

ФЕДЕРАЛЬНОЕ АГЕНТСТВО ПО ОБРАЗОВАНИЮ
АСТРАХАНСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ

РЕФЕРАТ

для сдачи экзамена по специальности 05.13.10 – управление в социальных и
экономических системах

на тему: «Система технико-экономического учёта мультисервисной сети
передачи данных»

Выполнил:

Спандерашвили Дмитрий Викторович

Астраханский государственный университет

Кафедра Информационных систем

Факультет Математики и информационных технологий

аспирант

Астрахань-2005

Содержание

Содержание.....	2
ВВЕДЕНИЕ.....	3
1. ИНФОРМАЦИЯ В ТЕЛЕКОММУНИКАЦИОННОЙ КОМПАНИИ.....	5
2. ПОТОКИ ИНФОРМАЦИИ.....	8
3. СБОР ИНФОРМАЦИИ.....	11
4. ТИПЫ ОТЧЁТНОСТИ.....	13
5. МОДЕЛЬ ХРАНЕНИЯ ИНФОРМАЦИИ В ХРАНИЛИЩЕ.....	21
6. РЕАЛИЗАЦИЯ.....	23
ЗАКЛЮЧЕНИЕ.....	25
ЛИТЕРАТУРА.....	26

ВВЕДЕНИЕ

В течение последних лет бизнес-процессы в телекоммуникационных компаниях претерпели значительные изменения. К наиболее значимым факторам оказывающими влияние на работу компаний–операторов телекоммуникаций, а также на предоставление телекоммуникационных услуг в целом, являются высокий уровень конкуренции, а также постоянное совершенствование технологий, появление новых типов услуг, общий рост потребности со стороны конечных пользователей в обеспечении определённого качества услуг.

Конкуренция со стороны других компаний заставляет обеспечивать более высокий уровень работы с клиентами; осуществлять контроль за группами потенциальных абонентов с целью наиболее быстрого определения возникающего спроса на определённые услуги, и выработку наиболее приемлемого способа предоставления данной услуги в зависимости от возможностей и потребностей конкретной группы потенциальных абонентов. Также важен контроль удовлетворения потребностей существующих абонентов.

Другим фактором оказывающими влияние на работу телекоммуникационной компании является возрастающая динамика развития технологий связи. Появление новых, более перспективных и мощных технологий позволяет как обеспечивать более высокое качество предоставления существующих услуг, так и предлагать клиентам новые услуги. Таким образом первичной в выборе направления развития компании остаётся не новизна технологии или авторитетность фирмы-разработчика, но перспективность технологии с точки зрения повышения качества предоставляемых услуг и возможности предложения абонентам новых, не пересекающихся с существовавшими, типов услуг.

Для получения более полной отдачи от внедрения новых технологий компания-оператор должна предугадывать интересы абонентов ещё на этапе принятия решения о закупке определённого типа оборудования. Не менее

важна задача анализа правильности принятия предыдущих решений: определение популярности услуг, предоставляемых уже закупленным ранее оборудованием,- подобный анализ может позволить избежать повторных ошибок, либо, опираясь на принятое ранее удачное решение, предугадывать популярность появляющихся технологий и проводить своевременное их внедрение.

Для решения задач анализа предоставляемых и прогноза популярности планируемых к предоставлению услуг одно из самых важных мест занимает сбор оперативной и достоверной статистической информации о деятельности компании.

Владение достоверной и удобной для анализа информацией становится одним из важнейших факторов эффективного развития телекоммуникационной компании. Эффективное преобразование данных из различных регистрирующих систем в структурированную информацию, а также обеспечение хранения информации в формате, способном предоставить бизнес-пользователю срезы информации в различных измерениях, в том числе с поддержкой исторических запросов, способно значительно облегчить и ускорить процесс принятия решений о направлении развития компании.

1. ИНФОРМАЦИЯ В ТЕЛЕКОММУНИКАЦИОННОЙ КОМПАНИИ

Информация циркулирующая в телекоммуникационной компании является неоднородной, как по типу описываемых бизнес-процессов так и по характеру обработки накопления и представления.

По проведённому анализу информации данных телекоммуникационной компании, можно предложить следующую классификацию данных по описываемым бизнес-процессам (рис 1):

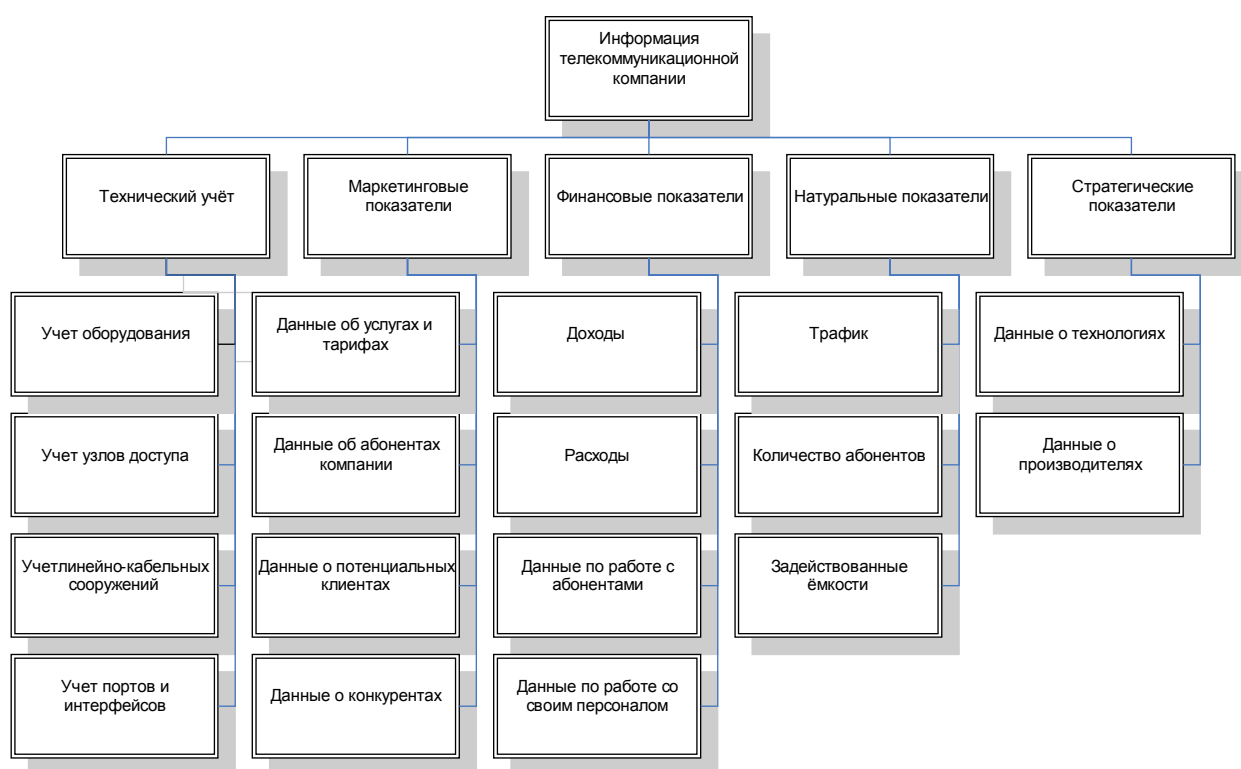


Рисунок 1. Классификация данных по описываемым бизнес-процессам.

Информация описывающая различные виды деятельности телекоммуникационной компании имеет определённую периодичность фиксации, которая определяется скоростью изменения данных, а также необходимой степенью актуальности при её анализе.

Данные по типам периодичности их фиксации можно разделить на следующие типы, отражённые на диаграмме:

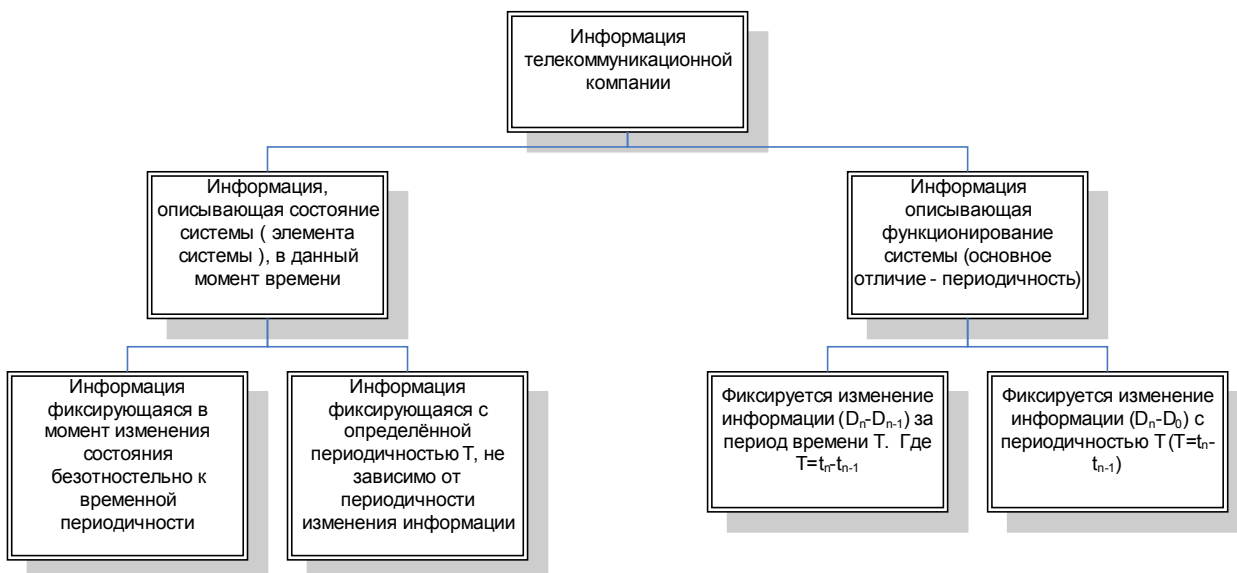


Рисунок 2. Классификация данных по типам периодичности её фиксации.

Итак, на диаграмме показано 4 основных типа периодичности фиксации информации в системе:

- Фиксация информации без какой-либо определённой периодичности в зависимости от изменения состояния фиксируемого параметра, как только данное изменение происходит новое значение параметра фиксируется как актуальное. Данный тип периодичности можно охарактеризовать нерегламентированностью скорости изменения информации.
- Фиксация состояния параметра наблюдаемого объекта с определённой периодичностью T . Данная периодичность не зависит от изменения состояния параметра объекта. Фиксация происходит независимо от того произошло изменение или нет. Особенность данного типа периодичности заключается в возможности наличия неоправданных ресурсных расходов на

фиксацию, если скорость фиксации значений параметра немного выше максимально возможной скорости его изменения; а так же в возможности наличия потерь в актуальности информации если скорость фиксации ниже чем максимально возможная скорость изменения состояния параметра.

- Третий тип периодичности относится к параметрам, описывающим функционирование системы, а не её состояние. Данный тип периодичности относится к фиксации объёмов определённых данных изменяющихся в период времени $t_{n1}-t_{n2}$, притом фиксируется изменение именно между этими двумя моментами, безотносительно к предыдущим состояниям параметра. Значение параметра к данному случаю может быть положительным (при росте объёмов), отрицательным (при снижении объёмов), либо нулевым (при неизменности объёмов).
- Четвёртый тип периодичности является обобщением третьего типа. В данном типе фиксация информации происходит так же с интервалом времени T , однако фиксируются значения данных не в интервале $t_{n1}-t_{n2}$, а в интервале $t_{n1}-t_{n0}$, где t_{n0} – время начала отсчёта состояния данного параметра. Между третьим и четвёртым типом периодичности существует связь:

$$\Delta D = (D_{n_k} - D_{n_0}) = \sum_{i=0}^k (d_i - d_{i-1}) \quad \text{Где } D - \text{данные заполняемые с}$$

периодичностью четвёртого типа, а d данные заполняемые с периодичностью третьего типа.

Важной задачей в аспекте периодичности фиксации данных является учёт этой периодичности при проектировании структуры БД с поддержкой исторических запросов.

2. ПОТОКИ ИНФОРМАЦИИ

Ввиду того, что телекоммуникационные компании в большинстве своём покрывают достаточно обширные территории, а также по причине большого, и всё растущего, многообразия предоставляемых услуг, сбор статистической информации становится достаточно громоздким и запутанным процессом. Необходимо анализировать всё большее количество показателей. Параллельно с ростом количества показателей происходит увеличение количества контрольных точек на которых данные показатели снимаются с объекта наблюдения. Сложность новых технологий приводит к тому, что мониторинг состояния оборудования или услуг, основанных на данных технологиях, производится несколькими подразделениями компании.

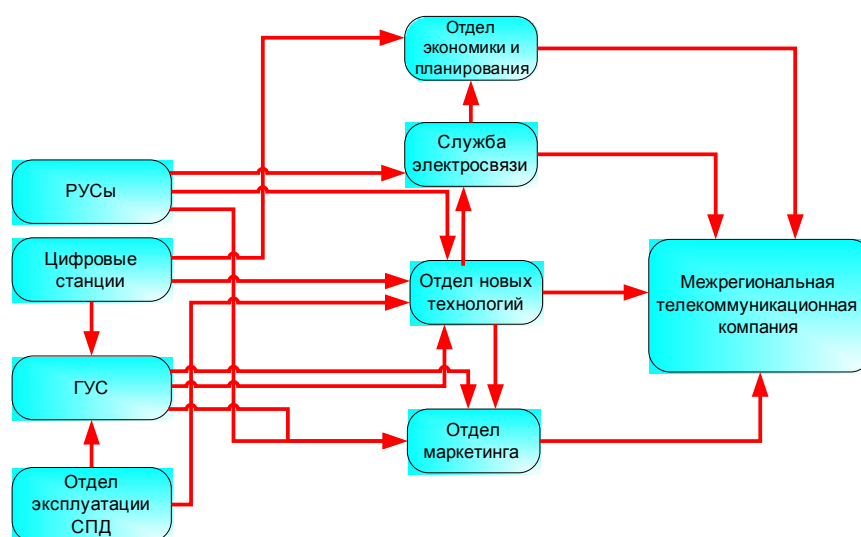


Рисунок 3. диаграмма потоков данных обмена статистической информацией до внедрения системы сбора статистической отчетности

Довольно часто возникает ситуация дублирования и неоднозначности статистической информации. На рисунке 3 приведена диаграмма потоков статистической информации в телекоммуникационной компании, полученная при анализе процесса сбора отчетности типичной региональной телекоммуникационной компании (чтобы чрезмерно не усложнять диаграмму некоторые потоки опущены).

Данная диаграмма потоков информации, а также тот факт, что передача информации происходит неоднородным образом (одни и те же

статистические отчёты могут передаваться и посредством электронных писем, и в виде факсимильных сообщений) позволяют сделать вывод о несовершенстве данной системы сбора статистической отчётности. Информация на данной диаграмме передаётся непосредственно между отделами организации, так что нет возможности организовать единое хранилище информации; каждый отдел взаимодействует с несколькими другими, при этом формы отчётов, предоставляемые одним отделом другому, зачастую представляют собой информацию, которую не требуется обрабатывать, а лишь передавать дальше, следующему звену цепочки обмена статистической информацией. Отсутствие центрального хранилища порождает необходимость ведения каждым отделом своего собственного архива, – очевидна избыточность трудовых затрат, а также возможность появления неоднозначной информации. Итак, недостатками применяемой схемы обмена информацией являются:

- Необходимость непосредственного взаимодействия каждого отдела с множеством других.
- Опасность появления неоднозначной информации.
- Чрезмерная избыточность трудовых затрат при ведении локальных архивов.

Использование единого хранилища данных не просто автоматизирует существующую технологию обмена информацией, но и вносит изменения в бизнес-процессы существовавшие на предприятии. Диаграмма, описывающая потоки данных при использовании системы сбора статистической информации, существенно отличается от рассмотренной ранее

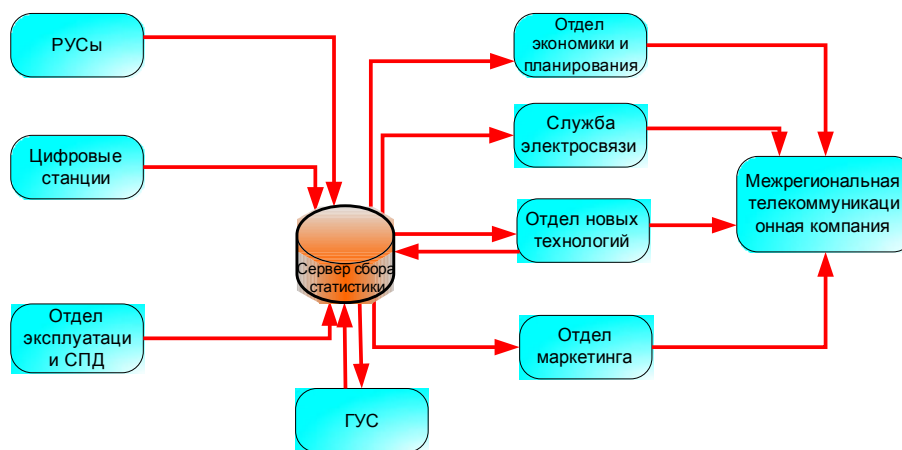


Рисунок 4. диаграмма потоков данных обмена статистической информацией после внедрения системы сбора статистической отчетности

Из диаграммы (Рис 4.) видно насколько упрощается обмен информацией. Устраняются все недостатки существовавшей модели обмена информацией. Централизованное хранилище информации устраняет дублирование и неоднозначность потоков данных. Применение системы сбора статистики устраняет необходимость ведения локальных архивов статистической информации, разрушает информационные связи, существовавшие между отделами и предоставляет каждому отделу единый интерфейс обмена информацией. Итак, достоинствами единой системы обмена информацией являются:

- Единый интерфейс обмена информацией.
- Устранение возможности появления неоднозначной информации.
- Устранение избыточности трудовых затрат.

3. СБОР ИНФОРМАЦИИ

Сбор и накопление информации на предприятии осуществляется различными способами, однако для информации описывающей различные аспекты функционирования компании можно отметить, что каждое подразделение, занимающееся контролем определённого вида деятельности, осуществляет сбор и накопление данных по данному виду деятельности. Данные могут собираться и фиксироваться вручную либо средствами автоматического мониторинга, OLTP системами. Данные накапливающиеся в подразделениях компании необходимы для оперативного управления контролируемой данным подразделением деятельностью. Однако для оперирования информацией в масштабах предприятия данные фильтруются и передаются в центральное хранилище данных предприятия.

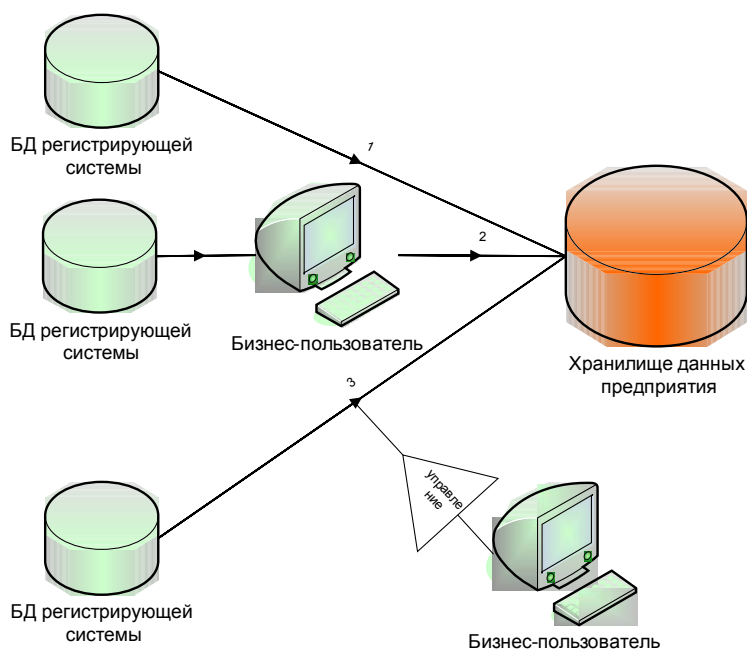


Рисунок 5 Способы сбора информации в хранилище данных предприятия

Способы передачи информации из локальных хранилищ подразделений предприятия в центральное хранилище можно разделить на три типа (рис 5):

- Ручной перевод информации. Приданном типе ввода информации, данные из локальных хранилищ, фильтруются сотрудником подразделений, отбираются наиболее значимые для предприятия и переносятся в

центральное хранилище данных предприятия. Данный тип переноса информации применим в случае специфичности формат локальной базы данных, в случае если при отборе данных используются эвристические, не формализуемые методы, а также в случае если необходимо внести информацию, локальное хранение которой на данный момент не автоматизировано.

- Автоматический ввод информации. Данный тип переноса информации наиболее удобен, так как отбор перенос и занесение информации осуществляется с заранее запланированной периодичностью автоматически, без участия человека
- Полуавтоматический ввод. Данный тип характерен тем что при переносе информации кокой либо процесс происходит под управлением человека. Например человек может инициировать перенос информации, принимать участие в фильтрации информации либо в её размещении в центральном хранилище данных предприятия.

В реальных условия применяется все три типа переноса информации, однако необходимо стремиться, чтоб как можно больше данных переносились в центральное хранилище автоматически.

4. ТИПЫ ОТЧЁТНОСТИ

Известно несколько различных архитектур корпоративной отчетности, начиная с создания отчетов напрямую из учетных систем и заканчивая комплексной средой хранилища, где данные очищаются, затем приводятся к стандартному формату и загружаются в репозиторий.

Первый тип сбора отчётности – **Отчетность в учетных системах** (рис 6). Для этого типа характерно применение инструментария, где предусмотрено собственное подключение к базе данных OLTP или используется общий протокол доступа к данным, например: ODBC, JDBC, OLE DB и т.п.

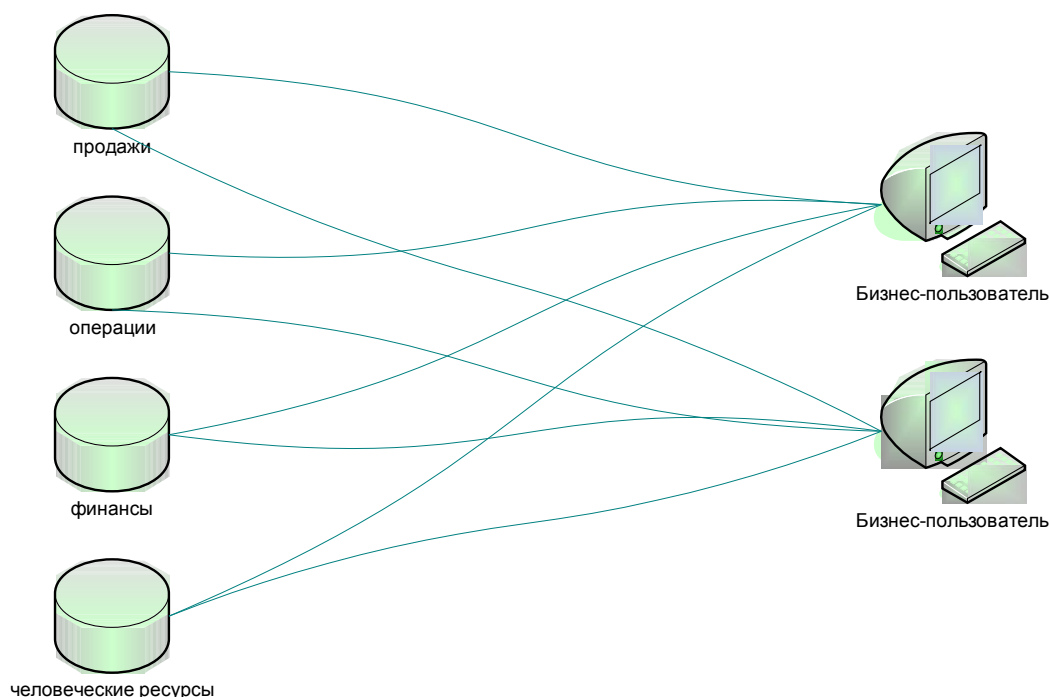


Рисунок 6. отчётность в учётных системах

Основное преимущество этого подхода - скорость получения информации. Необходимо только дождаться выполнения операции, введенной до начала запроса. Такая архитектура удобна для создания

срочных отчетов, т. е. в тех случаях, когда необходима информация об операциях, выполненных в течение последних 24 часов. Такие отчеты чаще всего дают список операций в порядке их выполнения в системе.

Данные, полученные из учетной системы - это необработанные сведения, введенные операторами, либо собранные автоматизированной системой. Их часто называют "неочищенными" (dirty), имея ввиду, что они недостаточно четко представляют необходимую информацию.

Еще одна сложность, связанная с созданием отчетов напрямую из исходных систем, заключается в том, что организация данных не всегда наглядна для бизнес-пользователей. Модель данных спроектирована для оптимизации ввода данных, а не для доступа. Следовательно, части специалистов придется отвлекаться на создание отчетов, удобных для восприятия остальными сотрудниками.

Также следует отметить, что прямые запросы к учетной системе идут в ущерб скорости выполнения транзакций.

Второй тип сбора отчетности – **Отчётность по дублированной базе учётной системы (Рис 7)**. Альтернатива описанному выше подходу - использование "автономной" (offline) базы данных (data store), которая представляет собой дубликат оперативной базы. Дублированная база обновляется по мере необходимости; в ней применяются те же средства отчетности. Запросы пользователей передаются в автономную базу данных, которая может быть реализована по разному, например в виде плоских файлов (flat files) или копий таблиц исходной базы.

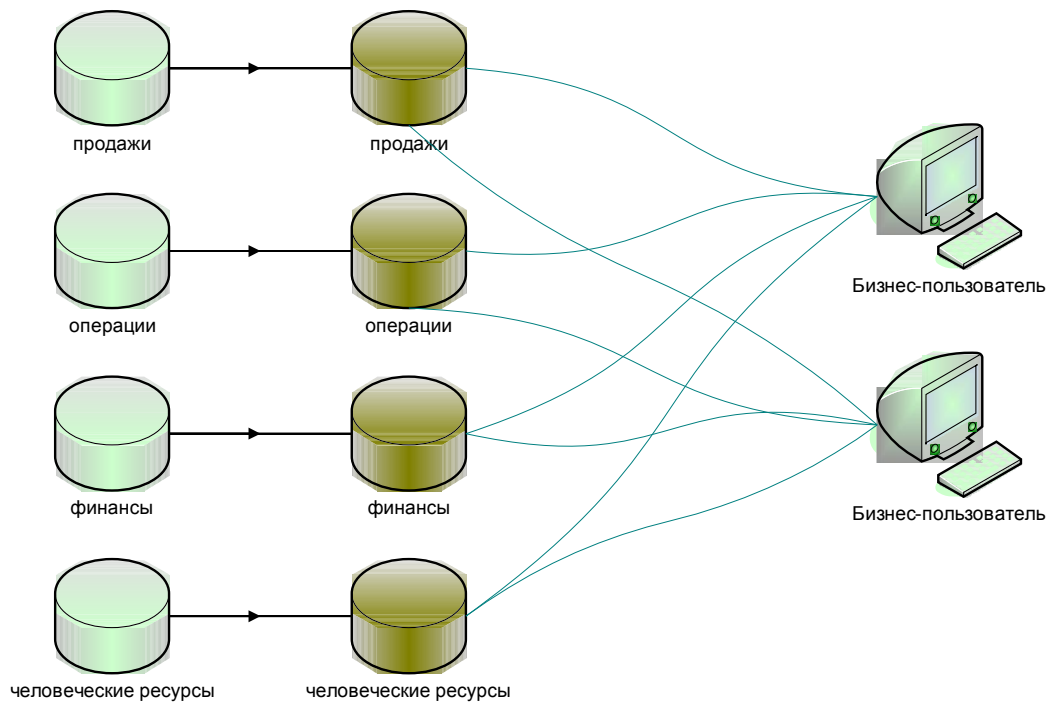


Рисунок 7. отчёт по дублированной базе учётной ситемы

Преимущество этого подхода в том, что выполнение запросов пользователей происходит в другой базе данных, а значит, обработка транзакций не прерывается. При этом системные ресурсы OLTP-сервера используются для быстрой обработки транзакций, а автономная база - только для задач отчетности. Однако многие из проблем, свойственных первому подходу, так и остаются не решенными, поскольку данные по-прежнему не "очищены", их организация не понятна пользователю, а информация за прошлые периоды, как правило, не доступна.

Третий тип сбора отчетности – **Отчётность с использованием витрин данных (Рис 8)**. Данный тип заключается в передаче данных транзакций в некоторые системы, специально разработанные для эффективного создания отчетов. Такие системы называются витринами данных (data mart), а

соответствующая технология оптимизации данных для отчетности - многомерным моделированием данных (dimensional data modeling).

Разработчики представляют данные уже на языке бизнес-терминологии, а не в виде формальных OLTP-структур. Тогда любой средний пользователь, не являющийся специалистом по учетным системам, сможет без особой подготовки получить из витрины данных относительно сложную информацию. Такой метод анализа получил название оперативной аналитической обработки, или OLAP (OnLine Analytical Processing). Он является мощным инструментом просмотра любых параметров коммерческой деятельности компании. Эти параметры также известны как измерения, в них заносится информация о клиентах, поставщиках, типе продукции, времени и географии продаж и т.д.

Чтобы перенести данные из учетной системы в витрину данных, необходимо задействовать дополнительные механизмы. Хотя эти методы могут замедлить доступ к данным, они оптимизируют данные, упрощая создание отчетов. Однако данные, перемещенные в витрину, зачастую остаются по-прежнему не обработанными и не очищенными.

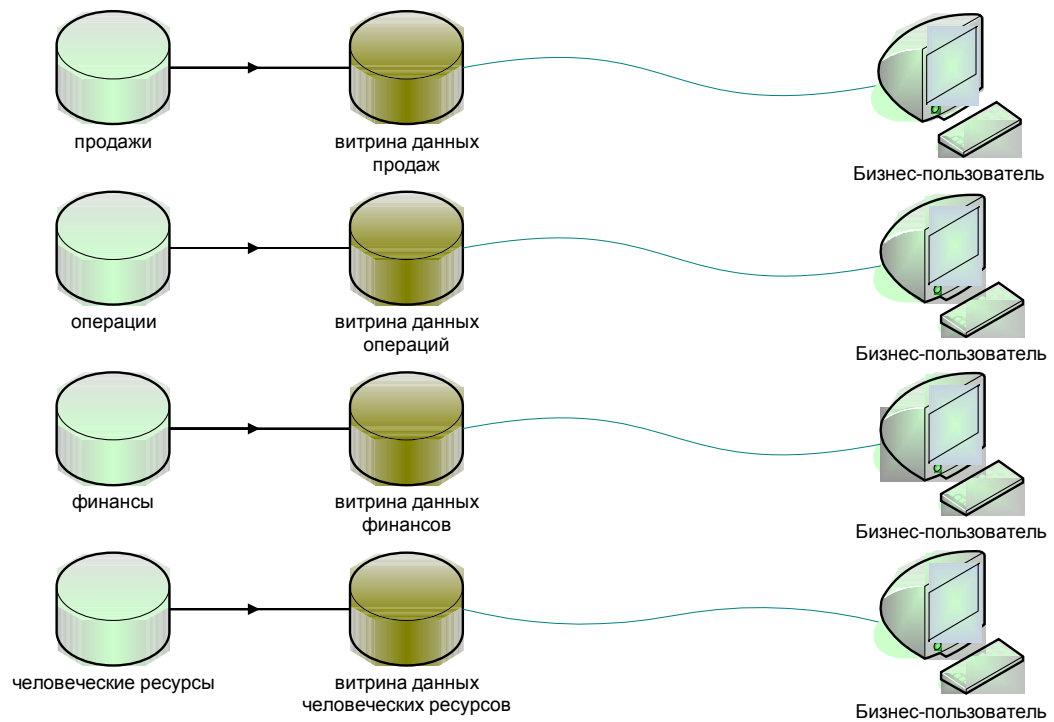


Рисунок 8. Отчётность с использованием витрин данных

Следует отметить, что учетная система и витрина данных - это две самостоятельные системы, каждая из которых использует свои ресурсы только для одной цели, а поэтому работает максимально эффективно. Еще одно преимущество этой технологии: в витринах можно хранить транзакции за длительные периоды, что позволяет анализировать различные временные тенденции, тогда как в OLTP-системах это невозможно. Основной недостаток этого подхода - ограниченность масштабируемости витрин данных. Для сохранения наглядности представления витрина обычно ориентирована на какую-нибудь одну предметную область (например: продажи, управление запасами, финансовый анализ или производство). Крупная организация может иметь до 20 витрин для каждого отдела. Причем

использование этой технологии не решает проблемы консолидации и очистки данных. Сложности здесь могут возникнуть при попытке, например, объединить данные о доходах из витрины продаж с данными по численности сотрудников из витрины человеческих ресурсов. Если поместить все эти сведения в одну витрину, то пользователи будут перегружены слишком большим количеством информации. Для таких целей лучше подходит Хранилище данных предприятия.

Четвёртый тип сбора отчётности – **Отчётность по хранилищу данных предприятия (Рис 9)**. Хранилище данных предприятия (Enterprise Data Warehouse - EDW) предназначено для объединения данных из различных учетных систем, обеспечивает консолидированные и очищенные данные для ряда витрин данных. Ключевым моментом этой технологии является применение бизнес-правил к исходным данным. Бизнес-правила определяют методы консолидации, стандартизацию кодов, очистку данных и отслеживание транзакций за прошедшие периоды. Таким образом, каждая из систем (учетная, Хранилище и витрина данных) будет выполнять те цели, для которых она была разработана.

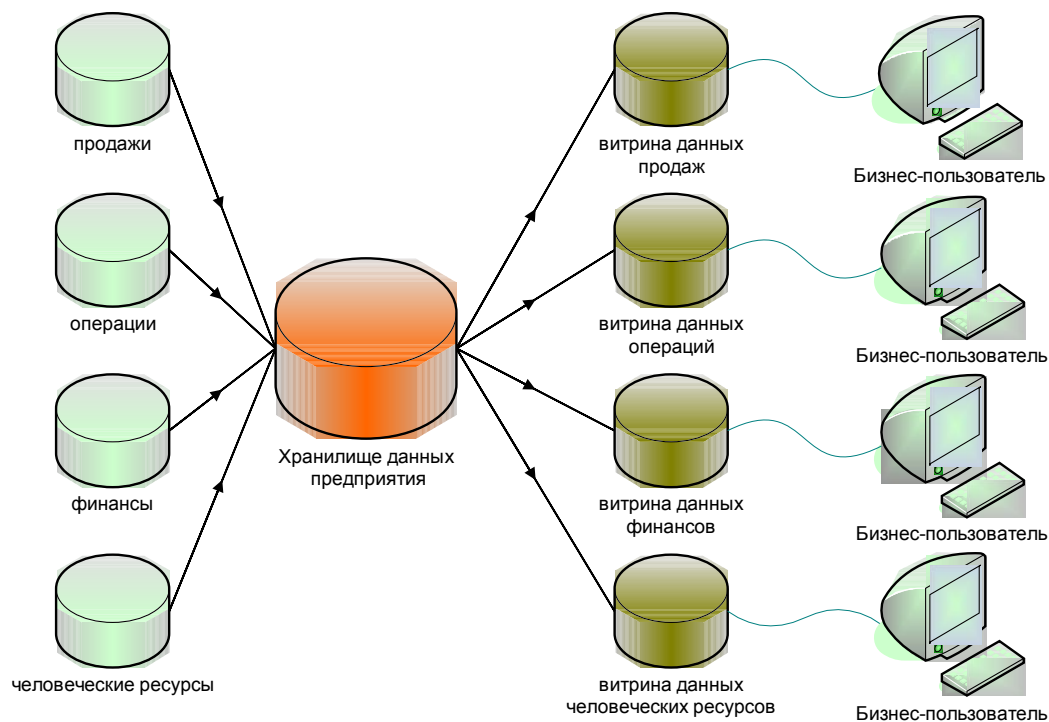


Рисунок 9. Отчётность по хранилищу данных предприятия

Часто данные передаются в Хранилище в стандартном формате. К каждому типу данных, загружаемых в хранилище, применяется специальный формат файла или XML-схема. Например, при пересылке информации о продажах в EDW, выполняется ее консолидация по всем исходным системам. Каждый клиент должен быть занесен в основной список, независимо от того, во скольких отделах он зарегистрирован.

После консолидации к данным применяются правила стандартизации и очистки. Таким образом, у пользователей появляется согласованный доступ к данным, не зависящий от источника. Записи, которые не отвечают требованиям, предъявляемым к качеству данных, получают статус "ожидания" ("on hold"), пока не будет выполнено некоторое действие,

которое устранил создавшуюся неопределенность. В соответствии с бизнес-правилами, таким данным можно присвоить также значение "по умолчанию" (default value).

После очистки и консолидации данные хранятся в реляционной СУБД. Часто используется та же технология, что и в учетных системах (например, продукт фирмы Oracle, IBM, SQL Server и т.п.). Данные хранятся в реляционном формате, позволяя эффективно использовать историю транзакций и время от времени менять дизайн. Помимо истории транзакций Хранилище данных также содержит изменения, вносимые в структуру бизнес подразделений. Так, продавцы часто переходят в другие отделы, а заведующие отделов прикрепляются к другим подразделениям. Но запись о конкретной продаже не подгоняется под изменившуюся структуру организации, так как каждая транзакция должна отражать ту реальную ситуацию, которая складывалась на момент ее выполнения. Таким образом, в Хранилище поддерживаются так называемые "медленно изменяющиеся измерения" (slowly changing dimensions).

Хранилище может служить единым источником регулярно обновляемой информации для витрин - на сегодняшний день это наиболее предпочтительных инструментов пользовательского доступа к данным.

5. МОДЕЛЬ ХРАНЕНИЯ ИНФОРМАЦИИ В ХРАНИЛИЩЕ

Хранилища строятся на основе многомерной модели данных, подразумевающей выделение отдельных измерений (время, география, клиент, счет) и фактов (объем продаж, доход, количество товара) с их анализом по выбранным измерениям. Многомерная модель данных физически может быть реализована как в многомерных, так и в реляционных СУБД. В последнем случае она выполняется по схеме «звезда» или «снежинка». Данные схемы предполагают выделение таблиц фактов и таблиц измерений. Каждая таблица фактов содержит детальные данные и внешние ключи на таблицы измерений.

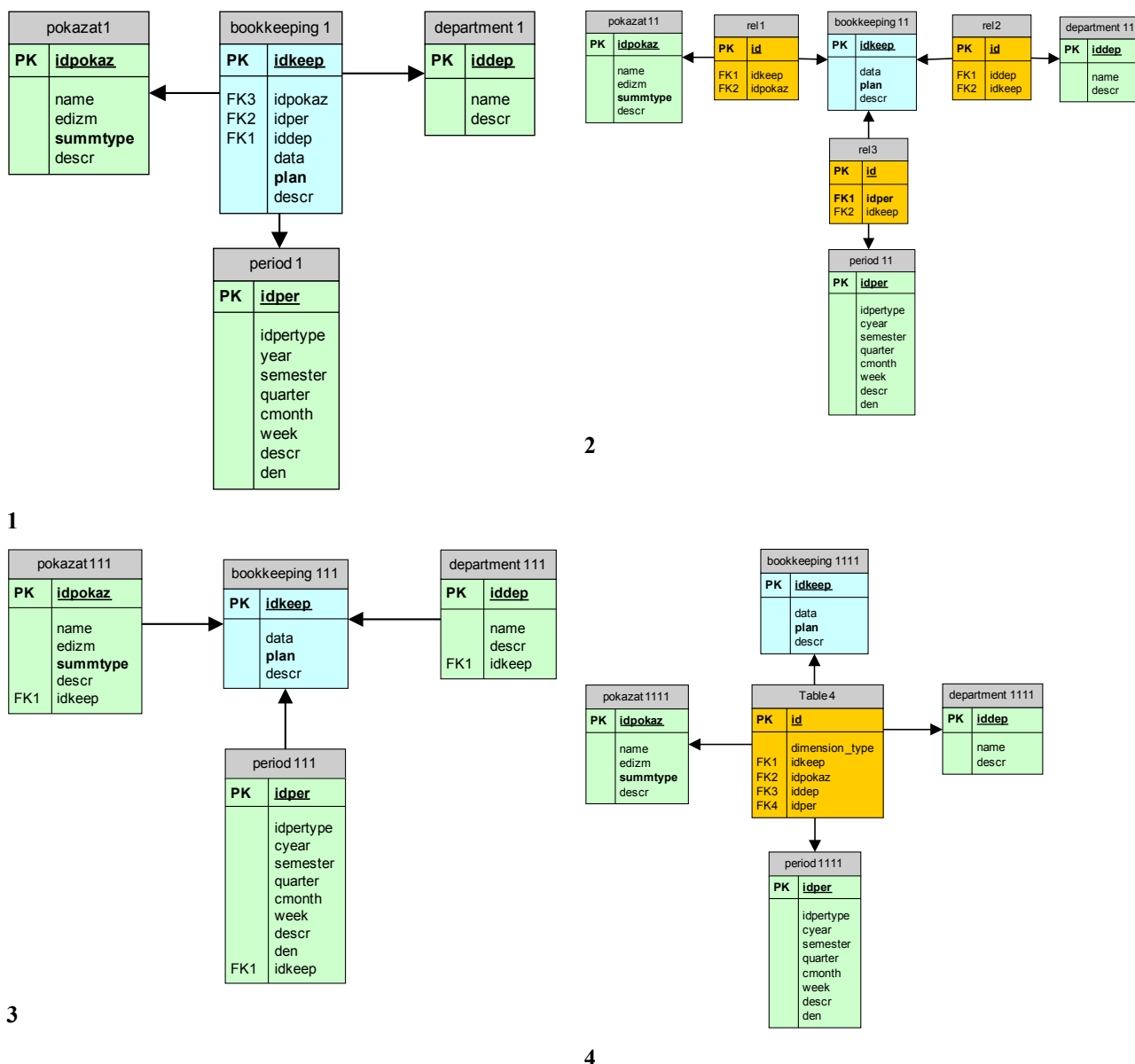


Рис10. Виды реализации модели хранения данных "снежинка"

В результате проведённого исследования выявлено 4 типа реализации модели хранения данных "снежинка" (Рис 10):

- Первый тип характерен хранением ключей измерений в таблице фактов. Данная реализация является наиболее просто, однако характеризуется статичностью и трудностью добавления новых измерений в модель. Однако добавление нового факта вызывает добавление всего одного факта в таблицу фактов.
- Второй тип характерен выделением N таблиц для хранения связей (N - количество измерений) является наиболее приспособленным для добавления новых измерений в модель. Однако при добавлении факта происходит $1+N$ добавлений в записей.
- Третий тип характеризуется хранением ключей фактов в таблицах измерений. Такой подход для добавления измерений. При добавлении факта происходит $1+N$ добавлений в записей (N - количество измерений). Недостатком является избыточное использование памяти при достаточно длинных записях в таблицах измерений.
- Четвёртый тип характерен выделением отдельной таблицы для хранения связей таблицы фактов и измерений, является удобным для добавления новых измерений и при добавлении факта происходит добавлении всего 2-х записей в таблицу связей.

6. РЕАЛИЗАЦИЯ

В процессе исследования вопросов сбора и хранения информации в телекоммуникационной компании в региональном филиале ОАО "ЮТК" "Связьинформ" Астраханской области была разработана и введена в промышленную эксплуатацию система сбора статистической отчётности. Основным принципом архитектуры данной системы является возможность расширения. Вторым принципом – наличие централизованного хранилища информации. Данное хранилище представляет собой полнофункциональную реляционную СУБД с возможностью построения стандартных SQL запросов, возможностью "горячего" резервного копирования и конвертирования в формат наиболее популярных СУБД. Возможность конвертирования гарантирует сохранность информации в случае принятия решения о переходе на другую программно-аппаратную платформу. В результате проведённого анализа рынка, в качестве СУБД, удовлетворяющей данным требованиям, а также предоставляющей оптимальное соотношение цена-качество (так как данная СУБД является бесплатной), была выбрана кроссплатформенная СУБД Firebird. Движок базы данных Firebird, был разработан на основе исходного кода базы данных Interbase 1.0 выпущенной фирмой Borland под открытой лицензией 25 июля 2000 года. За счет усовершенствования программного кода (переноса с языка C на язык C++ и его серьёзной переработки) и улучшенного управления памятью, производительность при выполнении определённых задач возросла на 30–60 процентов. Данная СУБД имеет версии для наиболее популярных операционных систем, что в сочетании с возможностью осуществлять доступ к данным как с помощью ODBC так и JDBC позволяет говорить о её кроссплатформенности.

Доступ клиентов к базе данных осуществляется через web интерфейс, что даёт все преимущества распределённых web приложений:

- Возможность доступа из любой точки земного шара (что важно, например, при необходимости продемонстрировать

показатели работы организации на различного рода семинарах и конференциях)

- Отсутствие необходимости в установке какого либо специфического клиентского ПО, – достаточно стандартного web браузера.

В качестве программной "прослойки" между клиентом и базой данных выбран web сервер Tomcat. Серверное программное обеспечение построено по технологии сервлетов на языке программирования Java. Доступ к данным осуществляется с использованием технологии JDBC, которая реализована под самые разные операционные системы и является более быстрой и надёжной чем технология ODBC. Предложенная архитектура является полностью кроссплатформенной что позволяет, при необходимости с лёгкостью мигрировать с одной программно-аппаратной платформы на другую.

Данная архитектура представлена на нижеприведённой схеме (Рис 11.)

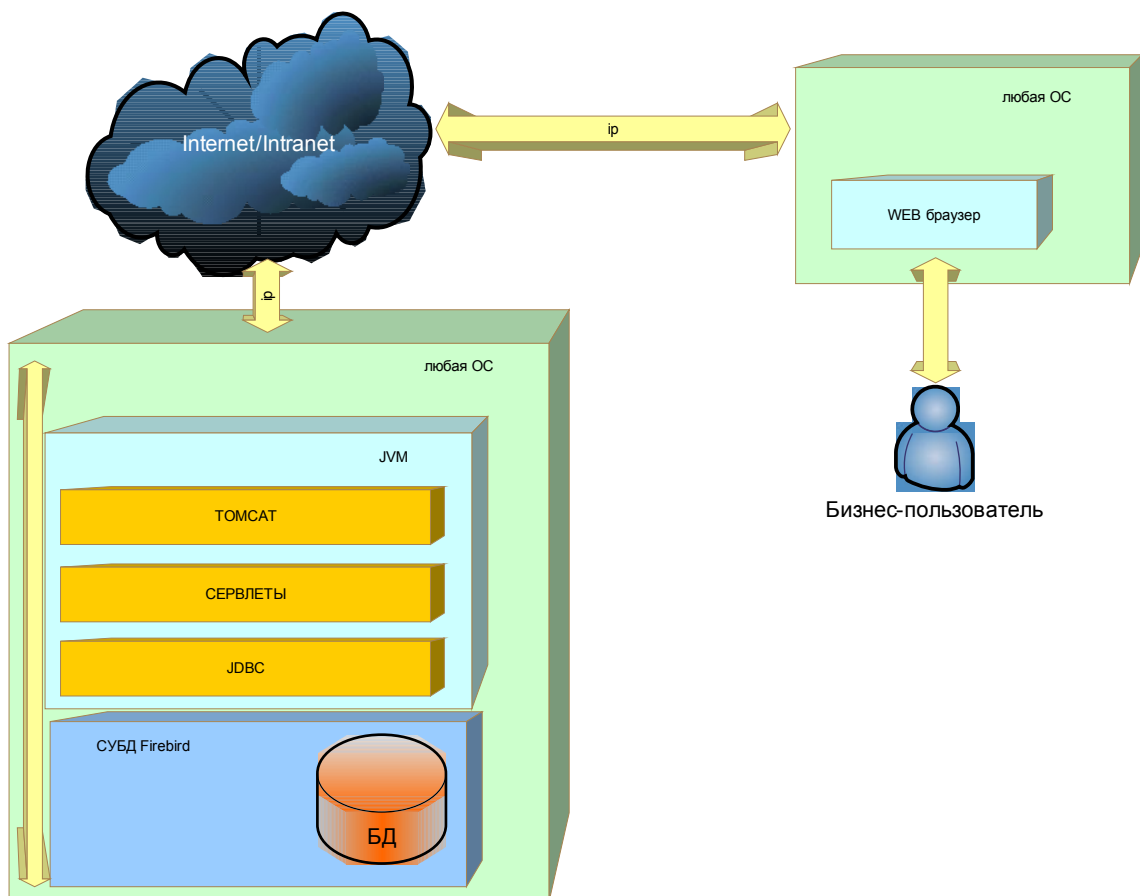


Рисунок 11. Архитектура системы сбора статистической отчётности.

ЗАКЛЮЧЕНИЕ

В наше время более чем когда-либо требуются качественные и гибкие средства отчетности. Компаниям необходима архитектура, позволяющая эффективно использовать данные для принятия сложных решений. И, несмотря на широкий выбор средств, наиболее надежным способом достижения этой цели является реализация Хранилища данных. Расширяющиеся компании используют все новые и новые источники данных, а поэтому именно основательная инфраструктура Хранилища - это оптимальное средство для поддержки принятия решений.

ЛИТЕРАТУРА

1. Э. Гамма, Р.Хелм, Р. Джонсон, Дж. Влиссидес. Приёмы объектно ориентированного проектирования:паттерны проектирования – Спб: Питер, 2004.- 366 с.:ил.
2. Marty Hall. Core Servlets And JavaServer Pages – SUN Microsystems [PDF]
3. Marty Hall. More Servlets And JavaServer Pages – SUN Microsystems [PDF]
4. Java™ Servlet Specification, v2.2 – SUN Microsystems [PDF]
5. ArgoUML Manual – Tigris. [PDF]
6. Владимир Катафеев. Data Warehouse & OLAP Хранилища данных - место OLAP в информационной структуре Предприятия – Компания "Key Century" [PDF]
7. Эндрю Гро (Andrew Grohe). Архитектуры отчетности. – <http://citforum.ru>