УДК 372.8:002
К 23

ПРИНЦИПЫ И МЕТОДЫ РАСЩЕПЛЕНИЯ ИНФОРМАЦИИ В РАСПРЕДЕЛЕННЫХ БАЗАХ ДАННЫХ

Карипжанова А.Ж., докторант, II курс.
Научный руководитель: Сагиндыков К.М. к.т.н., доцент, заведующий кафедры «Информатики и информационной безопасности»
Евразийский национальный университет имени Л.Н.Гумилева г. Астана

Идея распределения данных по множеству компьютеров, связанных сетью [1], связана с возможностью решения нескольких перманентно актуальных задач, среди которых можно выделить следующие: 1) расширение возможностей безопасного хранения информации: 2) оптимизация Big Data; 3) обеспечение прозрачности (т.е. невидимости) информации. Эти же задачи опосредуют облачные системы, которые и есть по сути, распределенные базы, причем проблема безопасности в них нивелирует остальные задачи.

В этом направлении может оказаться весьма перспективным разрабатываемый нами в Казахском гуманитарно-юридическом инновационном университете (г. Семей) метод расщепления информации, использующий, как основу, принцип распределения данных [2]. На базе этой технологии расщепления информации в распределенных базах данных уже разработаны гибридные облачные архитектуры, кибер-безопасность систем защиты, защищенные каналы связи и иные инновационные решения по хранению информации в распределенных базах данных [3].

Основным в нашем решении является инновационный метод расщепления данных в системе распределенного хранения. В результате применения метода новый класс кодов расщепляет Big Data в большое количество файлов, каждый из которых не может содержать даже одного бита исходной информации. Разделенные файлы распределяются по множеству серверов, запрограммированных на самовосстановление и самосохранение, что обеспечивает постоянную сохранность данных и безопасность. Отдельные расщепленные данные сами по себе не несут осмысленной информации. Другой аспект связан с тем, что конфигурацию расщепления/восстановления можно составить таким образом, чтобы восстановление данных могло быть выполнено с применением только части расщепленных данных, то есть можно обеспечить устойчивость к потере данных.

Систему нельзя расшифровать, потому что это не шифрование. Она может выдержать любые атаки. Многослойная инфраструктура безопасности и уникальная устойчивость к гибели мест хранения обеспечивает гарантированную защиту от самых изощренных хакерских атак. Благодаря

инновационному иерархическому протоколу доступа, технология предотвращает несанкционированный доступ к данным и хищение информации инсайдерами. Наконец, технология гарантирует 100% безопасной передачи данных через открытые интернет-каналы и частные сети.

Нами используется авторская система распределенного хранения с применением кодов расщепления/реконструкции устойчивых к частичным утерям мест хранения, созданная на базе патентованных технологий. В разработанной технологии применяются коды четности с устойчивостью к множественным отказам. Целевая платформа, на которой функционирует система, – MS Windows XP/Vista/7/8 с .NET Framework v.4/4.5. Система распределенного хранения с расщеплением данных позволяет хранить и обрабатывать клиентские данные на распределенных узлах. Данные разделяются на части и распределяются по узлам системы.

Это новый подход в хранении данных, не имеющий близких аналогов. Принципиальное отличие от существующих аналогов именно в новой парадигме в сфере безопасности, в возможности реализовать внутренне не противоречивую и актуальную модель безопасности хранимых и обрабатываемых данных с неизмеримо более высокой степенью защиты от внешнего вторжения.

Существует всего несколько примеров реализаций подобной идеи расщепления в распределенных базах данных – это платформа Clever Safe (распределенное хранение на серверах фирмы с использованием аппаратно-программного комплекса на базе iSCSI и собственного патентованного алгоритма) [4], технология Symform Cooperative Storage Cloud, использующая клиентские компьютеры как часть системы, а также использующая RAID-96 – собственный вариант RAID [5], и Wuala – похожая на Symform технология с использованием алгоритма Соломона-Рида [6]. Все эти системы предлагают услуги распределенного хранения на своих серверах и применяют различные алгоритмы расщепления/восстановления, предполагая загрузку через интернет данных на серверы фирм. Для защиты от несанкционированного доступа применяются методы шифрования. Сами услуги либо платные, с установкой специализированных аппаратных устройств (Clever Safe), либо востребуют встречные услуги в виде предоставления свободного места на диске своего компьютера, которое система использует для своих целей (диск как составная часть системы распределенного хранения), и/или обеспечения постоянного доступа к компьютеру через интернет извне. Примером стандартной системы обеспечения безопасности хранения может служить, например, портал aws.amazon.com: сервис облачного хранения S3 (Simple Storage Service) и EBS (Elastic Block Store) [7], где используются технологии кластеризации и репликации.

Наиболее близким аналогом, который выбран в качестве прототипа нашей системы, является Object Storage – архитектура хранения данных, которая управляет данными как объектами, в отличие от других архитектур

хранения как файловых систем, которые управляют данными в виде иерархии файлов и блоков хранения [8].

Сравнение параметров нашей системы распределенного хранения с расщеплением данных и параметров Object Storage позволяет сделать следующие заключения:

1) Безопасность информации.

В нашей технологии используется ключ, который не может быть декодирован, потому что это не кодирование. Соответственно, технология лучше всего подходит для публичных систем, в том числе облачных, для хранения конфиденциальной информации, обеспеченного локального облачного хранения. Для сравнения: в технологии Object Storage используется стандартное кодирование.

2) Хранение данных.

В нашей технологии 100% исходных данных могут быть восстановлены даже тогда, когда до 70% файлов теряются. Что практически устраняет необходимость резервного копирования и репликации. А в технологии Object Storage (как и в Block-based Storage) требуются многократные повторы операций и регулярное резервное копирование.

3) Масштабируемость.

В технологии распределенного хранения с расщеплением данных – встроенный уровень масштабируемости. Так как инфраструктура растет, хранение данных и безопасность улучшаются автоматически и экспоненциально. В технологии Object Storage – отличная масштабируемость, но без автоматического увеличения хранения и безопасности. Для сравнения: в технологии Block-based Storage расширение блока хранения данных на базе системы на несколько петабайт отрицательно влияет на прочность, оказывает давление на инфраструктуру хранения и увеличивает затраты на управление.

4) Обмен данными.

В нашей технологии – обеспеченный протокол обмена через развитую систему доступа к информации и обмену, гибкая система безопасности от иерархического доступа к одноранговому. Возможность реализации различных пользовательских конфигураций. А технология Object Storage не предоставляет протокол общего доступа. Обмен данными возможен только с операционной системой или со специализированными продуктами сторонних производителей.

5) Поддержка метаданных.

В нашей технологии Клиент поддерживает метаданные, пользовательские метаданные и метаданные иерархии. Есть поддержка иерархической системы хранения. Администратор ограничен в операциях, связанных с перераспределением данных в архитектуре хранения параметров, таких как добавление новых узлов. Клиент контролирует доступ к данным, добавляя или удаляя пользователей. Для сравнения: в технологии Object Storage системный администратор Data-центра набирает метаданные. Как горизонтально распределенная система хранения, эта технология не

поддерживает иерархическую структуру. Иерархический доступ создается операционной системой.

6) Доступ персонала к информации.

В технологии распределенного хранения с расщеплением данных нет одного мета-ключа для доступа ко всем хранимым данным. Каждый ключ имеет конкретного пользователя. Гибкий протокол безопасности относится к категориям доступа и обмена данными. В технологии Object Storage один мета-ключ обеспечивает доступ ко всем хранимым данным.

Основное экономическое преимущество системы, разработанной нами, – обеспечение повышенной безопасности данных и более низкая стоимость хранения данных по сравнению с конкурирующими решениями. Она гарантирует защиту от внешних и внутренних атак, а также предлагает уникальное решение для защиты от потери данных – оригинальные данные могут быть восстановлены на 100% даже при 98% потере сплит-файлов, что значительно снижает финансовые затраты на обеспечение безопасности.

С точки зрения безопасности, наш метод предлагает инновационные гибридные решения, обеспечивает защиту против современных методов криптоанализа, несанкционированного доступа к данным, хищения данных инсайдерами, спуфинга и DoS/DDoS-атак, что не достижимо аналогами. Сделав хранение информации более защищенным, метод будет иметь трансформирующее воздействие на облачные вычисления, информационную безопасность, хранение данных и практически любой другой аспект, который будет способствовать развитию и укреплению глобальной экономики и выполнит свою обязанность по обеспечению неприкосновенности частной жизни и конфиденциальности.

Список литературы:

1. Дейт К.Дж. Введение в системы баз данных / 8-е издание. – М.: Вильямс, 2005. – 1328 с.

2. Kurmanbaev E.A., Syrgabekov I. N., Zadauly E. Karipzhanova A.Zh., Urazbaeva K.T. Information Security System on the Basis of the Distributed Storage with Splitting of Data // International Journal of Applied Engineering Research. – 2017. – Vol. 12. – № 8. – pp. 1703-1711.

3. Задаулы Е., Курманбаев Е., Сыргабеков И. Инновационная система безопасности на базе распределенного хранения информации с расщеплением данных // Patriot Engineering. – №2 (7). – 2015. – С. 111-119.

4. How Cleversafe Works. – http://www.cleversafe.com/overview/how-cleversafe-works.

5. The Smartest Cloud Security. – http://www.symform.com/how-it-works/security.

6. Comparison of encryption schemes. – http://www.wuala.com/en/learn/technology.

7. Security Resources. – http://aws.amazon.com/security/security-resources.

8. Mesnier M., Gregory R.G., Riedel E. Object-Based Storage // IEEE Communications Magazine. – 2003. – August. – P. 84-90.