ПЕРМСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ ЭКОНОМИЧЕСКИЙ ФАКУЛЬТЕТ

Дисциплина «Информационные технологии в экономике»

Лекция-25 Технологии интеграции информационных систем

Изучаемые вопросы

- 1. Понятие интеграции. Схемы интеграции («точка-точка», «консолидация данных во внешнем хранилище», «центральный узел обмена данными», «центр консолидации и обмена данными»). Режимы интеграции (реального времени (on-line), приближенный к реальному времени (off-line)). Уровни интеграции.
- 2. Методы интеграции данных (консолидация, федерализация, распространение). Технологии интеграции (Middleware, ETL (Extraction, Transformation, Loading), EII (Enterprise Information Integration), ECM (Enterprise Content Management), EAI (Enterprise Application Integration), EDR (Enterprise Data Replication)).
- 3. Стандарты интеграции (COM/DCOM, CORBA, EDI, JavaRMI, XML). Портальные решения. Системная интеграция.

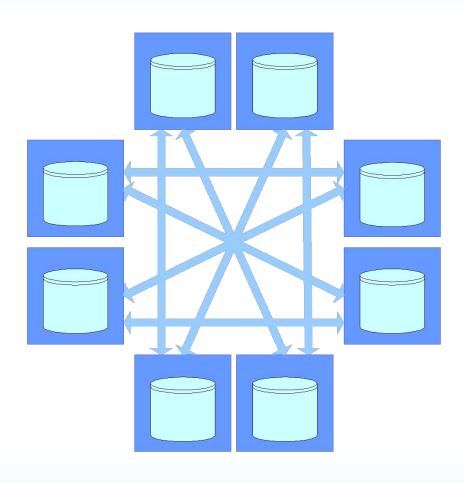
Понятие интеграции

Интеграция - понятие, означающее состояние связанности отдельных дифференцируемых частей и функций системы в целое, а также процесс, ведущий к такому состоянию.

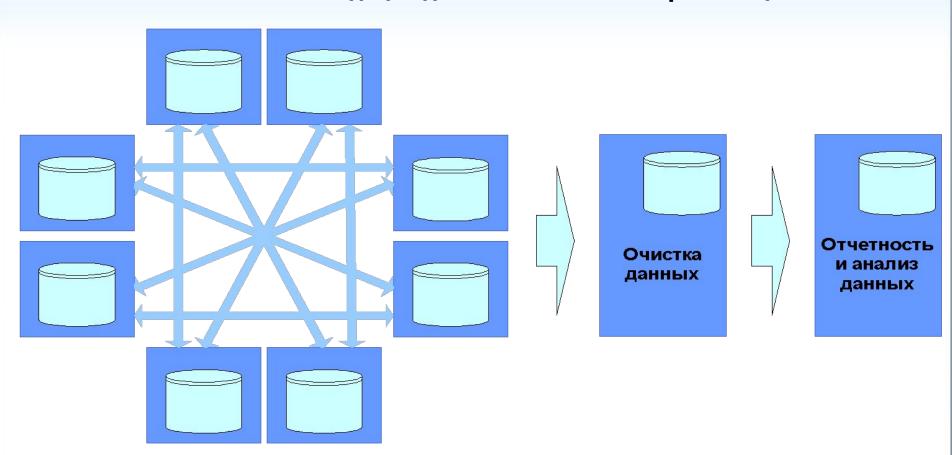
Проблемы обеспечения качества данных:

- отсутствие синхронизации между метаданными (что приводит к разной интерпретации бизнес-сущностей и технологических процессов в разрозненных системах);
- отсутствие синхронизации между справочниками, ведущимися независимо друг от друга в нескольких информационных системах;
- фрагментация, т. е. наличие в каждом из используемых приложений только части информации, необходимой для полноценного анализа и принятия решений;
- противоречивость той части данных, которая ведется параллельно в нескольких информационных системах;
- дублирование записей внутри каждой прикладной системы.

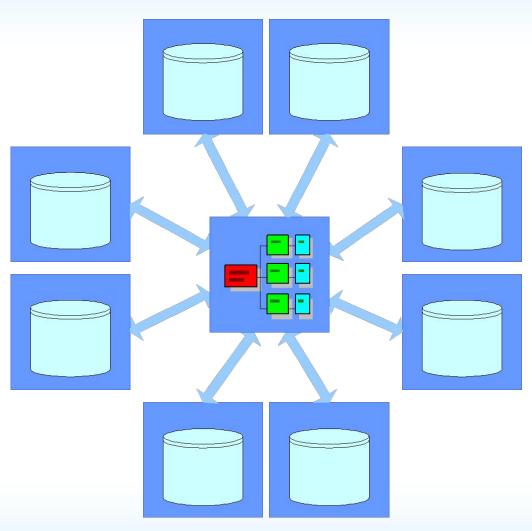
Схема "точка-точка"



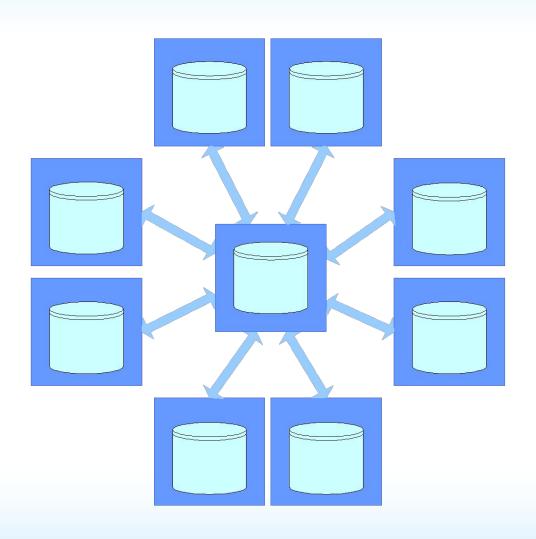
Консолидация данных во внешнем хранилище



Центральный узел обмена данными



Центр консолидации и обмена данными



Режимы интеграции

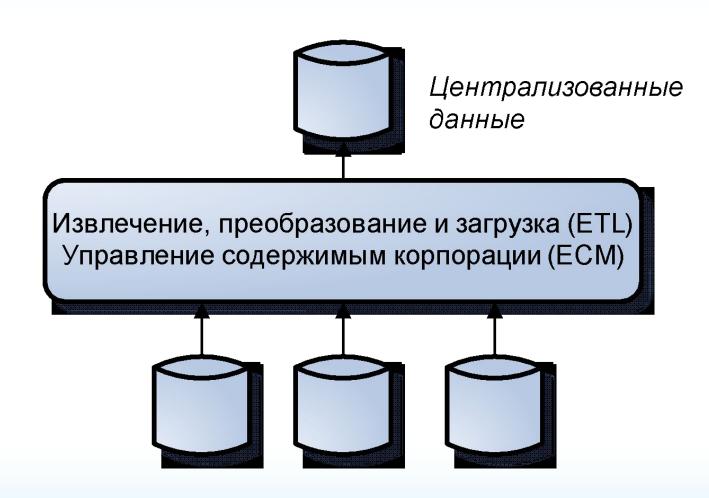
Nº	Режим	Временная задержка	Описание
1	Режим, приближенный к реальному времени	Несколько секунд, минут, часов, дней	В зависимости от потребностей бизнеса отставание между моментом обновления информации в первичных системах и временем, когда данные изменения появляются в конечном месте хранения может составлять существенную величину.
2	Режим реального времени	В установленных пределах (обычно, доли секунды)	Данные не отстают от источника. Такие решения обычно являются очень дорогими.

Уровни интеграции

Nō	Уровень	Описание
1	Интеграция платформ (Platform Integration)	Для интеграции систем - базовой архитектуры, аппаратного и программного обеспечения - необходимо интегрировать разнесенные части гетерогенной сети. Интеграция платформ касается процессов и инструментов, с помощью которых эти системы могут осуществлять безопасный и оптимальный обмен информацией. В результате, данные могут беспрепятственно передаваться по различным приложениям. Например, определение того, как нужно надежно передавать информацию с NT- на UNIX-машину, является чрезвычайно непростой задачей по интеграции всей корпоративной системы.
2	Интеграция данных (Data Integration)	Залогом успешной интеграции приложений и бизнес-процессов является интеграция данных и систем баз данных. Прежде чем приступать к интеграции, необходимо идентифицировать (определить местонахождение) и каталогизировать данные, построить модель данных. По завершении этих трех шагов данные можно совместно использовать/распространять в системах баз данных.
3	Интеграция приложений (Application Integration)	На этом уровне интеграции целью является объединение данных или функций одного приложения с другим, благодаря чему обеспечивается интеграция, близкая к реальному времени. Интеграция приложений используется - и это далеко не полный список - для интеграции В2В, внедрения СRМ-систем, которые интегрированы с корпоративными серверными приложениями, web-интеграции и построения web-сайтов, которые поддерживают многочисленные бизнес системы. Кроме того, может потребоваться проведение специальной интеграции, особенно когда требуется интегрировать существующее приложение с вновь устанавливаемым ERP-приложением.
4	Интеграция бизнес-процессов (Business Process Integration, BPI)	При интеграции бизнес-процессов компания должна определять, реализовывать и управлять процессами обмена корпоративной информацией между различными бизнес-системами. Благодаря этому организация может упростить операции, сократить расходы и улучшить реагирование на запросы клиентов. Элементы включают управление процессами, моделирование процессов и технологический процесс, который охватывает различные задачи, процедуры, архитектуры, требуемую входную и выходную информацию, а также средства, необходимые для каждого шага в бизнес-процессе.
5	Интеграция между организациями (Business-to-Business, B2B)	На этом уровне интеграции целью является обеспечение безопасного, надежного информационного обмена между различными организациями и их информационными системами. Обеспечивается пересылка информации за пределы организации и автоматизируются бизнес-процессы в рамках "расширенных организаций", которые включают поставщиков, партнеров, потребителей продуктов и услуг и т. д.

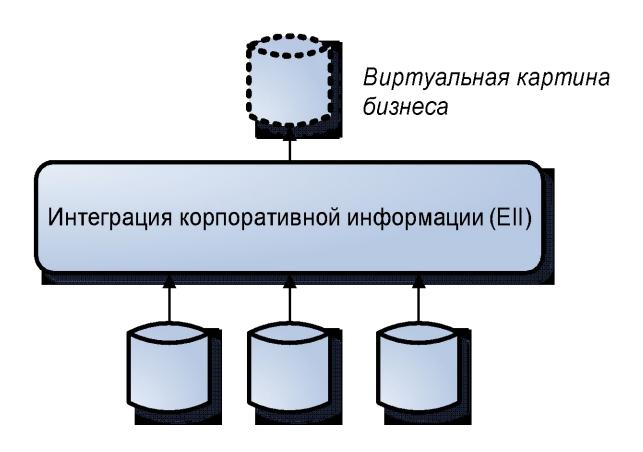
Методы интеграции данных

Консолидация данных



Методы интеграции данных

Федерализация данных



Методы интеграции данных



Nō	Технология интеграции	Описание
1	Интеграция платформ (Middleware)	Обеспечивают связь между процессами в распределенных гетерогенных системах путем обмена сообщениями и удаленного вызова процедур. К основным сервисам таких систем относятся установление соединений между программными объектами, передача сообщений, удаленный вызов процедур/инициализация объектов.
2	Интеграция данных (ETL - Extract, Transformation and Loading Data)	Продукты класса "интеграция данных" (ETL) предназначены для извлечения требуемой информации из разнообразных источников, преобразования и обработки извлеченных данных по заданным правилам и загрузки результатов обработки в целевые базы данных.
3	Интеграция информации (Enterprise Information Integration)	Интеграция в режиме реального времени несопоставимых типов данных из многочисленных систем в виртуально унифицированное, согласованное и точное представление, которое предназначено для последующего изучения и обработки данных.
4	Интеграция приложений (Enterprise Application Integration)	Продукты класса "интеграция приложений" базируются на перечисленных выше средствах и предназначены для организации взаимодействия между готовыми многофункциональными приложениями. Они представляют собой комбинацию универсальных механизмов взаимодействия ПО, готовых средств для интеграции популярных на рынке бизнес-приложений (ERP-, CRM-, SCM-систем), встроенного инструментария для управления взаимодействием на основе бизнес-правил, средств мониторинга и администрирования взаимодействия.
5	Интеграция бизнес-процессов (Business Process Integration)	Продукты класса "интеграция бизнес-процессов" в основном нацелены на решение задач моделирования, организации, мониторинга и администрирования рабочих процессов (work-flow) на базе средств интеграции компонентов и приложений.

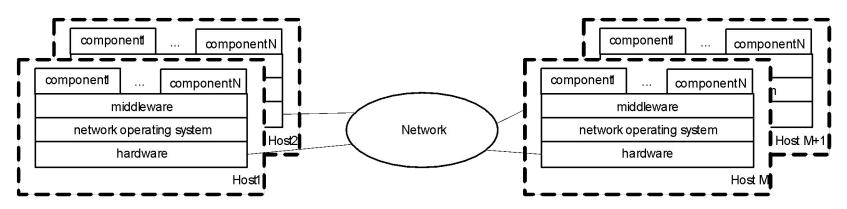
Интеграция платформ

Интеграция платформ базируется на понятии промежуточного программного обеспечения (middleware).

Существующие определения:

Промежуточное программное обеспечение (middleware)

- программное обеспечение производителей компьютеров, которое приспосабливается под конкретные потребности установки (1970г.)
- слой между сетевыми операционными системами и приложениями, который помогает справляться с неоднородностью и распределенностью



Ключевые абстракции промежуточного программного обеспечения:

- вызовы удаленных процедур (Remote Procedure Calls, RPC);
- распределенные транзакции;
- промежуточное программное обеспечение, ориентированное на сообщения (message oriented middleware, MOM).

Интеграция платформ: вызовы удаленных процедур (Remote Procedure Calls, RPC)

Основная идея удаленного вызова процедуры состоит в том, чтобы обеспечить программистам приложений возможность **вызова процедуры**, которая развернута и выполняется на некотором **удаленном** хосте, таким же образом, как если бы вызывалась локальная процедура.

Решаемые задачи:

- система RPC должна отображать сложные типы данных уровня приложения, передаваемые в качестве аргументов вызова и результатов, в транспортное представление, которое может передаваться с использованием основанных на потоках байтов транспортных протоколов. В литературе этот процесс отображения называется маршалингом (marshalling).
- от системы RPC требуется преобразовывать представления данных, поскольку нельзя предполагать, что на клиенте и сервере будут использоваться одни и те же аппаратура, операционная система и язык программирования. Например, для алфавитно-цифровых данных могут использоваться разные кодировки, такие как ASCII или ECBDIC; на мейнфреймах для представления длинных числовых данных может использоваться прямой порядок байт в машинном слове, а в персональных компьютерах используются представления с обратным порядком байт; в разных языках программирования по-разному представляются символьные строки и т.д.

Интеграция платформ: распределенные транзакции

Распределенные транзакции — это последовательности нескольких вызовов удаленных процедур, являющиеся атомарными, согласованными, изолированными и долговечными.

В отличие транзакций, используемых в системах управления базами данных, для поддержки распределенных транзакций требуется промежуточное программное обеспечение, которое иногда называют мониторами обработки транзакций (transaction processing monitors, TPM).

В этих ТРМ реализуется двухфазный протокол фиксации, который позволяет транзакционным менеджерам ресурсов, таких как базы данных или очереди сообщений, сначала решить, в состоянии ли они зафиксировать транзакцию, и только после того, как все участвующие в распределенной транзакции менеджеры ресурсов соглашаются произвести фиксацию, ТРМ запрашивает у них выполнение фиксации.

Распределенные транзакции требуются для создания надежных, отказоустойчивых и параллельно используемых систем.

Интеграция платформ: промежуточное программное обеспечение, ориентированное на сообщения (message oriented middleware, MOM)

МОМ поддерживает **очереди сообщений** (message queue), через которые компоненты распределенной системы могут обмениваться сообщениями **асинхронным** и **надежным** образом.

В очередях сообщений сообщения могут сохраняться, так что они могут пережить сбои промежуточного программного обеспечения и получателя.

В очередях сообщений обычно поддерживаются двухточечные коммуникации и коммуникации на основе «публикации/подписки».

В очередях сообщений также поддерживаются распределенные транзакции, так что сообщения могут ставиться в очередь и изыматься из очереди под транзакционным управлением.

Промежуточное программное обеспечение очередей сообщений, вызовов удаленных процедур и транзакций интегрировано в продуктах серверов приложений. В серверах приложений обеспечивается распределенная компонентная модель, в которой поддерживаются вызовы удаленных операций.

Интеграция данных: Извлечение, преобразование и загрузка данных (Extract, Transformation and Loading Data, ETL)

ETL - это технология, которая преобразует данные (обычно с помощью их пакетной обработки) из операционной среды, включающей гетерогенные технологии, в интегрированные, согласующиеся между собой данные, пригодные для использования в процессе поддержки принятия решений.

ETL - технология, ориентированная на базы данных и хранилища данных (витрину данных, операционный склад данных и т.д.).

Технология ETL оказывается наиболее полезной в тех случаях, когда необходимо создать Хранилище данных, содержащее хорошо документированные и надежные данные для исторического анализа, например, для анализа временных рядов или многомерных запросов.

Эта технология используется для интеграции ключевых справочных данных:

- удаление дублирующихся данных,
- осуществление процессов проверки качества данных и т.п.

Эти инструменты используются для создания отдельных витрин данных, обслуживающих конкретный отдел или бизнес-процесс или предназначенных для каких-либо долгосрочных целей.

Инструменты ETL дают пользователю возможность запустить повторяющиеся процессы для большей слаженности действий и возможности их многократного использования. Такие процессы включают создание точных технических метаданных, поддерживающих общую целостность среды business intelligence (BI).

Интеграция информации (Enterprise Information Integration, EII)

EII — это технология для интеграции в режиме реального времени несопоставимых типов данных из многочисленных источников как внутри, так и за пределами корпорации.

Инструменты EII обеспечивают универсальный уровень доступа к данным и используют технологию поиска информации (pull technology) или возможности работы по запросам.

Технология EII ориентирована на конкретных сотрудников, которые получают информацию через инструментальную панель или отчет

Технология EII лучше всего подходит в тех случаях, когда необходимо создать общий шлюз (gateway) с едиными языком и точкой доступа к несогласованным источникам данных. Такие инструменты предоставляют приложениям и конечным пользователям возможности более гибкого, а также незапланированного доступа к данным, при этом не требуя постоянного использования данных или долговременных целей для получения этого доступа.

Помимо традиционных реляционных баз данных, инструменты EII могут работать с XML- и LDAP-файлами, плоскими файлами и другими нереляционными данными.

Эти инструменты также способны представлять реляционные данные в формате XML или формате webсервисов. Особенно полезны инструменты EII, если есть необходимость добавить к справочным данным Хранилища дополнительные детали, в частности, детальную информацию в реальном времени (например, сопоставление исторических данных с текущей ситуацией).

Интеграция приложений (Enterprise Application Integration)

EAI — это технология, с помощью которой организация добивается централизации и оптимизации интеграции корпоративных приложений, обычно используя те или иные формы технологии оперативной доставки информации (push technology), которая управляется внешними событиями (event-driven);

Технология EAI наиболее функциональна тогда, когда необходимо связать приложения в реальном времени для автоматизации бизнес-процессов.

Второй случай применения EAI - это ситуация, когда необходимо, чтобы изменения, внесенные в одно приложение (обычно это небольшой набор записей), были отражены во всех других. Эта технология очень хорошо справляется с задачей фиксации изменений и их переноса в соответствующие приложения или системы.

Корпоративные системы управления контентом (Enterprise Content Management, ECM)

ЕСМ предполагают построение инфраструктуры управления слабоструктурированной цифровой информацией в виде электронного архива документов, поддерживающих специфику работы с разнородной информацией.

Для реализации этого комплекса задач в области управления корпоративной информацией обычно предлагается интегрированное решение, сочетающее в себе:

- систему ввода и обработки электронных образов бумажных документов (сканирование);
- электронный архив;
- инфраструктуру хранения и передачи информации.

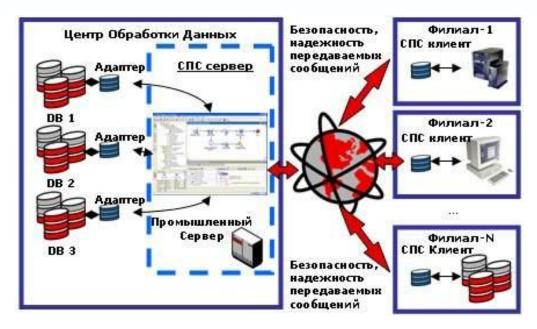
Извлечение, преобразование и загрузка данных в реальном режиме времени (Real time ETL, RT ETL)

RT ETL используется для поддержки интеграции в реальном времени, когда пакетный подход к извлечению операционных данных заменяется на процессы, которые постоянно отслеживают состояние исходных систем, захватывают и преобразуют изменения в данных по мере их возникновения, затем загружают их в Хранилище в режиме, максимально приближенном к реальному времени. Постоянный сбор данных позволяет анализировать прибыль и ценовые элементы в любых временных рамках. Тенденции можно анализировать с любой выбранной периодичностью и без задержки.

Тиражирование данных предприятия (Enterprise Data Replication, EDR)

EDR – процесс формирования и воспроизведения многочисленных копий данных на одном или нескольких узлах. Тиражирование дает ряд преимуществ, таких как высокая производительность, высокая степень готовности, обеспечение сохранности данных на уровне транзакций, гибкая архитектура.

Интеграция бизнес-процессов (Business Process Integration, BPI)



Интеграция бизнес-процессов решает следующие задачи:

- Создание единой транспортной магистрали на основе универсальной системы передачи сообщений между приложениями;
- Гарантированная передача данных между различными приложениями в унифицированном виде в реальном времени;
- Адаптация данных в различных системах путем использования специализированных адаптеров для различных приложений (с возможностью разработки новых адаптеров для любых систем, используемых Заказчиком).

Для обеспечения интеграции данных необходимо выбрать стандартные форматы для данных.

Стандартами интеграции являются те форматы, которые поддерживают использование и распространение информации и бизнес данных, т.е. стандарты являются основой для проведения интеграции корпоративных приложений.

К ним относятся

- COM/DCOM,
- CORBA,
- EDI,
- JavaRMI
- XML

COM/DCOM

СОМ (англ. Component Object Model — Объектная Модель Компонентов; произносится как [ком]) — это технологический стандарт от компании Microsoft, предназначенный для создания программного обеспечения на основе взаимодействующих распределённых компонентов, каждый из которых может использоваться во многих программах одновременно.

Основным понятием, которым оперирует технология СОМ, является СОМ-компонент. Программы, построенные на технологии СОМ, фактически не являются автономными программами, а представляют собой набор взаимодействующих между собой СОМ-компонентов. Каждый компонент имеет уникальный идентификатор (GUID) и может одновременно использоваться многими программами через СОМ-интерфейсы — наборы функций и свойств.

Выпущенная в 1996 году технология **DCOM** (англ. Distributed COM — распределённая COM) основана на технологии RPC и является развитием COM. DCOM позволяет COM-компонентам взаимодействовать друг с другом по сети. Главным конкурентом DCOM является другая известная распределённая технология — CORBA.

CORBA

CORBA (сокр. от англ. Common Object Request Broker Architecture — общая архитектура брокера объектных запросов; по-русски чаще произносится как [ко́рба]) — это технологический стандарт написания распределённых приложений, продвигаемый консорциумом ОМG. Задача СОRBA — осуществить интеграцию изолированных систем, дать возможность программам, написанным на разных языках, работающим на разных узлах сети, взаимодействовать друг с другом так же просто, как если бы они находились в адресном пространстве одного процесса.

CORBA объединяет программный код в объект, содержащий информацию о функциональности объединяемого кода и интерфейсах доступа. Готовые объекты могут вызываться из других программ (или объектов CORBA), расположенных в сети.

СОRBA использует язык описания интерфейсов (OMG IDL) для определения интерфейсов взаимодействия объектов с внешним миром. CORBA описывает правила отображения из IDL в язык, используемый разработчиком CORBA-объекта. Стандартизованы отображения для Ada, C, C++, Lisp, Smalltalk, Java, COBOL, PL/I и Python. Также существуют нестандартные отображения на языки Perl, Visual Basic, Ruby и Tcl, реализованные средствами ORB, написанных для этих языков.

EDI (Electronic Data Interchange)

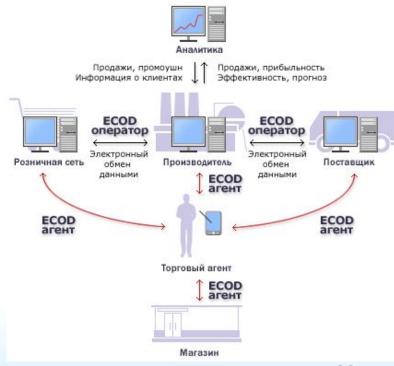
EDI представляет собой сочетание трех направлений: бизнеса, обработки и обмена данными.

Основа EDI — протокол прямого обмена деловыми и коммерческими документами между вычислительными средами в межкорпоративной среде (в отличие от систем документооборота).

Суть EDI заключается в создании стандартизованных документов и представлении их в виде, удобном для компьютерной обработки.

Одними из основных преимуществ EDI являются:

- Экономичность значительное снижение объемов бумаг, подлежащих обработке, приводит к снижению расходов на персонал и административное управление. Освободившиеся сотрудники могут быть переориентированы на решение других задач.
- Оперативность большие объемы коммерческих данных могут быть в течение нескольких минут переданы из одного компьютера в другой. На них будет получен ответ, что позволит немедленно удовлетворить запросы клиентов.
- Точность использование EDI исключает появление ошибок, неизбежных при наборе данных вручную.



Java RMI (Java Remote Method Invocation)

Java RMI — это решение для создания распределенных систем, которое работает для клиентских и серверных приложений, написанных на языке Java.

Распределенные объекты обеспечивают мощную модель для создания объектно-ориентированных сетевых приложений, интерес к которым в последние годы неуклонно растет. В данной модели клиентская часть должна иметь возможность вызывать методы объектов, расположенных на удаленных серверах так, как если бы эти объекты располагались локально. Данная модель может быть реализована несколькими способами. Одним из них является Java Remote Method Invocation (RMI).

XML (eXtensible Markup Language)

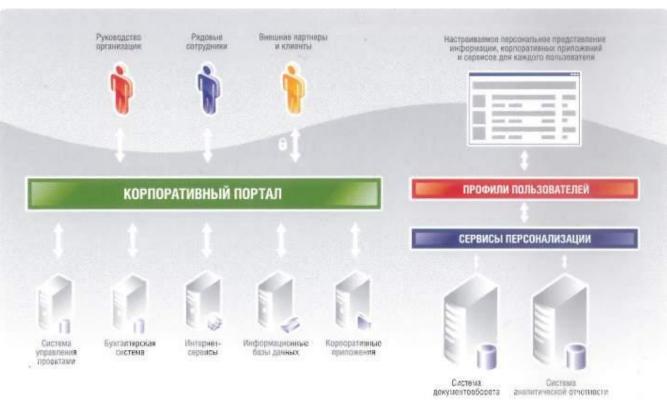
XML (англ. eXtensible Markup Language — расширяемый язык разметки; произносится [экс-эм-эл]) — рекомендованный Консорциумом Всемирной паутины (World Wide Web Consortium, W3C)) язык разметки, фактически представляющий собой свод общих синтаксических правил.

XML предназначен для хранения структурированных данных (взамен существующих файлов баз данных), для обмена информацией между программами, а также для создания на его основе более специализированных языков разметки (например, XHTML), иногда называемых словарями.

Целью создания XML было обеспечение совместимости при передаче структурированных данных между разными системами обработки информации, особенно при передаче таких данных через Интернет.

Портальные решения

- Централизованный доступ и корпоративным информационным системам.
- Удобные средства публикации данных.
- Обеспечение мгновенного доступа к необходимой информации.
- Персонализация доступа и информационным ресурсам.
- Единая в масштабах всей организации система поиска информации.
- Обеспечение совместной работы сотрудников.
- Функциональные сервисы, автоматизирующие рутинные операции
- Средства аналитической обработки данных



Порталы обеспечивают доступ различных групп пользователей: от руководства до рядовых сотрудников ко всему спектру информационных ресурсов организации, с учетом персональных предпочтений по представлению данных.

Любые корпоративные системы могут быть доступны через портал, таким образом формируется единая точка доступа ко всем приложениям. При необходимости может быть организован доступ внешних по отношению к организации пользователей (например клиентов или поставщиков) к порталу с обеспечением необходимого уровня безопасности.

Модульная архитектура портала позволяет решать различные задачи, специфичные для функциональных подразделений организации.

Портальные решения



Системная интеграция

Под системной интеграцией понимается:

- проектирование и разработка некой информационной системы, объединяющей в функционально полное решение программные и аппаратные средства, позволяющие заказчику добиться максимально возможного взаимодействия и эффективности различных бизнес-процессов;
- выбор экономически оправданного, интегрированного информационнотелекоммуникационного решения для реализации конкретных задач заказчика, его комплексная реализация и сопровождение в течение жизненного цикла системы;
- работы по созданию и запуску в эксплуатацию необходимых клиенту систем (ПО или программно аппаратных комплексов), формируемых, как правило, из независимо разработанных компонентов, либо включаемых в состав уже функционирующих систем;
- комплекс работ, предоставляющих заказчику системные (взаимосвязанные и законченные) решения в части технологий, транспортной среды и оборудования, обеспечивающие эффективный бизнес-процесс оператора;
- разработка специализированных решений, включающих поставку аппаратной платформыи разработку ПО, а также интеграцию разработанного решения с другими бизнесприложениями заказчика;
- интеграция различных аппаратных и программных средств в единые подсистемы, а также разработка, производство, монтаж, поддержка и обслуживание программно-аппаратных комплексов, предназначенных для решения определенных заказчиком задач;
- •IT-аутсорсинг передача компанией ряда полномочий и задач (например, поддержание и обслуживание IT структур) сторонней компании-аутсорсеру (обычно это многолетний либо разовый контракт на предоставление услуг с определенным уровнем качества);
- способность обеспечить заказчиков всеми необходимыми механизмами для эффективного управления бизнесом.

Системная интеграция

Система требований к системному интегратору

- Наличие соответствующего ресурса (консультанты, аналитики, архитекторы, инженерно-технический персонал, производственно-техническая база).
- Способность найти в пределах выделенного бюджета оптимальное для данного клиента техническое решение, надежно обеспечивающее существующие потребности и дающее возможность прогнозируемого развития.
- Способность структурировать бизнес-процессы, разрабатывать единую концепцию информационных систем для компаний любого уровня и сложности информационных потоков.
- Способность обеспечить комплексный подход к построению телекоммуникационных и информационных систем различного масштаба, выполняя все виды работ в процессе создания и эксплуатации системы связи:
 - ■технический аудит телекоммуникационных систем с выработкой рекомендаций по дальнейшему развитию;
 - предпроектное обследование, экспертная оценка, сбор и формирование требований к системе, подготовка технического задания;
 - инжиниринг;
 - проектирование;
 - поставка оборудования и средств связи;
 - интеграция оборудования различных производителей;
 - выполнение монтажных и пусконаладочных работ;
 - сдача заказчику системы "под ключ";
 - обучение персонала заказчика;
 - гарантийное и послегарантийное сопровождение.
- Опыт реализации проектов различного масштаба на базе оборудования различных производителей.
- Опыт собственных разработок (как оборудования, так и программного обеспечения).
- Способность создать качественно новые возможности и инструменты для работы, кардинально улучшающие бизнес-процессы и обеспечивающие заказчику новые конкурентные преимущества.
- Способность предложить заказчику на выбор несколько вариантов решения задачи.
- Способность обеспечить высокий уровень качества работ и услуг.
- Оптимально, если интегратор большая компания, имеющая возможность выполнить крупные и сложные проекты и позволить клиенту воспользоваться различными финансовыми схемами;
- Оптимально, если мобильная компания, способная быстро реагировать на возникновение новых проблем заказчика. Современные крупные заказчики территориально распределены, и вопрос мобильности интегратора становится ключевым.

Системная интеграция

Категории системных интеграторов

Широкопрофильный интегратор	Компания, в линейке решений которой есть предложения по самым разным технологическим и отраслевым направлениям и которая может самостоятельно решать широкий спектр проблем заказчиков, включая узкоспециализированные задачи.
Отраслевой интегратор	Компания, специализирующаяся на одном или нескольких клиентах из одной отрасли (например, банковской). Решает специфические проблемы своих клиентов, как правило, очень хорошо разбираясь в этой специфике.
Интегратор узкого профиля	Компания, имеющая очень высокий уровень экспертизы в отдельных (как правило, специфических и сложных) технологиях.
Инженерный интегратор	Компания, занимающаяся физическим созданием локальных сетей и инженерной инфраструктуры (как правило, работает на подряде у более крупных компаний).