**3.1. ESB шина – основные функции. Продукт IBM WebShpere MessageBroker – основные функции.**

**ESB (Enterprise Service Bus)** — дословно можно перевести как «сервисная шина предприятия». ESB описывает вполне реальный программный продукт, в задачи которого входит упрощение вызова службы за счет управления всеми взаимодействиями на пути от потребителя службы к поставщику и обратно. Двумя наиболее часто упоминаемыми возможностями ESB являются преобразование сообщений и их маршрутизация. На шину ESB возложена важнейшая задача обеспечения взаимодействия систем из слабосвязанных сервисов в сети. Корпоративная сервисная шина поддерживает Web-сервисы, реализуя протокол SOAP (Simple Object Access Protocol, Простой протокол доступа к объектам) и используя язык WSDL (Web Services Description Language, Язык описания Web-сервисов) и спецификацию UDDI (Universal Description, Discovery and Integration, Универсальное описание, обнаружение и интеграция).

**Основные функции ESB**

* Обеспечение интерфейсов взаимодействия
* Отправка сообщений и маршрутизация
* Преобразование данных
* Сенсоры событий
* Управление политиками
* [Виртуализация](http://www.tadviser.ru/index.php/%D0%92%D0%B8%D1%80%D1%82%D1%83%D0%B0%D0%BB%D0%B8%D0%B7%D0%B0%D1%86%D0%B8%D1%8F)

**Архитектура ESB**

Основа архитектуры ESB — это идея использования общей интеграционной инфраструктуры всеми корпоративными приложениями на базе обмена сообщениями. Все приложения взаимодействуют через одну точку, которая, в случае необходимости, обеспечивает сохранность обращений, преобразование данных и транзакции. При этом целью интеграции приложения является создание единственного модуля (или адаптера), который отвечает за «подключение» приложения к ESB. Последующую обработку сообщений и их маршрутизацию в другие системы, ESB выполняет на основании установленных бизнес-правил самостоятельно. Этот подход обеспечивает превосходную гибкость, простоту масштабирования и переноса, поэтому в случае замены одного из приложений подключенного к шине, перенастраивать остальные не нужно.

**Достоинствами ESB является:**

* Организация размещения существующих систем осуществляется быстрее и дешевле.
* Повышение гибкости.
* ESB основывается на общепризнанных стандартах.
* Наличие большого количества конфигурации для интеграции.

**К числу недостатков ESB относят:**

* Сложность реализации
* Требует больших ресурсов.

**IBM® WebSphere® Message Broker** - это решение для корпоративной сервисной шины (Enterprise Service Bus, ESB), поддерживающее широкий спектр коммуникационных протоколов и форматов сообщений. Оно обеспечивает выполнение преобразований, маршрутизации и дополнения с использованием различных технологий. Ключевой компонент WebSphere Message Broker - встроенный инструментарий разработки программного обеспечения. Инструментарий WebSphere Message Broker является плагином Eclipse, основанным на IBM Rational Application Developer. Он поможет разрабатывать потоки и наборы сообщений, администрировать преобразователь сообщений и включенные в него компоненты рабочей среды: группы исполнения и работающие в них потоки сообщений.

WebSphere Message Broker содержит возможности разработки потоков сообщений и расширенную поддержку Web-сервисов, включая генерирование и использование WSDL.

Возможность создавать потоки сообщений с нуля, из существующих WSDL- или XSD-файлов или набора сообщений.

Возможность группировать и отображать связанные проекты в виде Broker Development. Это будет полезно, если вы хотите видеть в одном месте какой-то определенный поток и соответствующий набор сообщений.

Возможность тестировать поток сообщений, запущенный в узлах MQInput или HTTPInput.

**3.2. Методы трансформации и транспортировки данных, ETL – преобразования. Определение типов данных, определение узлов преобразований, построение карт преобразований данных.**

Под аббревиатурой **ETL — extraction, transformation, loading**, то есть извлечение, преобразование и загрузка, скрываются три основных процесса, используемые при переносе данных из одного приложения или системы в другие.

Цель практически любого ETL-приложения состоит в том, чтобы своевременно предоставить данные его пользователям. Традиционно предприятия использовали программы ETL для переноса информации из унаследованных приложений в новые.

Однако с активным распространением Internet процессы ETL стали все шире применяться и для поддержки Web-приложений. Например, поставщик может использовать средства ETL для «закачки» в Web-систему данных, необходимых для проверки состояния обработки заказов из внутреннего унаследованного приложения.

В общем и целом приложения ETL извлекают информацию из исходной базы данных, преобразуют ее в формат, поддерживаемый базой данных назначения, а затем загружают в нее преобразованную информацию.

**Извлечение данных**

Для того чтобы инициировать процесс ETL, применяются программы извлечения данных для чтения записей в исходной базе данных и для подготовки информации, хранящейся в этих записях, к процессу преобразования. Чтобы извлечь данные из исходной базы данных, можно выбрать один из трех вариантов — создать собственные программы, обратиться к готовому специализированному инструментарию ETL или использовать сочетание и того и другого.

**Преобразование данных**

После того как процедуры извлечения завершили сбор данных, процедуры преобразования могут подготовить эти данные к размещению в новом месте. Существует несколько методик преобразования, в том числе объединение, перевод значений, создание полей и фильтрация. Предположим, что имеется некий гипотетический киоск данных, на примере которого будут рассмотрены эти процессы.

Киоски данных представляют собой специализированные приложения, которые дают возможность конечным пользователям анализировать самые разнообразные тенденции на весьма интуитивном уровне. Например, аналитик по вопросам маркетинга может использовать киоск данных с информацией о продажах для того, чтобы проанализировать доход в расчете на продукт за каждый из последних пяти лет. В отличие от средств обработки заказов, производства и других операционных приложений, киоски данных не требуют детальной информации. Фактически резюмированные данные даже предпочтительнее, поскольку в этом случае уменьшается время ответа и упрощается использование.

Перед тем как загрузить киоск данных, они, как правило, **агрегируются.** Процедуры агрегирования (aggregation) заменяют многочисленные детальные записи на относительно небольшое число кратких записей. Например, предположим, что данные о продажах за год занимают в нормализованной базе данных несколько тысяч записей. После агрегирования данные преобразуются в меньшее число кратких записей, которые будут перенесены в киоск с данными о продажах.

**Перевод значений (value translation)** — еще одна распространенная методика преобразования данных. Операционные базы данных хранят закодированную информацию для того, чтобы сократить избыточность данных и удовлетворить требования к хранению.

Поскольку киоски данных содержат резюмированную информацию и рассчитаны на то, что их просто использовать, программисты обычно заменяют закодированные данные на более понятные описания, решая эту задачу с помощью написанных ими программ. Однако инструментальные средства ETL более эффективны, поскольку позволяют для декодирования данных использовать списки перевода значений.

**Создание полей (field derivation)** — третья методика, применяемая для преобразования данных. При создании полей для конечных пользователей создается и новая информация. Например, предположим, что операционная база данных содержит одно поле для указания количества проданных товаров, а второе — для указания цены одного экземпляра. Для того чтобы определить доход, пользователям не придется считать его самим, поскольку программисты могут создать специальное поле для указания дохода во время преобразования данных.

Четвертая процедура преобразования — **очистка (cleansing)** используется очень многими. Программисты применяют алгоритмы очистки данных для того, чтобы отсеять неточные данные из других систем. Например, процедуры очистки обычно проверяют, содержат ли числовые поля именно числовые данные, насколько указанные даты и числа корректны и осмысленны, и выполняют подобные действия. Процедуры очистки могут также использоваться в тех случаях, когда одно уникальное значение представлено в базе данных различным образом. Например, IBM может обозначаться как IBM Co., International Business Machines, IBM и т. д. Во время очистки различные версии одного и того же элемента данных заменяются на одно значение.

**Загрузка данных**

После того как данные преобразованы для размещения в базе данных назначения, программисты используют процедуры загрузки для записи информации в новую базу данных. Во время этого этапа необходимо определить, будут ли данные переноситься периодически или постоянно. Периодическое пополнение данных выполняется регулярно, например ежедневно, еженедельно или ежемесячно. Если пользователям необходима текущая информация, постоянное обновление позволяет загружать данные в базу данных назначения в оперативном режиме.

**Инструменты манипуляции данными**

Продукты ETL позволяют переносить информацию из одной среды в другую, не создавая специальных программ для открытия файлов, чтения записей, объединения данных и других подобных операций.

**Достоинства**: предлагает унифицированный доступ к различным базам данных; позволяет быстро переносить данные между системами

**Недостатки**

: относительно высокая стоимость