АКТУАЛЬНЫЕ МЕТОДЫ ОБЕСПЕЧЕНИЯ ЦЕЛОСТНОСТИ ДАННЫХ ОБЛАЧНОЙ СИСТЕМЫ ХРАНЕНИЯ

CURRENT METHODS ENSURE DATA INTEGRITY OF CLOUD STORAGE SYSTEMS

A. Kuzmin

Summary. It considers problems of data integrity in the cloud storage system. The analysis of actual methods of ensuring data integrity when you use for their treatment for cloud storage. The proposed generalized approach to the selection of current threats methods data integrity when you use for their treatment for cloud storage.

Keywords: integrity, information security, cloud computing, information security

Кузьмин Александр Ростиславович
Аспирант, Санкт-Петербургский национальный исследовательский университет информационных технологий, механики и оптики
alexander.kouzmin@gmail.com

Аннотация. Рассмотрена проблематика обеспечения целостности данных в облачной системе хранения. Проведен анализ актуальных методов обеспечения целостности данных при использовании для их обработки облачных систем хранения. Предложен обобщенный подход к выбору актуальных угроз методов обеспечения целостности данных при использовании для их обработки облачных систем хранения.

Ключевые слова: целостность, защита информации, облачные вычисления, информационная безопасность

Введение

В настоящее время существуют три модели предоставления облачных услуг:

SaaS(Software-as-a-Service) — модель предоставления программного обеспечения как услуги, заключается в предоставлении возможности использования прикладного программного обеспечения, работающего в облачной инфраструктуре и доступного на различных клиентских устройствах. Контроль и управление физической и виртуальной инфраструктурой облака, включая вычислительные сети, сервера, операционные системы, системы хранения, средства защиты информации, прикладное программное обеспечение осуществляется облачным провайдером. Предоставляемое прикладное программное обеспечение в данной модели принадлежит провайдеру.

PaaS (Platform-as-a-Service) — модель предоставления платформы как услуги, под платформой понимается облачная инфраструктура с предустановленным базовым программным обеспечением, которое состоит, как правило, из операционных систем, систем управления базами данных, сред исполнения языков программирования, средств тестирования. Платформа применяется как для размещения ранее разработанного программного обеспечения, так и для разработки нового. Контроль и управление физической и виртуальной инфраструктурой облака, включая вычислительные сети, сервера, операционные системы, системы хранения, средства защиты информации осуществляется облачным провайдером. Предоставляемое базовое программное обеспечение в данной модели принадлежит провайдеру.

IaaS (Infrastructure-as-a-Service) — модель предоставления инфраструктуры как услуги для самостоятельного управления ресурсами обработки, хранения, сетями и другими вычислительными ресурсами, наличие возможности устанавливать и запускать произвольное программное обеспечение, включая операционные системы, ограниченного контроля за набором доступных сетевых сервисов, в т.ч. межсетевым экранированием. Контроль и управление физической и виртуальной инфраструктурой облака, в том числе вычислительной сети в целом, серверов, систем хранения осуществляется провайдером. Физическая и виртуальная инфраструктура принадлежит провайдеру.

Частное облако (private cloud) — модель, при которой принадлежащая одному владельцу облачная инфраструктура имеет множество пользователей аффилированных с владельцем, например, структурные подразделения компании-владельца облака. Частное облако может находиться в собственности или только в управлении владельца или третьей стороны, с которой у владельца заключен договор, физически инфраструктура облака может находиться как внутри, так и за пределами территории владельца.

Публичное облако (public cloud) — модель, при которой инфраструктура предназначается для свободного использования любым пользователем. Публичное облако находиться в собственности, управлении и эксплуатации организаций любой формы собственности, в т.ч. комбинацией таких организаций. Физически публичное облако находится в юрисдикции владельца, который является поставщиком облачных услуг.

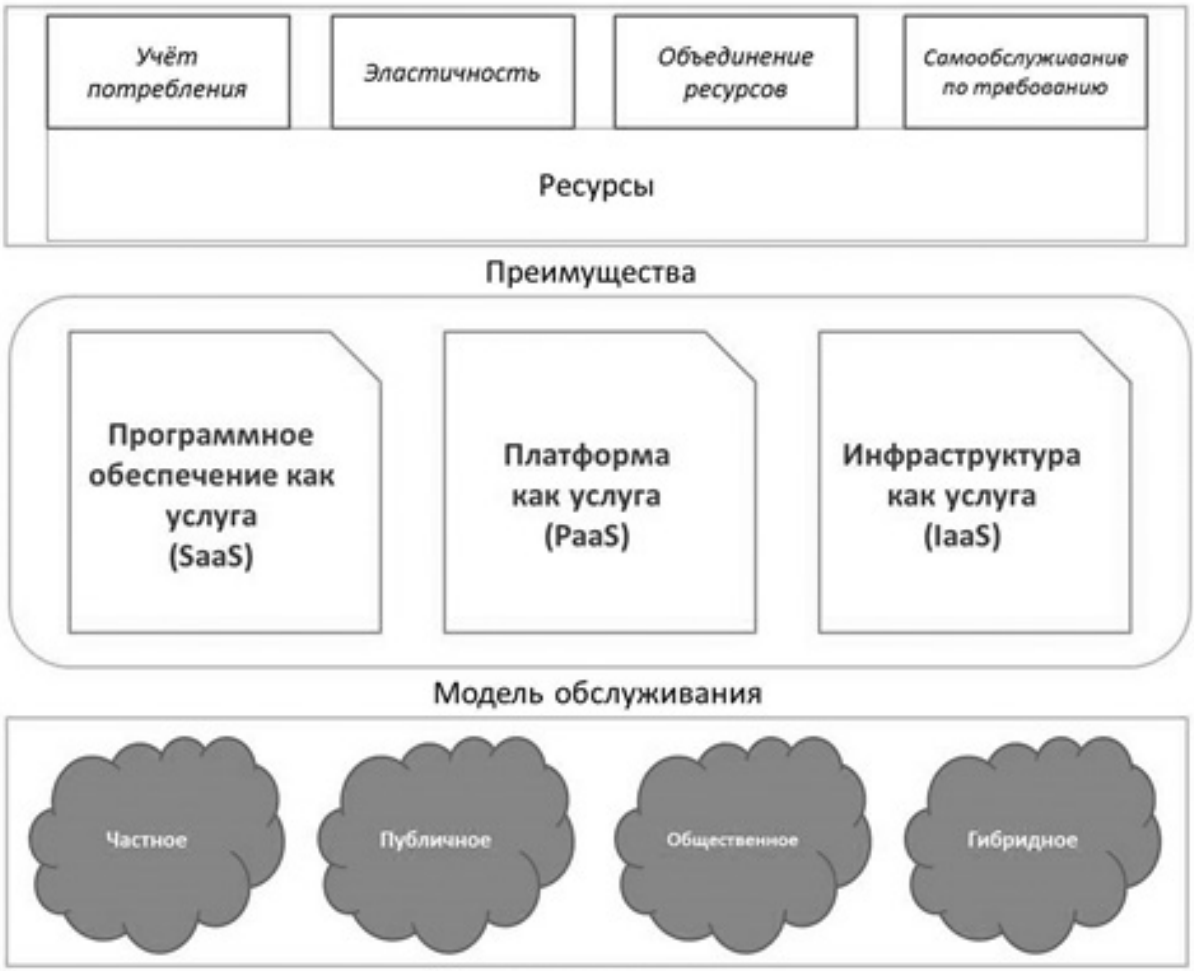

Рис. 1. Визуализация модели облачных вычислений NIST

Общественное облако (community cloud) — модель, при которой инфраструктура, предназначается для использования конкретным сообществом пользователей из организаций, имеющих общие задачи. Общественное облако может находиться как в совместной собственности, управлении и эксплуатации, так и в собственности, управлении и эксплуатации одной или более организаций сообщества или третьей стороны в т.ч. возможны комбинации собственности и управления, эксплуатации. Физически общественное облако может находится как внутри, так и вне юрисдикции владельца.

Гибридное облако (hybrid cloud) — модель-комбинация из двух или более моделей функционирования облачных инфраструктур (частной, публичной или общественной). Не имеет массового распространения, используется, в основном, для балансировки нагрузок между различными облачными инфраструктурами.

Для облачных технологий характерны следующие преимущества:

Самообслуживание по требованию — позволяет самостоятельно определять и выделять потребности в вычислительных мощностях.

Объединение ресурсов — поставщик облачных услуг управляет объединенными ресурсами для обслуживания большого числа пользователей в единый как программный, так и аппаратный пул для динамического перераспределения мощностей между пользователями в условиях постоянного изменения потребности в мощности.

Эластичность — вычислительные мощности и другие сервисы могут быть, как расширены, так и сокращены в любой момент времени в автоматическом режиме.

Учёт потребления — поставщик облачных услуг автоматически вычисляет потреблённые ресурсы.

Передача данных в облачную систему хранения является экономически выгодным шагом для большин-

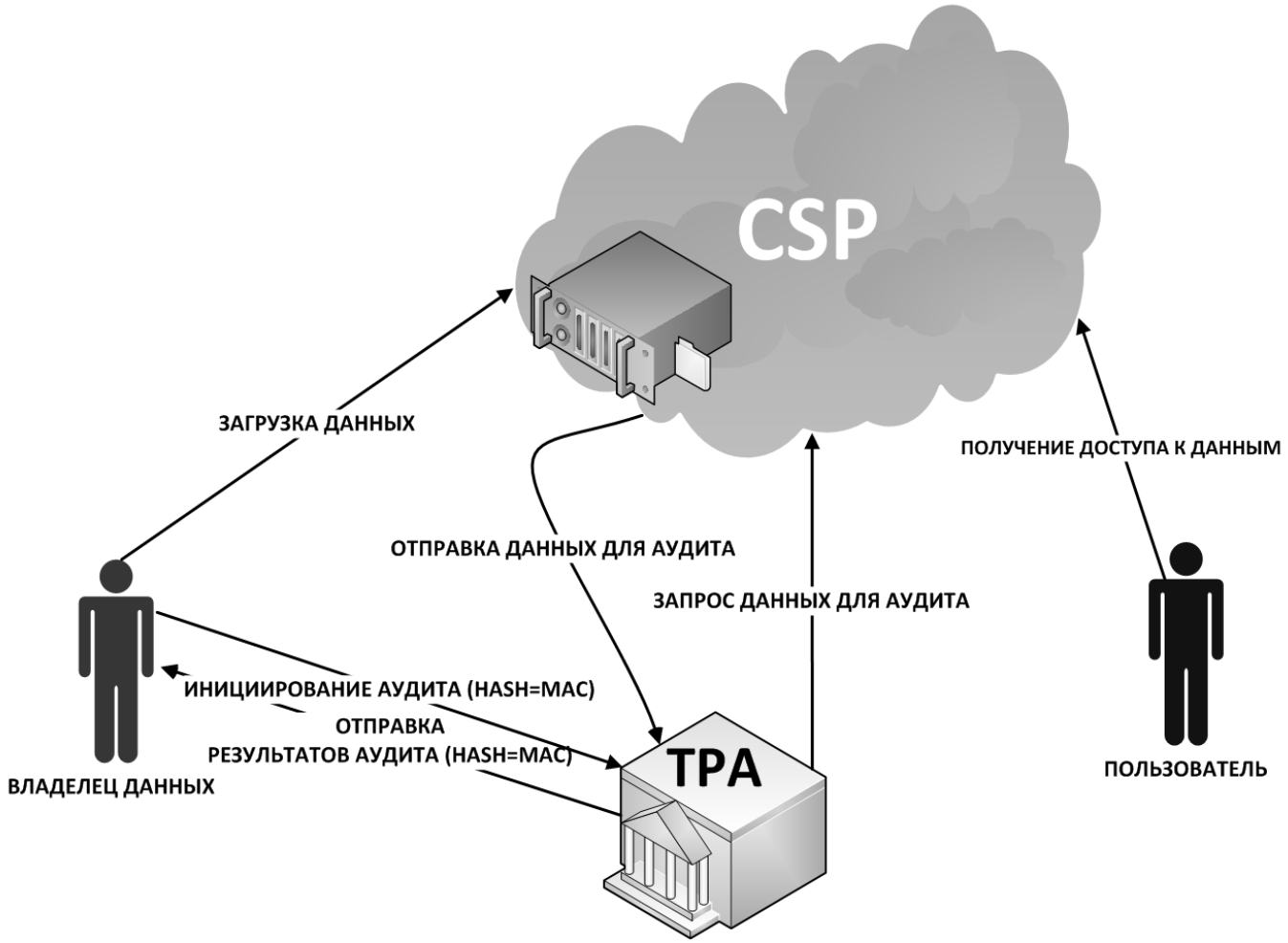

Рис. 2. Схема процесса проверки целостности данных при помощи внешнего аудитора (TPA)

ства организаций любой формы собственности и масштаба. Наряду с обеспечением конфиденциальности и доступности данных переданных в облако, остро встает вопрос обеспечения целостности, в том числе, фиксации факта изменений данных отданных на обработку в облачную систему хранения. Использование частного, т.е. эксплуатируемого непосредственно владельцем данных облачного хранилища не может уберечь от преднамеренного изменения хранящихся в нем данных. Как отмечает Cloud Security Alliance, организация, объединяющая крупнейших поставщиков услуг облачных технологий, потребителей и производителей средств защиты информации, в своих рекомендация по обеспечению безопасности облачных инфраструктур, целостность данных и аудит их изменений должны быть сохранены при переносе данных в облачные хранилища. Особо подчеркивается необходимость постоянного мониторинга состояния целостности данных, необходимость создания средств подобного контроля [1]. Опросы консалтинговой фирмы Ernst & Young проведенные в 2015 году среди крупнейших мировых компаний показывают, что угрозы потери данных и нарушений, связанных с целостностью данных имеют наивысший приоритет и волнуют наибольшее число респондентов [2].

В настоящее время применяемые методы обеспечения целостности данных облачного хранилища подразумевают наличие третьей стороны, помимо владельца (owner) и облачного провайдера услуг (cloud service provider (CSP), внешнего аудитора (Third Party Auditor (TPA). Как показывают исследования, данная технология является весьма уязвимой для атак со стороны злоумышленников [3].

Описание проблемной области

Многие потребители услуг облачных систем хранения полагают, что шифрование данных поможет уберечь их от потери или преднамеренного искажения. Несмотря на необходимость в существенных вычислительных ресурсах для обеспечения своевременного расшифро-

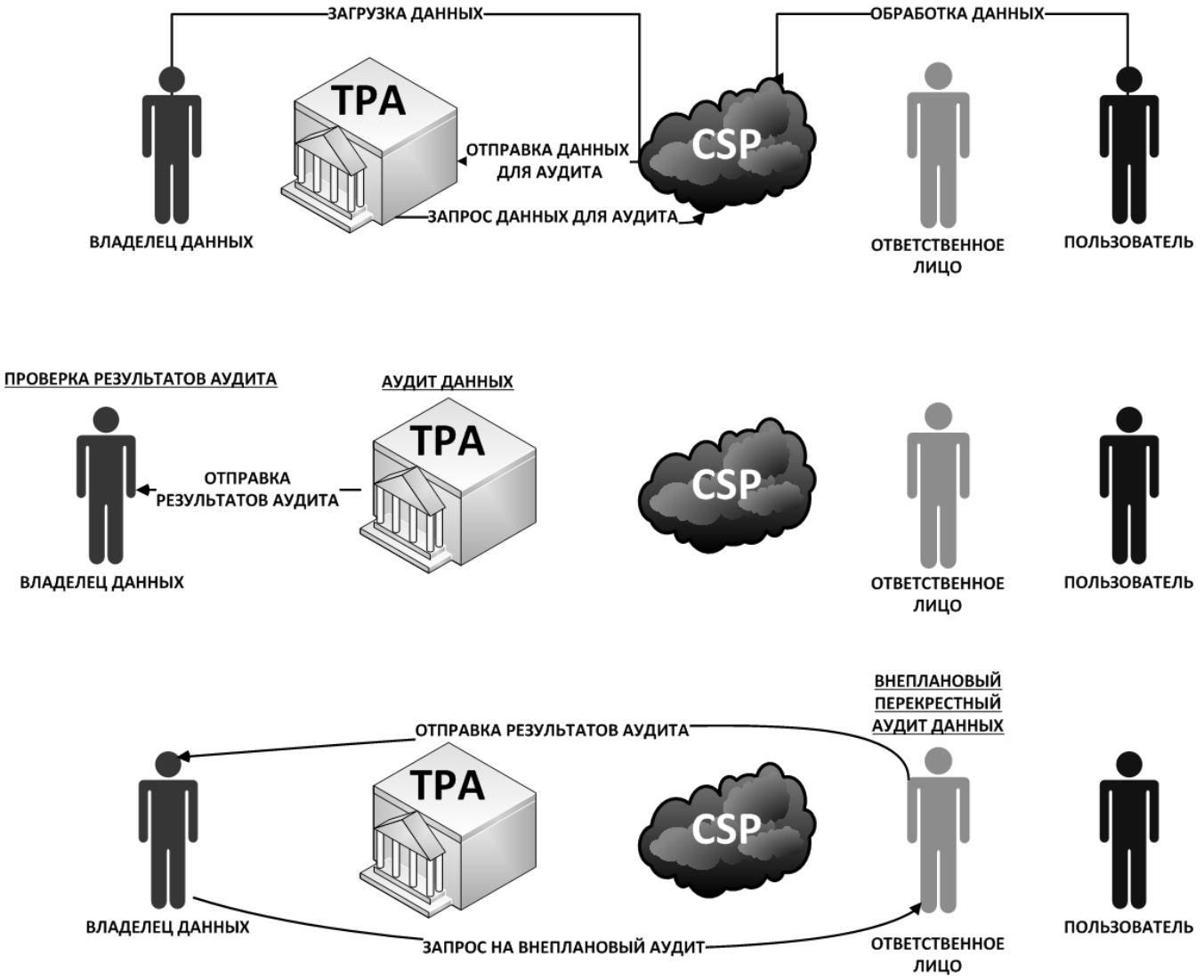

Рис. 3. Схема процесса проверки целостности данных при помощи внепланового, перекрестного аудита.

вания данных для их дальнейшей обработки, в общем случае, шифрование не может обеспечить целостность переданных данных. Которая, может быть нарушена, в т.ч. и не преднамеренными действиями, вызванными, например, сбоями в работе оборудования. В настоящее время, существует два основных метода проверки целостности данных переданных в облачное хранилище, с помощью внешнего аудитор и метод доказуемости владения и его вариаций.

Внешний аудитор (Third Party Auditor (TPA)

При использовании данного метода владелец данных проверяет хэш-сумму файла, которая также выполняет функцию имитовставки для формирования кода аутентичности сообщения (Message Authentication Code (MAC) при отправки запроса внешнему аудитору (Third Party Auditor (TPA). У которого имеется хэш-сумма полученная, непосредственно, в облачной системе хранения данных, к которой у TPA есть доступ на основании трехстороннего договора с владельцем и облачным провайдером (CSP). Если MAC владельца и TPA совпадают, то целостность файла считается подтвержденной. Недостатками данного метода являются потребность в стороннем канале связи, а также, подверженность атакам типа «человек посередине». Кроме того, появление третьей стороны при обработке данных, что повышает риск реализации целого спектра угроз со стороны злоумышленников.

Существует также реализация технологии проверки целостности при помощи нескольких TPA, при которой

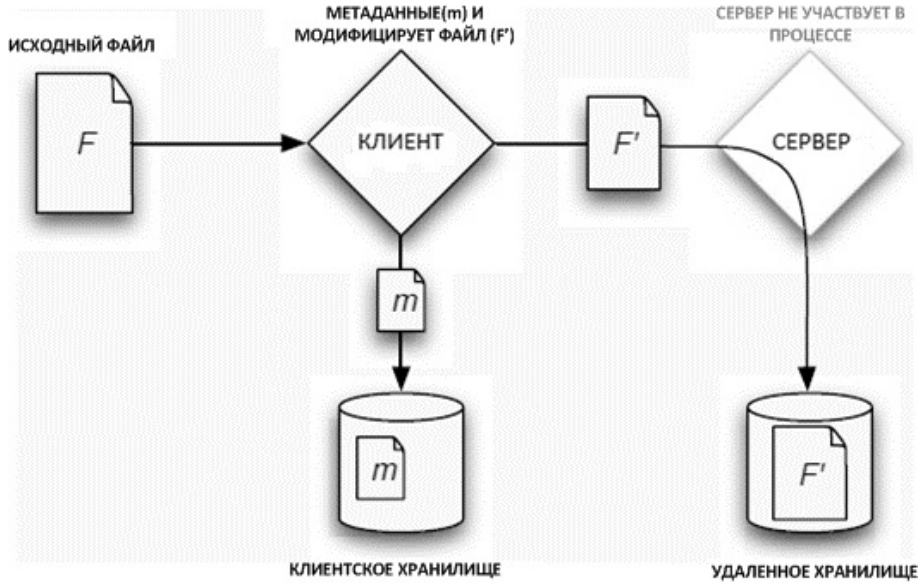

Рис. 4. Предварительная генерация метаданных на стороне пользователя.

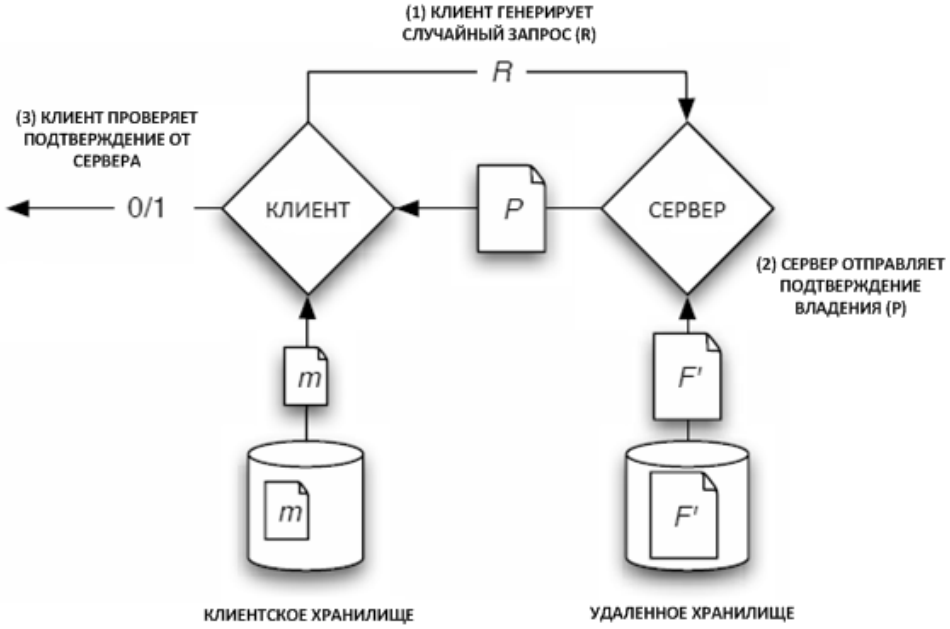

Рис. 5. Проверка факта владения данными.

аудит инициируется через разные промежутки времени, происходит перекрестный аудит с участием TPA и ответственного лица, на стороне обработки данных [4]. Как приведено на рисунке ниже.

Доказуемость владения данных (Provable Data Possession (PDP)

Метод проверки факта хранения данных на удаленном сервере без их извлечения. Данная модель базируется на вероятностном доказательстве факта хранения путем выбора случайных блоков данных с удаленной системы хранения.

К преимуществам данного метода можно отнести минимальный задействованный трафик для проверки целостности данных, а также отсутствие третьей стороны в процессе проверки. Проведенные исследования также показали, что данная модель имела ограничения из-за недостаточной скорости чтения/записи физических устройств хранения информации (HDD), в тоже время, не было выявлено существенных задержек из-за крипто-

графических вычислений. Стоит отметить, что исследования проводились на устройствах типа Hard (magnetic) disk drive (HDD), в то время как скорость чтения/записи физических устройств хранения информации типа Solid-state drive (SSD) существенно выше [5]. Метод использует гомоморфное шифрование для генерации метаданных файла передаваемого в облачное хранилище [6]. На предварительном этапе, перед отправкой файлов в облачное хранилище, пользователь их модифицирует, добавляя метаданных и, вместе с тем, сохраняя метаданные у себя в клиентском хранилище как показано на Рисунке 4.

В дальнейшем при проверке факта неизменности файла в облачном хранилище, пользователь формирует запрос к серверу хранилища на сравнение метаданных файла в хранилище с метаданными находящимися в клиентском хранилище (Рисунок 5).

Дальнейшим развитием метода PDP является метод E-PDP, который отличается от первого протоколом, способным в 185 раз быстрее генерировать метаданные файла и, фактически, ограничивается только скоростью работы устройств чтения/записи физических устройств хранения информации [7]. Существуют, также, вариации метода PDP, такие как Proof of Retrievability (POR), можно перевести как «доказуемость восстанавливаемости», отличается от PDP возможность осуществления внешнего аудита переданного в облачное хранилище файл [8], а также метод «Доказуемость владения» (Proof of Ownership), который, заключается в добавлении в файл некоторого «секрета» перед отправкой в облачное хранилище, для последующего сравнения и, таким образом, доказательства владения пользователем [9].

Основным недостатком перечисленных выше методов является необходимость наличия хранилища метаданных на стороне пользователя, вне облачного хранилища, что увеличивает число возможных объектов для атаки злоумышленником. Кроме того, данные методы существенно затрудняют обеспечение целостности динамических файлов, т.е. файлов в которые вносятся санкционированные изменения, в т.ч. другими авторизованными для этого пользователями.

Подход к выбору актуальных угроз описанных методов

Исходя из описания методов, наиболее актуальными для них угрозами являются угрозы связанные с перехватом и/или изменением ключевой информации подтверждающей обеспечение свойства целостности переданных в облачное хранилище файлов. Если говорить об отличиях, то для метода TPA, также актуальной угрозой является компрометация во время взаимодействия с внешним аудитором. В случае с методом PDP наиболее подверженным угрозам информационной безопасности элементом является клиентское хранилище, т.к. к нему не всегда могут быть применены более высокие требования по обеспечению безопасности, по сравнению, с облачным хранилищем данных.

В общем случае, к обработке актуальных угроз указных методов применимы стандартные подходы их описания. Так, под частотой (вероятностью) реализации угрозы понимается определяемый экспертным путем показатель, характеризующий, насколько вероятным является реализация конкретной угрозы безопасности того или иного метода обеспечения целостности. Вводятся четыре вербальных градации этого показателя:

- маловероятно — отсутствуют объективные предпосылки для осуществления угрозы;
- низкая вероятность — объективные предпосылки для реализации угрозы существуют, но принятые меры существенно затрудняют ее реализацию;
- средняя вероятность — объективные предпосылки для реализации угрозы существуют, но принятые меры обеспечения безопасности недостаточны;
- высокая вероятность — объективные предпосылки для реализации угрозы существуют, и меры по обеспечению безопасности не приняты.

При составлении перечня актуальных угроз безопасности применяемого метода каждой градации вероятности возникновения угрозы ставится в соответствие числовой коэффициент Y_2, а именно:

0 — для маловероятной угрозы;
2 — для низкой вероятности угрозы;
5 — для средней вероятности угрозы;
10 — для высокой вероятности угрозы.

С учетом изложенного коэффициент реализуемости угрозы Y будет определяться соотношением

$$Y = (Y_1 + Y_2) / 20 .$$

По значению коэффициента реализуемости угрозы Y формируется вербальная интерпретация реализуемости угрозы следующим образом:

- если $0 \le Y \le 0,3$, то возможность реализации угрозы признается низкой;
- если $0,3 < Y \le 0,6$, то возможность реализации угрозы признается средней;
- если $0,6 < Y \le 0,8$, то возможность реализации угрозы признается высокой;
- если $Y > 0,8$, то возможность реализации угрозы признается очень высокой [10].

Заключение

Задача обеспечения целостности данных в облачной системе хранения нуждается в появлении новых методов ее решения. Описанные в данной статье методы имеют ряд существенных недостатков и ограничений по функционалу. Так, если придерживаться основных принципов обеспечения целостности, сформулированных Кларком и Вилсоном [11], целостность это корректность транзакций, аутентификация пользователей, минимизация привилегий, разделение обязанностей, аудит произошедших событий, объективный контроль, управление передачей привилегий. А, следовательно, новые методы обеспечения целостности данных в облачной системе хранения должны соответствовать этим принципам, кроме того, позволять контролировать целостность в режиме реального времени и не зависеть от третьих лиц. Наиболее перспективным для решения данного функционала, вероятно, станет применение технологии цепочки блоков транзакций, также известной как блокчейн.

ЛИТЕРАТУРА

1. CSA Security Guidance for Critical Areas of Focus in Cloud Computing V3.0. 2011 Cloud Security Alliance
2. Creating trust in the digital world. EY's Global Information Security Survey 2015. 2015 EYGM Limited
3. Dr.Nedhal A. Al-Saiyd, Nada Sail. Data integrity in Cloud computing security. Journal of Theoretical and Applied Information Technology. 31.12.2013. Vol. 58 № 3 pp. 570–581
4. Sultan Aldossary, William Allen. Data Security, Privacy, Availability and Integrity in Cloud Computing: Issues and Current Solutions. (IJACSA) International Journal of Advanced Computer Science and Applications, Vol. 7, No. 4, 2016 pp.485–498
5. A Virtual Storage Environment for SSDs and HDDs in Xen Hypervisor. Yu-Jhang Cai, Chih-Kai Kang y and Chin-Hsien Wu Department of Electronic and Computer Engineering National Taiwan University of Science and Technology, Taipei, Taiwan. ACM SIGBED Review, Volume 11 Issue 2, June 2014, Pages 39–44
6. Н. Варновский, А. Шокуров. Гомоморфное шифрование // Российская Академия наук Институт Системного Программирования, 2006, с. 27
7. Giuseppe Ateniese, Randal Burns. Provable Data Possession at Untrusted Stores. 14th ACM Conference on Computer and Communications Security(CCS2007)
8. Kevin D. Bowers, Ari Juels, and Alina Oprea RSA Laboratories, Bedford, MA, Proofs of Retrievability: Theory and Implementation, CCSW '09: Proceedings of the 2009 ACM workshop on Cloud computing security, 2009
9. S. Halevi, D. Harnik, B. Pinkas, and A. Shulman-Peleg, «Proofs of ownership in remote storage systems,» in Proceedings of the 18th ACM conference on Computer and communications security. ACM, 2011, pp. 491–500.
10. Методика определения актуальных угроз безопасности персональных данных при их обработке в информационных системах персональных данных» (утв. ФСТЭК РФ 14.02.2008)
11. D. Clark, D. Wilson. A compassion of Commercial and Military Computer Security Policies.— Thr 1987 IEEE Symposium on Security and Privacy, 1987