- Алгоритм аутентификации, определенный протоколом Secure Socket Layer Version 3 (SSLv3), выполняет идентификацию объекта. Достоверность результатов такой проверки определяется степенью доверия к CA, поэтому локальный администратор принимает эти сертификаты, используя их затем для проверки цепочек сертификатов.
- Объект может делегировать подмножество своих прав (например, процесс, порождающий другой процесс) третьей стороне, создавая временные средства идентификации, называемые посредником

(proxy). Сертификаты посредников формируют цепочку, которая начинается с CA, а затем наращивается, когда сначала пользователь, а затем посредники пользователя подписывают сертификаты. Путем проверки цепочки сертификатов процессы, инициированные одним и тем же пользователем на различных узлах, могут аутентифицировать друг друга, проводя проверку обратно по цепочке сертификатов до тех пор, пока не будет найден исходный сертификат пользователя.

- Каждый ресурс может задавать свои правила для определения того, как надо реагировать на входящие запросы. Первоначально GSI использовала просто список контроля доступа, но в текущей версии реализованы более развитые методы.
- Протокол аутентификации выполняет проверку подлинности глобальных имен участвующих сторон, но GSI должна преобразовать это имя в локальное (например, регистрационное имя или имя доверителя Kerberos), прежде, чем локальная система защиты сможет его использовать. GSI выполняет эту процедуру на локальном узле, сверяясь с простым текстовым файлом соответствий, который устанавливает связи между глобальными и локальными именами.
- Стандартный интерфейс GSS-API обеспечивает доступ к функциям защиты. GSI использует OpenSSL или SSLeay (свободно распространяемую реализацию SSLv3) для своих протоколов аутентификации и поддержки сертификатов посредников. SSLv3 широко применяется для обеспечения безопасности в Web.

Несмотря на относительную простоту, данная архитектура соответствует всем критически важным требованиям пользователей и систем.

- С точки зрения пользователей, глобальное имя и полномочия посредников означают, что пользователю для получения доступа ко всем ресурсам необходимо только один раз пройти процедуру аутентификации, а полномочия посредников и процедура делегирования полномочий позволяют программам, работающим от имени пользователя, обращаться к ресурсам. Использование стандартов X.509, SSLv3 и GSS-API стимулирует разработку общего инструментария, ориентированного на GSI и более сложных приложений.
- С точки зрения узлов, архитектура не требует пересмотра локальной инфраструктуры защиты; вместо этого узлы просто устанавливают относительно простые серверы, поддерживающие GSI, которые используют хорошо известные стандарты. Узлы управляют правилами работы с помощью списка контроля доступа и файла соответствий, поэтому администраторам удобно работать с CSI и они готовы развертывать ее параллельно с SSH и другими механизмами удаленного доступа.

Несмотря на то, что система GLOBUS является наиболее развитым примером построения GRID-платформ, она все еще не является полноценной сетевой операционной средой [6]. Интерес к сетевым архитектурам P2P на уровне лидеров-производителей компьютерного рынка в настоящее время только складывается. Общепризнанного коммерчески значимого инструментария пока нет. Отсюда - важный вывод. Проблема программирования глобально-распределенных ресурсов находится в начальной фазе своего практического решения.

В [7,8,9] предложен новый подход к интегральным решениям глобально-распределенных задач в парадигме единого, математически однородного поля компьютерной информации на основе исчисления древовидных структур. На этой основе разрабатываются новые принципы и технологии построения глобально-распределенных информационно-вычислительных систем в едином, математически однородном поле компьютерной информации [7, 8].

В рамках этого подхода интеграция компьютерных решений начинается на уровне математически однородных форм машинного представления структурированной информации. В едином поле компьютерной информации (программ и данных, представляемых древовидными структурами), композиция и интеграция функций и систем осуществляется с существенно меньшими затратами времени и интеллектуальных усилий.

Технологии электронных библиотек предназначаются для работы со сверхбольшими объемами слабоструктурированной информации – текстами (в большей части неразмеченными) графической информацией, а также аудио и видеоинформацией, где структура (метаинформация) не имеет явных выражений. Дополнительные и весьма значительные проблемы возникают в связи с чрезмерным разнообразием компьютерных форматов представления данных.

Главное преимущество единого поля компьютерной информации – простые, математически однородные и, при этом, универсальные формы представления данных и программ. В рассматриваемом случае такие формы представляются в виде компьютерного исчисления древовидных структур [9]. Предварительные исследования и эксперименты показали [8], что традиционно сложные проблемы быстрого поиска, обновления больших объемов данных, обеспечения структурной и функциональной целостности компьютерной информации, переносимости, интеграции, масшабируемости компьютерных решений в потенциале могут производиться с существенно меньшими затратами времени и средств.

В едином поле компьютерной информации исчезают многочисленные информационные барьеры, требующие преобразования из одних форм представления в другие, препятствующие распространению информации и её интеграции. В новых условиях требования к защите информации приобретают уже не только прикладное, но и фундаментальное, системообра-

зующее значение. Защищенность информации должна поддерживаться уже на самых нижних уровнях машинного представления компьютерной информации. Формы машинного представления программ и данные должны учитывать этот аспект. Особое значение имеют подходы к организации сверхнадежных информационно-вычислительных и управляющих процессов в больших и сверхбольших человеко-машинных системах [10]. Без принципиального решения вопросов защищенности данных, программ и процессов в едином поле компьютерной информации разрабатываемые в данной парадигме технологии не смогут на практике проявить свои качественно новые системообразующие возможности.

Программируемый метакомпьютинг [7] – критическая задача, предъявляющая к ныне действующей *парадигме изолированного компьютера* дисквалифицирующий набор требований. Им практически невозможно удовлетворить, сохранив в неприкосновенности классические постулаты. На смену классической парадигмы изолированного компьютера идёт парадигма единого, математически однородного поля компьютерной информации, которая *на уровне аксиоматики имеет дело с изначально распределёнными в WWW вычислениями*. В классической модели распределённые вычисления – результат сложнейших, многослойных композиций большого числа непростых компьютерных решений. В новой – атрибут математически однородного поля компьютерной информации, определяемого на уровне новой аксиоматики [7, 8].

Сейчас в модели исчисления древовидных структур строится экспериментальная версия программируемого метакомпьютинга [7] в математически однородном поле компьютерной информации. В рамках поддержки проекта рассматриваются возможности по обеспечению защиты информации при практическом использовании модели распределенных исчислений древовидных структур на основе языка и системы программирования ПАРСЕК [9]

С этой целью наряду с изучением организации защиты в существующих реализациях GRID технологий осуществляется разработка модели, позволяющей изначально, на этапе формирования основополагающих принципов реализации глобально-распределенных информационных систем в математически однородном поле компьютерной информации, обеспечивать комплексную защиту информации.

Литература

1. В. Петров, А. Пискарев, А Шеин. Информационная безопасность. Защита информации от несанкционированного доступа в автоматизированных системах, МИФИ, 1993 , 73 с.

2. Грушо А.А. Тимонина Е.Е. Теоретические основы защиты информации. Издательство Агентства “Яхтсмен”.1996 г
3. Кореньков В., Тихоненко Е. Организация вычислений в научных областях. Открытые системы, № 2, 2001. (http://www.osp.ru/os/2001/02/030.htm)
4. Р. Балтер, В. Уэлч, Д. Энгерт, Я. Фостер, С. Тюке. Инфраструктура аутентификации в национальном масштабе. Открытые системы, № 2, 2001.(http://www.osp.ru/os/2001/02/040.htm)
5. I. Foster, C. Kesselmann, G. Tsudik, S. Tuecke. A Security Architecture for Computational Grids. Proc. ACM Conf. Computers and Security. ACM Press, N.Y., 1998, p.83-91.
6. Коваленко В., Коваленко Е., Корягин Д., Любимский Э., Хухлаев Е. Управление заданиями в распределенной вычислительной среде. Открытые системы, № 2, 2001.(http://www.osp.ru/os/2001/05-06/022.htm)
7. Затуливетер Ю.С. Программируемый метакомпьютинг в математически однородном поле компьютерной информации. (Доклад, опубликованный в данном сборнике).
8. Затуливетер Ю.С. Информация и эволюционное моделирование. Труды Международной конференции “Идентификация систем и задачи управления”, “SICPRO’2000”, Москва, 26-28 сентября 2000г, Институт проблем управления РАН, с.1529-1573 (http://zvt.hotbox.ru/, http://zvt.by.ru/).
9. Затуливетер Ю.С., Халатян Т.Г. ПАРСЕК - язык компьютерного исчисления древовидных структур с открытой интерпретацией. Стендовый вариант системы программирования. -М., 1997 (Препринт/Институт проблем управления РАН), 71с.
10. Затуливетер Ю.С., Лубков Н.В., Карибский В.В. Проблема организации надежных процессов управления с применением ненадежных вычислительных сред. Тезисы докладов. Четвертая международная конференция "Проблемы управления в чрезвычайных ситуациях", г. Москва, 11 января, 1997, с.160-163. (http://zvt.hotbox.ru/, http://zvt.by.ru/)