Security, Risk and Governance

# Примеры возможных архитектур Micro Focus SecureData

## Содержание

стр.

# Введение

«Общие положения о защите данных» (GDPR), призванные защитить от утечек персональные данные граждан и резидентов Евросоюза, стали самым важным шагом в области обеспечения конфиденциальности данных за последние несколько десятилетий. Закон, вступивший в силу 25 мая 2018 года, предусматривает суровые штрафы — до 4% годового дохода для организаций, не соблюдающих его.

GDPR предписывает ряд конкретных мер по защите данных, принадлежащих субъектам ЕС, в том числе по обеспечению информационной безопасности, требует своевременно уведомлять об утечках и шифровать персональные данные, относящиеся к некоторым категориям конфиденциальности. Соблюдение GDPR во многом сводится к выполнению действий, гарантирующих надежную защиту данных.

Для защиты персональных данных GDPR рекомендует использовать псевдонимизацию и шифрование. Имеется огромное количество информации о том, какие именно данные нуждаются в защите, но общедоступных сведений о способах внедрения технологий и процессов, обеспечивающих безопасность данных, относительно немного.

В этом документе описаны типовые варианты применения псевдонимизации и шифрования, приводится обзор платформы Voltage SecureData от Micro Focus, а также входящих в ее состав базовых технологий. Кроме того, описываются архитектуры и стратегии, с помощью которых защищают персональные данные два наших клиента:

- крупный европейский оператор сотовой связи, использующий технологии SecureData для защиты информации об абонентах, хранимой в озере данных Hadoop;
- глобальный оператор платежной системы и эмитент банковских карт, защищающий данные в процессе переноса в облако и применяющий ту же архитектуру для обеспечения конфиденциальности персональных данных клиентов в своей локальной среде.

# Псевдонимизация и шифрование: в чем отличие?

Псевдонимизация и шифрование обозначены в GDPR в качестве конкретных способов, которые целесообразно применять для защиты персональных данных. Термин «псевдонимизация» широко используется как обозначение методов деидентификации данных (удаления сведений, позволяющих установить личность), с помощью которых создаются «псевдонимы» или суррогатные данные для бизнес-процессов. Примерами псевдонимизации являются шифрование на уровне полей и токенизация.

В GDPR намеренно не указаны конкретные методы шифрования и псевдонимизации, рекомендованные к применению. В докладе IDC «Выполнение требований GDPR с помощью новых методов шифрования и управления ключами» (Enabling GDPR Compliance Through Innovative Encryption and Key Management Approaches) упоминаются традиционные схемы шифрования, обеспечивающие нераспознаваемость данных и вместе с тем нарушающие привычный ход бизнес-процессов. Так или иначе, GDPR предусматривает наличие у методов шифрования двух важных свойств: возможности расшифровки данных при возникновении такой необходимости и сохранение условий для использования зашифрованных данных в рамках бизнес-процессов. Технология Voltage Hyper Format-Preserving Encryption (FPE) с шифрованием, сохраняющим исходный формат данных, позволяет выполнить эти и другие требования в масштабе предприятия.

# Варианты применения псевдонимизации и шифрования

- **«Безопасная аналитика» для хранилищ данных и решений на Hadoop.** Системы обработки Больших данных, в том числе платформы хранилищ данных Teradata, Vertica и Hadoop, способны предоставить организациям поистине безграничные возможности получения новых аналитических выводов и повышения операционной эффективности. Организации принимают и передают информацию, непрерывно анализируя и сохраняя в полях данных конфиденциальные сведения, в том числе имена, адреса домов и электронной почты, информацию о местонахождении, номера телефонов и банковских счетов. Чтобы такие платформы окупались, эти сведения должны быть доступны специалистам по исследованию и анализу данных.

  Но расширение круга лиц, имеющих доступ к конфиденциальным данным, порождает риск утечек, возникающих в результате манипуляций недобросовестных инсайдеров, ошибочных действий или из-за брешей в системах безопасности сторонних организаций. В глобальных компаниях, где данные из представительств, находящихся в различных странах ЕС, передаются в центральный репозиторий или озеро данных, возникают дополнительные проблемы, связанные с соблюдением GDPR и требований в отношении мест хранения данных. При использовании перечисленных выше платформ защита данных с помощью Voltage FPE позволяет выполнять аналитическую обработку деидентифицированной информации. Такой подход уменьшает риск утечек и позволяет предприятию обеспечивать выполнение требований GDPR и других регулятивных норм.

- **Переход в облако.** Чтобы получить заметные рыночные преимущества и сократить затраты, организации реализуют облачные стратегии развития ИТ-инфраструктуры. Однако переход на облачные сервисы создает определенные трудности для бизнеса и осложняет соблюдение требований регуляторов, обусловленные особенностями облачной архитектуры. Когда данные, хранящиеся в открытом виде, зашифровываются, уменьшается вероятность раскрытия конфиденциальных сведений (например, медицинских или финансовых) и появляется возможность существенно уменьшить как масштабы аудиторских проверок, так и расходы на соблюдение требований регуляторов.

- **Оперативная защита данных в эксплуатируемых системах.** Организации хранят и обрабатывают конфиденциальные данные в различных бизнес-приложениях, базах данных и других системах. Обычно они защищены средствами безопасности, предназначенными для инфраструктуры и сетей: межсетевыми экранами, списками контроля доступа и системами мониторинга активности. Технология шифрования данных на уровне полей таблиц не позволяет получить доступ к реальным персональным данным злоумышленникам, если последним удается обойти традиционные средства защиты. Расшифровывать такие данные и использовать их в режиме реального времени могут только доверенные приложения и пользователи, имеющие право доступа.

- **Средства разработки и тестирования.** Серьезные сложности с точки зрения обеспечения безопасности и управления рисками создает подготовка данных для сред разработки и тестирования. Когда информация берется из рабочих баз данных и используется в открытом виде, на незащищенных серверах и рабочих станциях накапливаются большие объемы конфиденциальных данных.

Контроль над ними теряется, и предприятие подвергается ненужному риску, который еще более увеличивается, если компания пользуется аутсорсинговыми услугами по разработке и контролю качества ПО. В условиях угрожающего роста

### Анонимизация данных с помощью Voltage FPH — хеширования, сохраняющего исходный формат

В определенных ситуациях, в частности при обеспечении выполнения требований Статьи 17 GDPR («Право на забвение»), возможность восстановления маскированных данных может создавать ненужный риск или быть явно нежелательной, как, например, в случае обеспечения долговременного соблюдения права на забвение. Функция хеширования, сохраняющего исходный формат данных Voltage Format-Preserving Hash (FPH), обеспечивает, как и FPE, сохранение структуры, логики, применимости к фрагментам полей и т. п., добавляя к этим преимуществам невозможность восстановления исходных данных. В отличие от традиционных методов преобразования без возможности восстановления, таких как SHA-256, функция FPH предоставляет возможность высокопроизводительной работы с данными благодаря более гибкому и не разрушающему ИТ-ландшафт подходу к маскированию данных.

числа утечек и принятия регуляторами все новых требований, таких как GDPR, становится очевидной потребность в деидентификации конфиденциальных данных, переносимых из рабочих систем в среды разработки, тестирования и обучения.

## Критерии выбора технологий шифрования и псевдонимизации

Приведем ряд технологических факторов, которые следует иметь в виду организациям, выбирающим средства защиты персональных данных, осуществляемой с помощью шифрования и псевдонимизации.

**Шифрование с сохранением исходного формата данных Voltage FPE**
Организации, которые должны соблюдать требования GDPR, могли на протяжении многих лет хранить и обрабатывать конфиденциальную информацию в различных базах данных, приложениях и системах. Попытки защитить эту информацию с помощью традиционных методов шифрования приведут к созданию данных, которые окажутся несовместимыми с существующими схемами, структурами и требованиями к обработке. Шифрование структурированных полей с заданным форматом (например, содержащих имена клиентов, номера паспортов и других удостоверений личности, номера телефонов, данные геолокации и даты рождения) предусматривает существенное изменение схемы баз данных и приложений. Кроме того, для каждой операции анализа и обработки требуется расшифровка данных, в результате снижается общий уровень безопасности и появляются дополнительные расходы на управление ключами.

Voltage FPE — это принципиально иная технология, с помощью которой ориентированная на данные платформа Voltage SecureData обеспечивает высокую безопасность на всем протяжении жизненного цикла. Вот ключевые преимущества шифрования данных с использованием Voltage Hyper FPE:

- сохраняются формат и структура данных;
- сохраняется логическая структура данных, в частности контрольные суммы и допустимость дат;
- сохраняются неконфиденциальные значения в составе зашифрованных полей (частичные поля);
- по мере необходимости сохраняются связи с другими полями и целостность ссылок;
- сохраняются смысл данных и отношения между записями, что необходимо для выполнения аналитической обработки данных.

Эти преимущества позволяют приложениям, аналитическим процессам и СУБД использовать защищенные данные в подавляющем большинстве случаев, в том числе в распределенных системах, на различных платформах и с применением многообразных инструментов. При этом защита обеспечивается на уровне целого поля или его части с сохранением доступности неконфиденциальных фрагментов полей для приложений и маскированием конфиденциальных данных. При необходимости Voltage FPE сохраняет ссылочную целостность в наборах данных, позволяя ссылаться на защищенные данные и соединять их. Это особенно важно, когда телефонные номера, номера документов и другие традиционные идентификаторы используются в качестве ссылок между разнородными наборами данных.

В Voltage FPE алгоритм AES-FF1 реализован в соответствии со стандартом NIST SP-800-38G FPE1, в разработке которого компания Micro Focus принимала активное участие. Применение Voltage Hyper FPE гарантирует предприятиям надежную защиту и соответствие стандартам.

**Масштабирование с помощью Voltage Stateless Key Management**
Выстраивая защиту многочисленных приложений и различных типов конфиденциальных данных, организации сталкиваются с постоянно возрастающей сложностью масштабирования систем управления ключами шифрования. Традиционные системы этого класса хранят ключи в собственной базе данных или в защищенном хранилище, что создает проблемы с масштабированием, созданием резервных копий и аварийным восстановлением. Если в рамках разнородного или территориально распределенного ИТ-ландшафта используется шифрование на уровне поля, для традиционных систем управления ключами на основе хранилища требуется обеспечить непрерывное резервное копирование, синхронизацию и защиту — весьма обременительные процедуры, которые сами по себе увеличивают риски безопасности и несоблюдения требований регуляторов.

Технология управления ключами Voltage Stateless Key Management предоставляет динамически генерируемые ключи, которые создаются по мере необходимости в защищенном режиме, не хранятся и не нуждаются в использовании СУБД. Нет нужды осуществлять синхронизацию баз данных и их резервное копирование, а риск утраты ключей сведен к минимуму. Voltage Stateless Key Management интегрируется с существующей инфраструктурой управления учетными записями, например с внешними каталогами LDAP. Разрешение на расшифровку или детокенизацию может включать в себя информацию о ролях и группах пользователей, за счет этого упрощается администрирование с использованием системных политик управления учетными записями.

Наличие ролевого механизма контроля доступа к данным на уровне полей делает возможным просмотр и обработку только тех данных, для которых у пользователей и приложений есть соответствующие полномочия. Voltage Stateless Key Management полностью соответствует требованиям современных архитектур приложений благодаря простоте внедрения, высокой производительности, масштабируемости и возможностям распределенной обработки.

**Поддержка широкого круга платформ**
Соблюдение требований GDPR поддерживается для конфиденциальных персональных данных, которые размещаются на разнообразных платформах и системах, в том числе:

- в Windows, Linux, HP-UX, Solaris, AIX и на других платформах;
- в базах данных Oracle, DB2, Microsoft SQL Server и др.;
- на предназначенных для критически важных приложений платформах z/OS, HPE NonStop, Stratus Virtual Operating System (VOS) и др.;
- в хранилищах данных на основе платформ Teradata, Micro Focus Vertica и ведущих дистрибутивов Hadoop;
- на облачных платформах, например Amazon Web Services и Microsoft Azure;
- на мобильных устройствах под управлением iOS и Android.

Данные также защищаются в рамках процессов извлечения-преобразования-загрузки (ETL), таких, например, как NiFi, Sqoop, Informatica, IBM DataStage и Microsoft Server Integration Services (SSIS). Кроме того, в современных организациях данные анализируются с помощью различных инструментов бизнес-аналитики.

Основное преимущество Voltage Hyper FPE как технологии защиты на уровне полей состоит в том, что, применяя стойкое шифрование, можно обезопасить данные непосредственно в момент их появления в компании, после чего они остаются защищенными при хранении, перемещении и использовании в различных корпоративных системах. Если

[1] nvlpubs.nist.gov/nistpubs/SpecialPublications/NIST.SP.800-38G.pdf

для выполнения требований GDPR планируется применить шифрование, то необходимо выбирать решения, имеющие встроенные средства поддержки максимально широкого круга платформ и систем.

Развернутое решение Voltage SecureData обычно состоит из двух уровней.

- Уровень 1: виртуальное применение Voltage SecureData обеспечивает аутентификацию, авторизацию, управление ключами и политиками, а также интеграцию с аппаратными модулями безопасности, хранящими корневые ключи, которые применяются при создании временных ключей. Предоставляются защищенные веб-интерфейсы с двойным контролем входа для управления, мониторинга, аудита и использования развернутого решения. Поддерживается централизованное управление политикой преобразования форматов данных для шифрования и токенизации, доступны средства управления аутентификацией и авторизацией, а также централизованные механизмы аудита и мониторинга модулей, работающих на втором уровне.
- Уровень 2: предоставляется ряд гибких, простых в использовании программных интерфейсов API, средств командной строки и инструментов для обработки файлов, а также функции СУБД и компонент User Defined Function, с помощью которых данные можно шифровать и токенизировать. Эти инструменты доступны для ряда платформ, в том числе в нативных версиях решения для Windows, Linux, AIX, HP-UX, Solaris, дистрибутивов Hadoop, Teradata, Micro Focus Vertica, z/OS, HPE NonStop и Stratus VOS.

Поддержка многочисленных платформ и широкие возможности интеграции, которые предоставляет Voltage SecureData, позволяют клиентам выборочно выполнять шифрование и дешифрацию, если это нужно бизнес-процессам и приложениям. Поскольку Voltage FPE сохраняет смысл и логику данных, внедрение решения оказывается не слишком сложным и затратным: в подавляющем большинстве случаев вносить изменения в приложения и процессы не придется, так как они могут работать с зашифрованными данными. Таким образом, по сравнению с традиционными средствами шифрования, при использовании которых интеграция и управление ключами — весьма кропотливое и сложное дело, внедрение Voltage FPE выполняется намного проще.

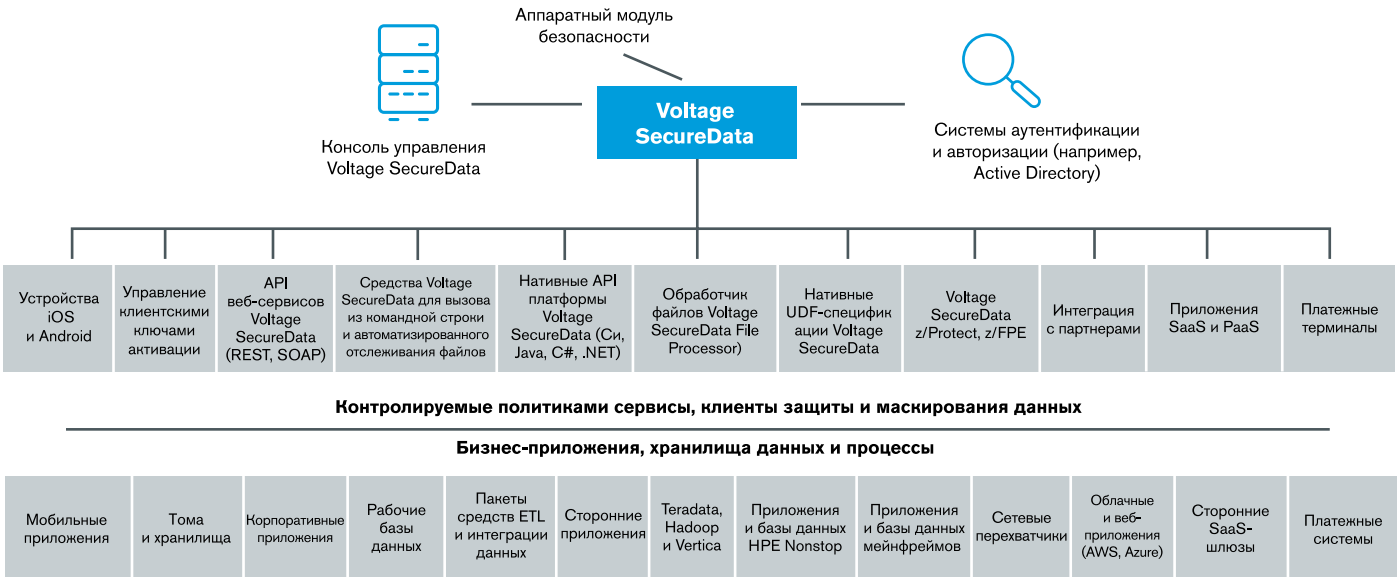


**Рис. 1.** Двухуровневая архитектура Voltage SecureData

**Развертывание компонента SecureData Sentry, ускоряющее получение нужных результатов**

С переходом на гибридные ИТ и все более широкое использование приложений SaaS не всегда удается обеспечить интеграцию на уровне API силами штатных разработчиков. Взаимодействуя с механизмами Уровня 2, брокер конфиденциальности данных Voltage SecureData Sentry перехватывает проходящие по сети конфиденциальные данные и обеспечивает их шифрование; при этом поддерживаются приложения, развернутые как локально, так и в гибридном облаке. SecureData Sentry упрощает переход на гибридные ИТ, ускоряет решение поставленных задач и получение нужных результатов за счет быстрого выполнения требований безопасности, а кроме того, позволяет выстроить цельную сквозную защиту данных без кардинальной модификации приложений и трансформации ИТ-архитектур.

# Примеры архитектур: шифрование Voltage FPE и GDPR

Крупный европейский оператор связи: деидентификация записей о вызовах в Hadoop для «защищенного» анализа

Крупный европейский оператор мобильной связи собирает огромные наборы данных о своих абонентах из нескольких стран ЕС. Эти данные перемещаются в центры обработки, находящиеся в Германии и Италии, где осуществляется их анализ на базе кластера Hadoop, состоящего из 140 нод. По оценкам оператора, ежедневный объем обработки составляет более 11 млрд записей.

### БИЗНЕС-ЗАДАЧА

- Защита огромных массивов данных, содержащих контактную информацию, номера IMEI и IMSI, информацию о местонахождении, абоненте, приложениях, SMS, вызовах и другие персональные данные.
- Выполнение GDPR, а также требований законов отдельных стран о месте размещения данных.
- Шифрование с сохранением исходного формата персональных данных, поступающих из различных европейских представительств с целью выполнения GDPR и законов о месте размещения данных с сохранением возможности их анализа — такая защита необходима для предотвращения мошеннического доступа, выявления шаблонов поведения пользователей и отладки сценариев реагирования на сбои в сети.

### РЕШЕНИЕ

Перед тем как загрузить в Hadoop персональные данные, содержащиеся в записях о вызовах, оператор выполняет их псевдонимизацию, используя для этого сертифицированную NIST технологию Voltage FPE. Перечислим компоненты, развернутые в рамках решения:

- **Серверы ключей Voltage SecureData.** На серверах, установленных в дата-центрах в Германии и Италии, развернут технологический модуль Voltage Stateless Key Management. Размещение серверов ключей в разных странах стало возможным благодаря особенностям архитектуры Voltage SecureData.
- **Обработчик файлов Voltage SecureData File Processor в зоне приема данных.** Во многих системах на базе Hadoop создается зона приема данных (landing zone), где все поступающие данные перед их размещением в HDFS подвергаются

[2] hortonworks.com/apache/ hdfs/

предварительной обработке, форматированию и нормализации2. Чтобы защитить персональные данные внутри файлов перед их загрузкой в Hadoop, используя для этого шифрование, сохраняющее исходный формат, оператор развернул Voltage SecureData File Processor на серверах в зоне приема.

Этот инструментарий шифрует конфиденциальные поля в структурированных файлах различных форматов, в том числе с разделителями-запятыми, XML, JSON, а также в файлах записей с разделителями и файлах с позиционным форматированием.

- **API-интерфейсы Voltage SecureData, встроенные в Apache Spark.** Для быстрой обработки данных (in-memory) при их загрузке в хранилище Hadoop оператор применяет Apache Spark. Клиенту не составило труда интегрировать в него основанные на Java API-интерфейсы Voltage SecureData, чтобы обеспечить шифрование конфиденциальных персональных данных перед их загрузкой в Hadoop с помощью Voltage FPE.

Использование Voltage FPE гарантирует ссылочную целостность и обеспечивает сохранение характеристик псевдонимизированных данных, включая их длину и тип. Вся аналитическая обработка выполняется с использованием псевдонимизированных данных — для этого не требуется деидентифицировать их и приводить в первоначальный вид.

## ПРЕИМУЩЕСТВА

Благодаря развертыванию Voltage SecureData оператор получил ряд преимуществ:

- реализована защита самых ценных и уязвимых систем, в том числе озера данных на основе Hadoop;
- при возникновении бреши в аналитической системе персональные данные не раскрываются, поэтому выполнять требования регуляторов об обязательном в подобных случаях уведомлении клиентов нет необходимости;

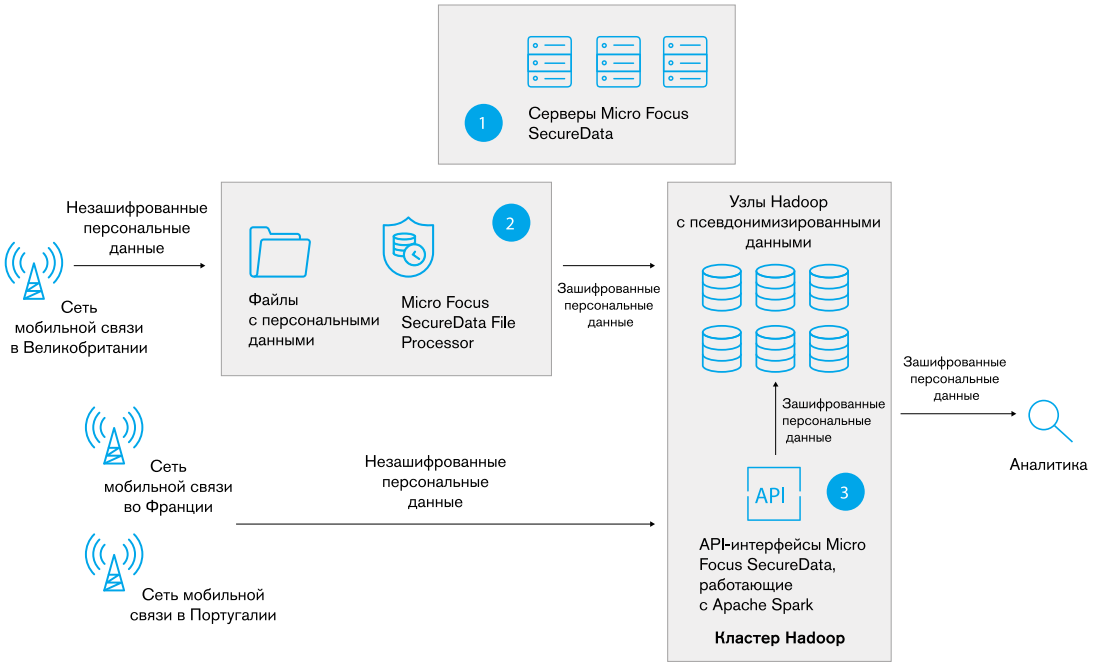


**Рис. 2.** Архитектура деидентификации данных в Hadoop

- обеспечивается соблюдение нескольких европейских законов о месте размещения данных, в том числе GDPR;
- единая масштабируемая платформа уровня предприятия используется для защиты конфиденциальных персональных данных внутри других платформ и систем.

### Глобальная платежная система: шифрование данных перед их переносом в облако

Чтобы уменьшить расходы и ускорить вывод на рынок новых продуктов и услуг за счет использования Agile-методов разработки, оператор глобальной платежной системы и эмитент карт перенес ряд приложений в общедоступное облако Azure. Как показало исследование, перенос в облако одного приложения, хранящего персональные данные, может обеспечить компании более 50% экономии. Реализованное на платформе .NET приложение для анализа операций с картами применялось несколькими подрядчиками, что порождало проблемы в области безопасности и приводило к дополнительным затратам, связанным с предоставлением доступа партнеров к внутренней сети компании. Если бы данные переносились в облако в незашифрованном виде, возникли бы риски, связанные с вероятностью утечки, решением вопросов о юрисдикции данных, а также с возможными нарушениями требований регуляторов, в том числе GDPR.

### Бизнес-задача

- Поддержка крупномасштабной гибридной инфраструктуры, содержащей устаревшие системы, корпоративные приложения и облачные платформы; в их числе — несколько операционных систем (z/OS, Windows, Linux и HPE NonStop) и СУБД (Oracle, Microsoft SQL и DB2). Работа с платформами хранилищ данных, например Teradata, и платформами хранения и обработки данных, такими как Hadoop.
- Мгновенная защита данных конкретных приложений сразу после их переноса в облако.
- Масштабирование с возможностью защиты миллиардов экземпляров персональных данных в сотнях приложений, которые их собирают, хранят и обрабатывают.

### Решение

Для защиты данных, перемещаемых в облако, компания внедрила Voltage Hyper FPE. На рис. 3 изображена архитектура развернутого решения.

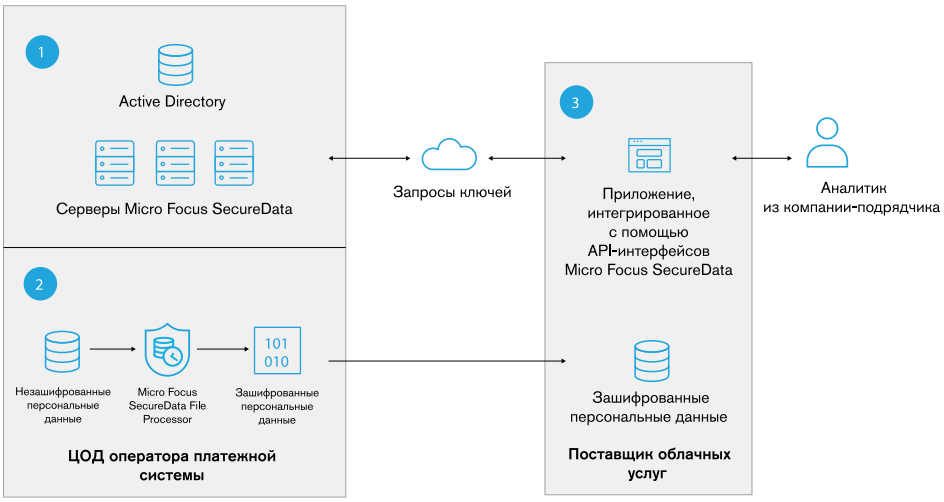


**Рис. 3.** Архитектура системы шифрования данных для облака

- Серверы ключей Voltage SecureData: в сети оператора платежной системы была развернута глобальная инфраструктура серверов ключей Voltage SecureData с балансировкой нагрузки. Их локальная установка обеспечила полный и постоянный контроль над ключами в соответствии с требованиями внутренней политики безопасности и внешних регуляторных актов.
- Перенос данных из приложений в облако: в базах приложений, перемещаемых в общедоступное облако, хранился значительный объем персональных данных. С помощью инструментария Voltage SecureData File Processor и компонента Command line было обеспечено их шифрование с сохранением исходного формата. Эта функциональность интегрирована в инструментарий ETL с помощью скриптов, выполненных в виде командных файлов.
- Интеграция облачных приложений с Voltage SecureData посредством API-интерфейсов: для предоставления аналитикам возможности расшифровывать персональные данные (если это позволено в рамках пользовательской роли), приложение для анализа операций с картами, основанное на платформе .NET, интегрировано с Voltage SecureData посредством API-интерфейсов. Чтобы загрузить ключи для расшифровки, эти API обращаются к серверам ключей, развернутых во внутренней инфраструктуре оператора платежной системы. Каждый вызов аутентифицируется с использованием корпоративной инфраструктуры каталогов Active Directory. Все вызовы заносятся в централизованный журнал и затем учитываются при подготовке уведомлений и различных отчетов.

**Преимущества**

Внедрение Voltage SecureData обеспечило ряд преимуществ:

- реализована защита данных приложений с возможностью масштабирования;
- несколько десятков приложений легко перемещаются в облако, при этом достигается значительное снижение расходов;
- обеспечивается выполнение требований внутренних стандартов безопасности и внешних регуляторных актов, в том числе GDPR.

Развертывание Voltage SecureData было расширено и охватило более 130 приложений, работающих на базе инфраструктуры оператора платежной системы, в том числе приложения на мейнфреймах и Hadoop, а также несколько приложений, функционирующих на базе распределенных операционных систем.

# Заключение

Чтобы обеспечить защиту персональных данных клиентов в соответствии с требованиями GDPR, организациям необходимо внедрять средства шифрования и псевдонимизации. Инновационные разработки, в частности Voltage Hyper Format-Preserving Encryption и Voltage Stateless Key Management, позволяют развертывать технологии защиты данных, масштабируемые в широком диапазоне, и, таким образом, минимизировать модификацию существующих платформ и систем.

Как правило, выполнение требований GDPR обеспечивается поэтапно: в первую очередь компании защищают конфиденциальную персональную информацию в самых уязвимых системах, в том числе в Hadoop, хранилищах данных и приложениях, развернутых в облаке. Накопленный опыт используется для защиты данных множества других приложений, платформ и систем.

Дополнительную контактную информацию и адреса представительств см. по адресу **www.microfocus.com**

www.microfocus.com