

МЕТОДЫ ПЕРЕНОСА ИНФОРМАЦИИ МЕЖДУ ДОМЕНАМИ.

Аннотация: В статье рассматриваются основные методы переноса данных между доменами, включая ETL, облачные сервисы, контейнеризацию и прямое копирование. Описаны примеры их применения и основные проблемы, такие как совместимость схем данных, обеспечение безопасности и управление большими объемами данных. Также предложены решения для эффективного управления процессом миграции, обеспечивающие целостность и безопасность данных.

Ключевые слова: Перенос данных, миграция данных, домены, ETL, облачные сервисы, прямое копирование, схемы данных, безопасность данных, целостность данных.

Fokt V.E

Second year graduate student

Povolzhskiy State University of Telecommunications and Informatics

Russia, Samara

METHODS OF TRANSFERRING INFORMATION BETWEEN DOMAINS

Annotation: This article discusses the primary methods of data transfer between domains, including ETL, cloud services, containerization, and direct copying. Examples of their application and key issues such as data schema compatibility, security assurance, and managing large data volumes are described. Solutions for effectively managing the migration process, ensuring data integrity and security, are also proposed.

Keywords: Data transfer, data migration, domains, ETL, cloud services, direct copying, data schemas, data security, data integrity.

В современном мире данные являются одним из самых ценных ресурсов для организаций. Слияния и поглощения компаний, модернизация ИТ-инфраструктуры, переход на новые программные платформы или оптимизация бизнес-процессов часто требуют переноса данных между различными

доменами. Этот процесс, называемый миграцией данных, включает перемещение данных из одной системы или базы данных в другую. Однако перенос данных между доменами сопряжен с рядом технических и организационных проблем, таких как несовместимость схем данных, обеспечение безопасности, управление большими объемами данных и сохранение целостности данных. В данной статье рассматриваются основные методы переноса данных между доменами и приводятся примеры их применения.

Методы переноса данных

ETL (Extract, Transform, Load)

ETL (Extract, Transform, Load) является одним из наиболее распространенных методов миграции данных. Процесс включает три этапа:

Извлечение (Extract): Данные извлекаются из исходной системы.

Преобразование (Transform): Данные преобразуются в формат, совместимый с целевой системой.

Загрузка (Load): Данные загружаются в целевую систему.

Пример использования ETL

Компания А решила мигрировать свою клиентскую базу данных из локальной системы в облачное хранилище данных Amazon Redshift. Для этого они использовали инструмент Talend, который автоматизировал процесс ETL. Данные были извлечены из старой базы данных MySQL, преобразованы в формат, подходящий для Redshift, и загружены в облачное хранилище. Этот процесс обеспечил плавный переход без потери данных и с минимальными простоями.

Облачные платформы, такие как AWS, Microsoft Azure и Google Cloud, предлагают инструменты для автоматизации и масштабирования процесса миграции данных. Эти сервисы предоставляют высокую производительность, надежность и безопасность данных.

Пример использования облачных сервисов

Компания В мигрировала свою базу данных с локального сервера на платформу Microsoft Azure с помощью Azure Data Migration Service (DMS). DMS автоматизировал процесс миграции, обеспечив бесшовный перенос данных с минимальными усилиями со стороны ИТ-отдела компании. В результате компания получила доступ к облачным возможностям Azure, таким как масштабируемость и улучшенная безопасность.

Контейнеризация

Контейнеризация, с использованием таких технологий, как Docker, позволяет изолировать приложения и их зависимости, что упрощает перенос данных между различными системами и платформами. Контейнеры обеспечивают совместимость и портативность приложений.

Пример использования контейнеризации

Компания С использовала Docker для контейнеризации своей базы данных PostgreSQL перед миграцией на новую платформу Kubernetes. Это позволило изолировать базу данных и её зависимости, обеспечив простоту и гибкость процесса миграции. Контейнеры Docker были развернуты на новой платформе без необходимости изменений в исходной структуре данных, что ускорило процесс миграции.

Прямое копирование

Прямое копирование данных из одной системы в другую может быть выполнено с помощью различных инструментов и скриптов. Этот метод

требует тщательного планирования и контроля для обеспечения целостности и безопасности данных.

Пример использования прямого копирования

Компания D использовала команду `rsync` для прямого копирования файловых данных из одного сервера на другой. Rsync позволил компании выполнять инкрементальные обновления, копируя только изменённые файлы, что значительно сократило время простоя системы и ускорило процесс миграции. Перед началом копирования данные были зашифрованы с помощью OpenSSL для обеспечения их безопасности.

Основные проблемы и методы их решения

Совместимость схем данных

Различные системы и базы данных могут использовать различные схемы данных, что делает процесс миграции сложным. Для решения этой проблемы используются следующие методы:

Сопоставление схем данных: Сравнение и сопоставление элементов данных между исходной и целевой схемами.

Преобразование данных: Преобразование данных в формат, совместимый с целевой схемой.

Инструменты преобразования схем: Инструменты, такие как AWS Schema Conversion Tool (SCT), автоматизируют процесс преобразования схем данных.

Управление большими объемами данных

С увеличением объема данных процессы миграции становятся более трудоемкими. Основные методы решения этой проблемы включают:

Инкрементальная миграция: Перенос данных поэтапно для снижения нагрузки на системы.

Сжатие данных: Использование методов сжатия для уменьшения объема передаваемых данных.

Параллельная обработка: Применение параллельной обработки для ускорения процесса миграции.

Сохранение целостности данных

Обеспечение целостности данных при миграции критически важно. Методы включают:

Валидация данных: Проверка данных на каждом этапе процесса миграции.

Проверка контрольных сумм: Использование контрольных сумм для проверки точности данных до и после миграции.

Обеспечение транзакционной целостности: Гарантия, что транзакции данных являются целостными и атомарными.

Перенос данных между доменами представляет собой важную и сложную задачу, требующую тщательного планирования, использования современных методов и инструментов, а также обеспечения безопасности и целостности данных. Успешная миграция данных позволяет организациям эффективно управлять своими информационными ресурсами, повышать производительность и надёжность систем, а также соответствовать нормативным требованиям. Эффективное решение проблем, связанных с переносом данных, требует интеграции различных подходов и технологий, таких как ETL, облачные сервисы, контейнеризация и прямое копирование.

Использованные источники:

1. Миграция данных. При миграции данных может потребоваться повторное выполнение логики базы данных или приложения, изменение схемы базы данных хранимых процедур базы данных, <https://www.imperva.com/learn/data-security/data-migration/> .

2. Извлечение, преобразование, загрузка (ETL) Контейнеры ETL — это методы автоматической миграции данных для приема данных из различных источников в целевую систему. Конвейер извлекает данные из исходных систем, преобразует их и загружает в хранилище данных, такое как Snowflake., <https://www.starburst.io/data-glossary/etl/> .