Безопасность чувствительной информации в облаках

Обзор подходов EMC к решению задачи предотвращения утечек информации в облачных средах методами шифрования и токенизации.

Александр Чигвинцев – руководитель подразделения EMC RSA в регионе Россия и СНГ.

Введение

В облачной инфраструктуре владельцы информации утрачивают контроль над своими данными. Фактически они должны принять гарантии и условия внутреннего (private cloud) или внешнего (public cloud) сервис-провайдера по обеспечению безопасности информации.

По этой причине в настоящее время на рынке появляются технологии, позволяющие владельцам информации полностью контролировать доступ к данным – даже при условии их размещения в виртуальных инфраструктурах (облаках). В частности, подразделение информационной безопасности RSA корпорации EMC и сама EMC предлагают целый набор продуктов и технологий для решения подобной задачи. Следует заметить, что на сегодняшний момент широкое распространение и применение нашли методы защиты информации при ее передаче по различным каналам связи. Но методы защиты данных при их оперативном или долговременном хранении, а также при обработке приложениями пока не нашли такого же повсеместного применения в корпоративной среде. Однако статистика свидетельствует об обратном: "В 2008 ... 99% всех утечек произошли на уровне серверов и приложений" (Verizon Business 2009 Data Breach Study – April 15, 2009). По сведениям того же источника, "в 2008 г. было утеряно электронных записей больше, чем за 4 предыдущих года".

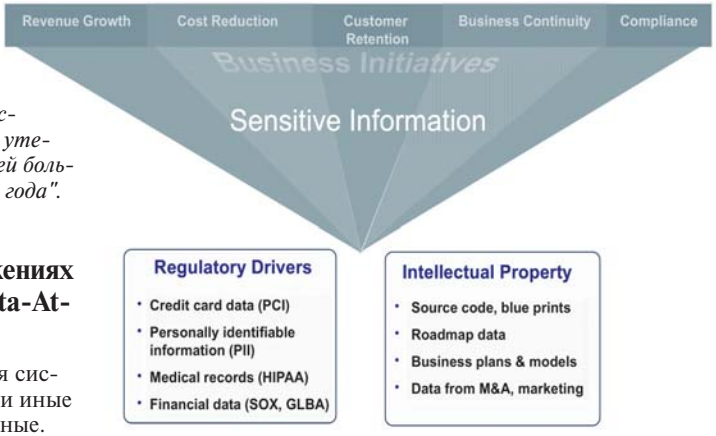

Рис. 1. Источники генерации чувствительной информации.

Защита данных при обработке в приложениях и при хранении (Data-At-Rest)

Любая информационная система обрабатывает те или иные конфиденциальные данные. "Конфиденциальность" данных диктуется как требованиями внешних или внутренних регуляторов, так и требованиями защиты интеллектуальной собственности предприятия (рис. 1). Например, требования защиты персональной информации, как и конфиденциальных данных, закреплены на законодательном уровне во всех развитых странах в почти всех сферах публичной жизни граждан.

Так, в соответствии с актом PCI DSS 3.4, номер кредитной карты (Credit Card Primary Account Number – PAN) должен быть защищен везде, где бы он ни хранился, включая: переносные носители, резервные копии, логи, файлы после передачи по сетям. Для этого предлагается использование следующих технологий: хэш-функций, усечения, индексации токенов, глубокой криптографии[1]. Также должны защищаться: медицинские записи (в соответствии с Актом HIPAA), персональная идентификационная информация (Personally identifiable information – PII, PII/Mass 201), финансовые данные (SOX, GLBA) и т.д.

В связи с распространением виртуальных и облачных технологий, в последние 1,5 года во многих регулирующих актах введены специальные пункты по защите чувствительных данных и в виртуальных средах.

В настоящее время технологии по защите чувствительных данных в наибольшей

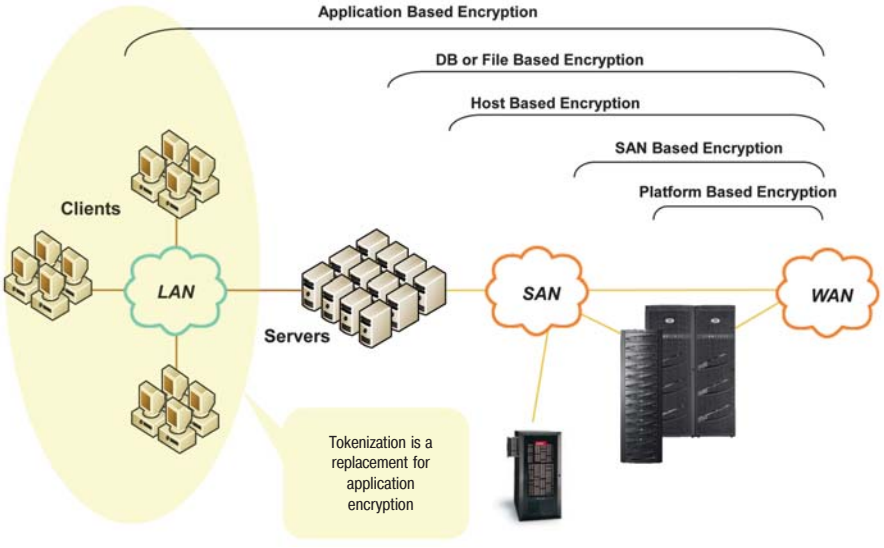

Рис. 2. Шифрование данных и токенизация.

1) "Render PAN, at minimum, unreadable anywhere it is stored (including data on portable digital media, backup media, in logs, and data received from or stored by wireless networks) by using any of the following approaches:
- Strong one-way hash functions (hashed indexes);
- Truncation;
- Index tokens and pads (pads must be securely stored);
- Strong cryptography with associated key management processes and procedures."

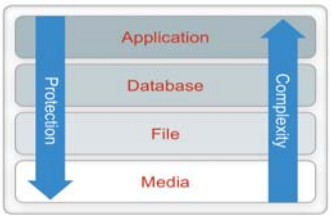

Рис. 3. С возрастанием уровня шифрования возрастает сложность реализации решения.

степени востребованы отраслями, находящимися под контролем регуляторов (как государственных, так и отраслевых): розничная торговля, предоставление медицинских и финансовых услуг.

Токенизация и шифрование

Технологии "кодирования" чувствительных данных – последний рубеж обороны вслед за методами идентификации, авторизации и DLP-технологий, если эти барьеры пройдены. Компания EMC предлагает два варианта подобной защиты информации: непосредственное шифрование на различных уровнях инфраструктуры и токенизация. В современном виде решения в составе портфеля EMC стали доступны с ноября 2010 г., когда был анонсирован продукт RSA Data Protection Manager. Сочетание средств защиты данных и методов управления ключами гарантирует не только усиленную защиту информации, но и за счет консолидации уровней управления – уменьшение текущих расходов на обеспечение безопасности.

RSA Data Protection Manager (прежнее название RSA Key Manager) по существу решает две задачи:

- ***корпоративное управление ключами***: управление ключами шифрования предприятия путем их интеграции в различные технологии шифрования данных при их хранении (дисковые массивы, магнитная лента и т.д.);
- ***токенизация:*** замена конфиденциальной информации подменяющим значением, или значением метки, для защиты таких данных, как номера кредитных карт, номера банковских счетов, номера социального страхования и другая информация, идентифицирующая личность. Истинное значение данных, соответствующих конкретному токену, хранится в БД на отдельном устройстве, управляющем токенами.

Традиционно шифрование используется на различных уровнях ИТ-инфраструктуры (рис. 2). Считается, что, чем ближе

Рис. 4. Данные после токенизации и шифрования.

Табл. 1. Сравнение особенностей технологий шифрования и токенизации.

	Токенизация	**Шифрование**
Развертывание	Требуется изменение только приложений, использующих данные с токенами. Каких-либо изменений в БД/файлы вносить не надо	Требуется изменение всех приложений, в которых используются данные с токенами, плюс все приложения, где изменение длины данных из-за шифрования влияет на другие данные/поля
Выполнение регулирующих требований	Только системы, которые имеют доступ к чувствительным данным и самому серверу токенизации в пределах границ согласия	Все системы в пределах границы согласия
Производительность *)	Централизованная модель, требующая понятной завершенности. Сетевые задержки требуется оценивать для влияния на производительность	Распределенная модель с высокой производительностью
"Off-line" использование	Требуется соединение с сервером токенизации или с распределенными серверами токенизации	Локально кэшируемый ключ позволяет автономное использование
Операционное влияние	Можно кастомизировать токен для снижения или ограничения потенциального операционного влияния	Формат зашифрованных элементов не может быть определен
Мобильность данных	Данные должны быть "детокенизированы", чтобы можно было их экспортировать вне кастомизированной управляемой зоны (Customer controlled domain)	Ключ может экспортироваться, что позволяет экспортировать зашифрованные данные

*) соотносится на типовых малого размера данных

модуль шифрования находится к приложению, тем надежнее работает решение. С другой стороны, чем выше по "ИТ-стеку" встраивается защита данных на основе шифрования, тем сложнее его реализовать и поддерживать (рис. 3). Кроме того, сертификация решений по шифрованию, требует много времени и значительных усилий по согласованию.

Также одним из минусов шифрования является то, что при криптографировании данных изменяется их формат (рис. 3), например, не только содержимое, но и длина. Это, в свою очередь, требует модификации всех приложений, на работу которых влияет такое изменение.

Для решения приведенных проблем была предложена альтернативная технология – токенизация (рис. 3). С одной стороны, процесс кодирования исходных данных происходит на уровне приложений (т.е. обеспечивается наилучшая защита исходной информации). С другой стороны, токенизированные данные сохраняют оригинальный формат, что ограничивает влияние этой технологии на приложение при сохранении высокого уровня защиты. Кроме того, метки (токены) способны сохранять некоторую часть исходных данных (например, последние четыре цифры номера социального страхования), так что другие приложения потенциально могут использовать метки, даже не получая доступа к реальной информации.

Среди основных преимуществ токенизации (табл. 1):

- расширение границ аудита (Tokens are OUTSIDE audit boundaries);
- отсутствие необходимости каких-либо изменений в интерфейсе пользователя / дизайне взаимодействия;
- отсутствие необходимости каких-либо изменений в приложениях типа "Pass-Through";
- отсутствие необходимости каких-либо изменений в схемах баз данных;
- сохранение производительности поиска в базе данных;
- способность поддержки множественных типов данных;
- минимизация требований для криптографической экспертизы;
- поддержка масштабируемой архитектуры в целях обеспечения высокой доступности и производительности.

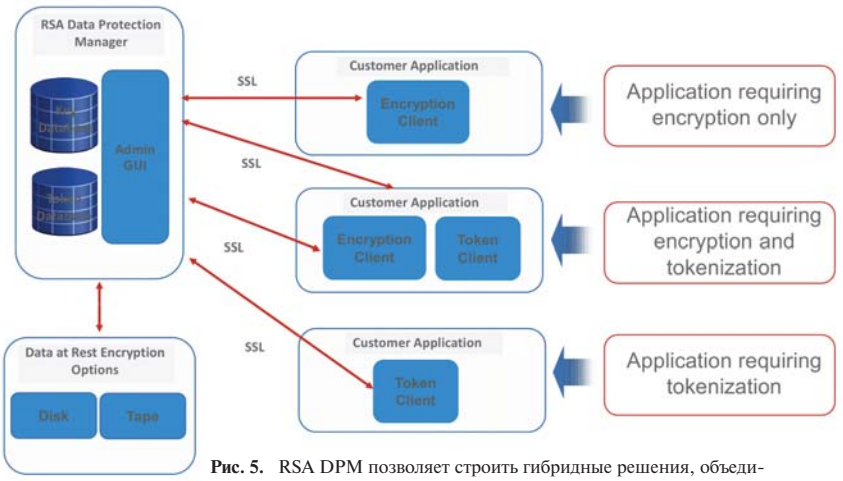

Рис. 5. RSA DPM позволяет строить гибридные решения, объединяя токенизацию и шифрование.

Сравнительные особенности технологий токенизации и шифрования приведены в табл. 1.

С учетом этих особенностей шифрование целесообразно в тех случаях, когда:

– сетевая связность и производительность являются узкими местами;

– устройство имеет минимальные возможности (например, POS-терминал и т.п.).

Токенизация может быть использована, где:

– сетевая связность является надежной компонентой;

– производительность не является критичной;

– системы имеют достаточный потенциал/мощность.

Для некоторых приложений оптимальная схема включает как шифрование, так и токенизацию. Такие гибридные решения позволяют оптимально использовать преимущества обоих методов – шифрования и токенизации для защиты чувствительных данных (рис. 4). При этом сам сервер токенизации и управления ключами шифрования может располагаться как на стороне клиента (например, в рамках распределенной облачной инфраструктуры), так и непосредственно в составе центра обработки данных сервис-провайдера.

Помимо методов токенизации и шифрования, EMC еще предлагает третий метод защиты данных – Format Preserving Encryption (FPE). Это технология, которая, шифруя данные, на выходе представляет их по типу токенизации. FPE имеет все те же самые ограничения, что и обычное шифрование при выполнении регулирующих требований: все системы должны быть в пределах границ. FPE разработан как “полевое” решение. Оно не может масштабироваться для использования шифрования файлов на дисках или на лентах.

Эти три метода позволяют существенно снизить как внутренние, так внешние риски утечки чувствительной информации, как в частных, так и публичных облаках, а также как при хранении информации, так и при ее обработке.

RSA DPM – управление токенизацией и ключами шифрования во всей инфраструктуре

RSA DPM является законченным решением для управления токенизацией и ключами шифрования всех компонент ИТ-инфраструктуры. RSA DPM выполняет следующие основные функции (рис. 6):

Рис. 6. RSA DPM представляет собой законченное решение со всей необходимой функциональностью для управления процедурами токенизации и шифрования.

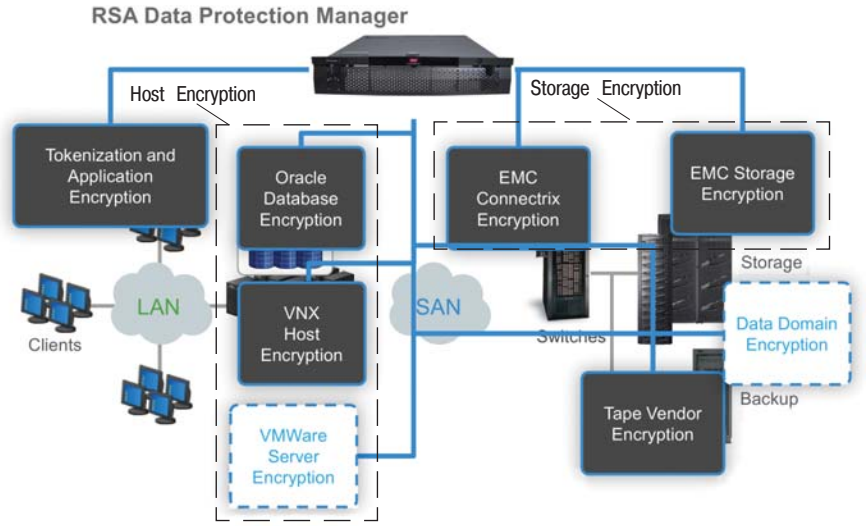

Рис. 7. RSA DPM поддерживает управление ключами шифрования для всех компонент ИТ-нфраструктуры: на уровне хоста, СХД/SAN, компонент резервного копирования/восстановления.

– защиту чувствительных данных от точек использования до хранения (в качестве клиентской части используются библиотеки для встраивания в приложения RSA BSAFE);

– централизованное назначение политик шифрования и токенезации клиентам с возможностью удаленного управления;

– полное максимально автоматизированное управление ключами и жизненным циклом токенов;

– масштабируемость управления ключами шифрования на каждом уровне в инфраструктуре.

В качестве системы поддержки шифрования данных RSA DPM фактически представляют собой специализированную систему по хранению и управлению жизненным циклом ключей для симметричных алгоритмов шифрования для всех компонент инфраструктуры: на уровне хоста, СХД/SAN, компонент резервного копирования/восстановления (рис. 7).

RSA DPM может поставляется в двух форм-факторах:

– *RSA DPM with Application Encryption* – ПО, которое устанавливается на операционные системы Windows, RedHat, Solaris и использует для хранения (KeyVault) ключевой информации коммерческие СУБД. В этом виде RSA DPM предоставляет разработчикам надежную, масштабируемую платформу для встраивания функций шифрования в приложение (на базе библиотек RSA BSAFE) и поддержки сервиса управления ключами;

– *RSA DPM for the Datacenter* – представляет собой программно-аппаратный комплекс (appliance), состоящий из KeyManager Server, установленного на аппаратной платформе (сервер). RSA DPM for the Datacenter поддерживает управление ключами шифрования для следующих устройств и программного обеспечения: EMC PowerPath with RSA Encryption, SAN коммутаторы Cisco и Brocade, Oracle TDE, Native Tape Encryption.

Как отмечалось, продукт может быть использован и в качестве системы поддержки токенизации – технологии кодирования данных без изменения их формата.

Примеры использования технологии шифрования в облачной инфраструктуре

Некоторые технологии шифрования данных, разработанные для корпоративных приложений, могут успешно использоваться и в облачных инфраструктурах. Это обусловлено тем, что даже при размещении приложений в “облаке” у пользователей остается возможность полностью контролировать и управлять процессом шифрования информации.

Одним из примеров такого рода решений может служить технология EMC PowerPath with RSA Encryption (рис. 8). В классическом виде EMC PowerPath – это так называемые multi-path драйвера под различные операционные системы для отказоустойчивого подключения сер-

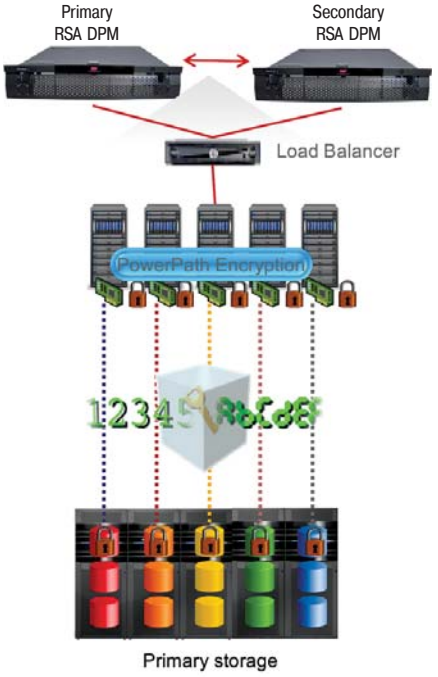

Рис. 8. Опциональный модуль шифрования в составе EMC PowerPath обеспечивает прозрачное, блочное шифрование данных при записи их на внешнюю систему хранения.

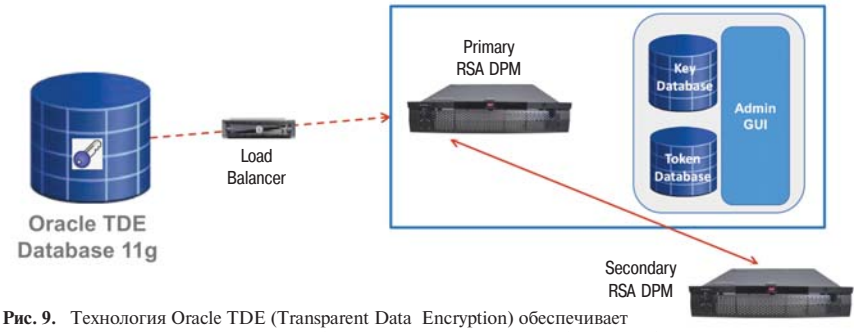

Рис. 9. Технология Oracle TDE (Transparent Data Encryption) обеспечивает прозрачное и гранулярное шифрование данных в таблицах СУБД Oracle.

веров к системам хранения. Поддерживается как использование каналов FC (fiber channel), так и IP (iSCSI).

Некоторое время назад в EMC PowerPath был добавлен опциональный модуль шифрования, который обеспечивает прозрачное, блочное шифрование данных при записи их на внешнюю систему хранения. Для каждого тома создается свой уникальный ключ шифрования, который автоматически передается и сохраняется в апплайнсе DPM for the Datacenter (для надежности, как правило, используются две системы DPM).

В случае облачных инфраструктур для надежной защиты данных достаточно разместить системы DPM на стороне пользователя. Теперь все данные будут храниться в облаке в зашифрованном виде, что исключает несанкционированный доступ к ним со стороны сервис-провайдера.

Стоит также отметить, что EMC PowerPath with RSA Encryption поддерживается и в виртуальной инфраструктуре – драйверы могут быть установлены в виртуальные машины и обеспечивать надежную защиту пользовательских данных.

Как было отмечено в начале статьи, подсистемы шифрования данных также встраиваются и в корпоративные приложения. В некоторых случаях эти решения в том или ином виде могут быть использованы и в облачных инфраструктурах.

Технология Oracle TDE (Transparent Data Encryption) обеспечивает прозрачное и гранулярное шифрование данных в таблицах СУБД Oracle (рис. 9). Для хранения ключевого материала, как и в предыдущем примере, могут использоваться системы RSA DPM. Как и раньше, размещение их за пределами облачной инфраструктуры сервис-провайдера обеспечивает достаточно надежную защиту пользовательских данных. В состав такой конфигурации входят:

- 1 Oracle TDE Database;
- 2 RSA Data Protection Managers;
- соответствующее число программных лицензий для БД хостов;
- Load Balancer (рекомендуется).

Примеры использования технологии токенизации в облачной инфраструктуре

Принцип, лежащий в основе технологии токенизации, позволит безболезненно переносить процессинг данных, попадающих под регулирование PCI DSS, в облачные инфраструктуры. Кроме того, даже в рамках традиционного ЦОД, токенизация позволяет резко сократить количество информационных систем процессингового центра, которые необходимо сертифицировать по требованиям PCI DSS, и, соответственно, сократить расходы компаний на приведение своей инфраструктуры в соответствие с требованием регулятора (рис. 10).

Решение RSA по токенизации обеспечивает всю необходимую функциональность, требуемую для работы приложений с защищенными данными:

- сбор части данных в простом текстовом формате;
- создание криптографически сильно защищенного токена в 1:1, совпадающего с форматом первичных данных;
- надежное хранение первоначальных данных в виде зашифрованного текста;
- поддержание механизма восстановления данных для любого токена;
- обеспечение сильной аутотентификации, авторизации и управления аудитом для всех чувствительных операций;
- поддержку шифрования данных локально в случае потери соединения в сети. После восстановления связи для операций токенизации становятся доступны зашифрованные данные.

Заключение

Традиционные технологии шифрования данных и новые разработки, базирующиеся на концепции сохранения формата защищаемых данных (токенизация), позволяют гарантировать сохранность конфиденциальной информации любого типа при переносе обработки данных в облачные инфраструктуры. Предоставление пользователям облачных сервисов – гарантии того, что они полностью управляют защитой своей информации и даже сервис-провайдер не в состоянии получить несанкционированный доступ к конфиденциальным данным, позволяет повысить лояльность компаний к облачным технологиям и, соответственно, ускорить их повсеместное внедрение.

Александр Чигвинцев,
EMC Россия и СНГ

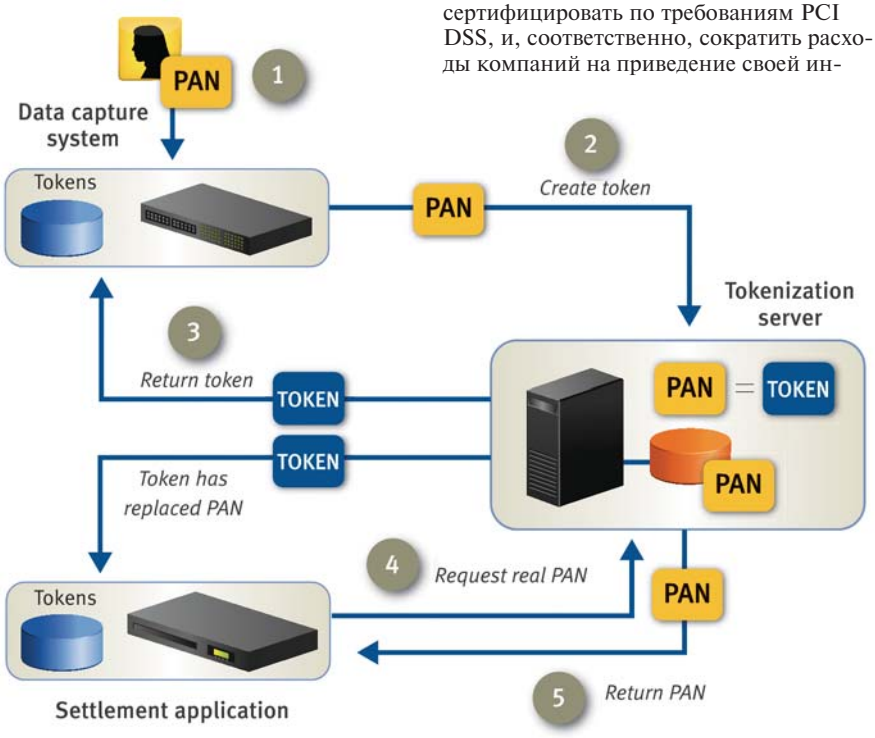

Рис. 10. Сценарий токенизации.

Сценарий токенизации

(1) Полные данные клиента, включая первичный номер счета (primary account number – PAN), вводятся в систему сбора данных.

(2) PAN посылается в простом текстовом формате серверу токенизации, где они записываются в базу данных токенов.

(3) Токен возвращается в систему сбора данных, где он заменяет PAN во всех данных клиента. Также токен посылается другим системам, типа приложений урегулирования.

(4) Если авторизованным приложениям необходим реальный PAN, они посылают запрос серверу токенизации.

(5) PAN возвращается авторизованным приложениям.