

# Тема: Разработка структуры хранилища данных для интеллектуальной обработки учетных данных биллинговой системы

Все права защищены

Ташлыков С.Г.

## Введение

За последние 10 лет строительство хранилищ данных, несмотря на сложность вопроса, стало важным компонентом современных систем поддержки принятия решений в крупных организациях. Само же Хранилище данных представляет собой предметно-ориентированную информационную базу данных, строится такое хранилище на базе систем управления базами данных и систем поддержки принятия решений.

Одна из основных проблем в строительстве хранилища данных заключается в выявлении и рассмотрении информационной потребности потенциальных пользователей. Часто они не в состоянии сформулировать свои требования. Решают эту проблему, обычно используемая модель бизнес-процессов.

В данной статье объясняется связь между бизнес-процессами модели данных и хранилищем данных на примере: упрощенной модели бизнес-процесса компании предоставляющей услуги телематики.

## 2. Основные методы построения хранилища данных (плюсы и минусы)

**Объектно-ориентированное построение хранилища:** способ заключается в опросе будущих групп пользователей хранилища, для получения полного представления о бизнесе компании. Основная проблема связана с тем, что потенциальные пользователи, чаще всего не могут четко сформулировать свои потребности в информации или вовсе их не представляют. Не каждый пользователь представляет себе требования к их представлению и соответственно не знаком с OLAP инструментами.

**Операционно-ориентированное построение хранилища:** стратегия построения хранилища заключается в анализе данных основных систем-источников информации и определения операций проводимых над ними.

Основная проблема заключается в ограниченной информации о данных систем-источников компании.

### **Процессно-ориентированное построение хранилища:**

Основное преимущество данной стратегии развития заключается в том, что модель бизнес-процесса содержит формальное описание информационных потребностей пользователя. Еще одним преимуществом является возможность создания правил/условий, при которых получение данных из систем-источников становилось бы возможным или более эффективным.

### 2.1) Анализ и проектирование

Для анализа и проектирования выбрана семантическая модель объектов (табл.1).

Основные блоки	Примечание
1 блок. Бизнес-план предприятия	На основе данных об окружающей среде формируется поведенческая активность. Здесь определяется вид предоставляемого товара/услуги и степень их продвижения.
2 блок. Бизнес - процессы	Совокупность взаимосвязанных мероприятий или задач, направленных на создание определенного товара /услуги для покупателя/потребителя
3 блок. Ресурсы	Здесь отражается специфика ресурсной базы, которая создана для выполнения задач, бизнес-процессов.

Табл.1 Архитектура предприятия

Для более детального понимания остановимся на втором блоке. Каждый бизнес-процесс можно отнести к одному из трех видов:

- бизнес-процессы, которые управляют функционированием системы;
- бизнес-процессы, участвующие в производстве/снабжении товара/услуги (операционные процессы);
- бизнес-процессы, описывающие последовательность задач, которые выполняются для достижения цели.

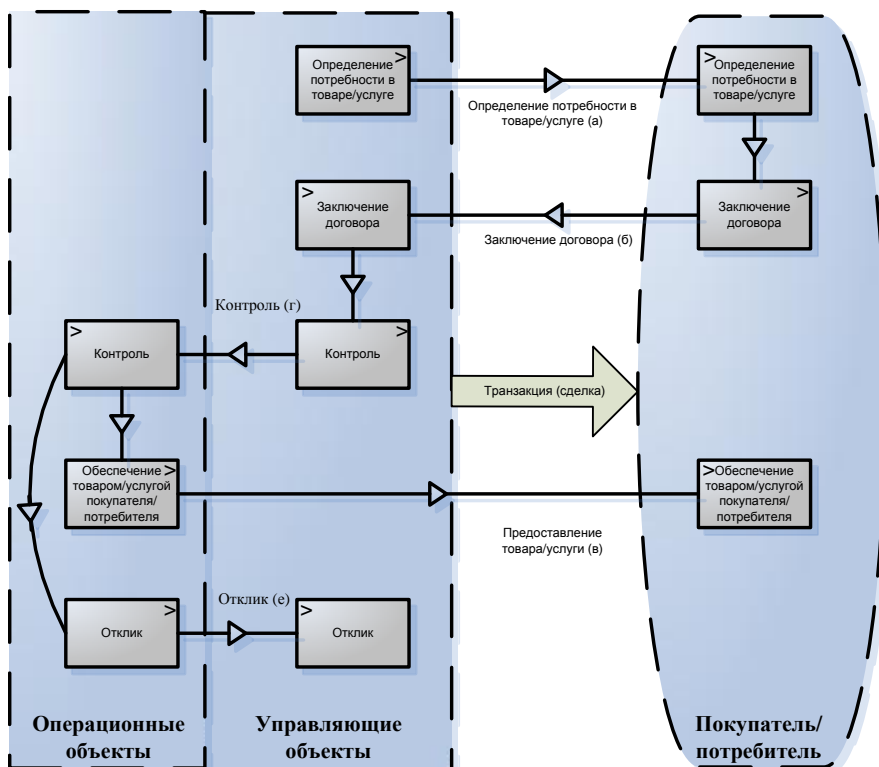
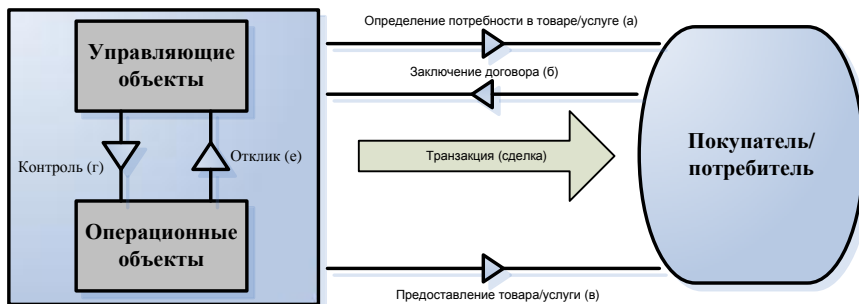


Рис.1 Схема взаимодействия связанных бизнес-процессов

Расширение понятия бизнес-процесс приводит к более полному представлению о компании и сфере ее бизнеса:

**Рассмотрим схему взаимодействия связанных бизнес-процессов (рис.1):** она состоит из бизнес-объектов и транзакций. Здесь под бизнес-объектом понимается встроенный бизнес-процесс, включающий в себя набор задач, выполнение которых приводит к общей цели - предоставление покупателю/потребителю спектра товаров/услуг. Бизнес-объекты связаны между собой операциями, и каждая транзакция выполняет ровно две задачи. Нужно понимать, что каждый бизнес-объект и соответствующая ему транзакция могут быть разложены на несколько бизнес-объектов и транзакций соответственно. Будем различать два основных принципа такого разложения:

**Обсуждение условий работы (сделки):** в процессе переговоров обе стороны договариваются об определенных условиях предоставления товаров/услуг, в связи с чем возникают три последовательных операции:

- определение потребности в товаре/услуге (а)
- оформление сделки (б)
- предоставление товара/услуги (в)

**Принцип работы схемы.** Бизнес-объекты делятся на две группы: управляющие и операционные. Они обеспечивают динамику в поведении бизнес-процессов. Схема состоит из задач и операций, которые должны выполняться бизнес-объектами в определенной последовательности. Для упрощения названия задач, используются условные обозначения "> задача" или "задача>", что означает отправку или получение соответствующей задачи/операции. Принцип работы схемы основан на принципе сетей Петри.

### 3. Конструкции хранилища данных на основе модели бизнес-процессов

#### 3.1 Введение в исследование на основе компании предоставляющей услуги телематики.

Рассмотрим применение выбранного подхода к типичной модели бизнес-процесса в компании, предоставляющей услуги телематики (рис. 2).



Рис.2 Бизнес-процесс компании

Модель основана на результатах собственного проекта по оптимизации процессов компании предоставляющей услуги телематики.

В модели можно выделить два основных процесса, каждый из которых сосредоточен на определенный результат:

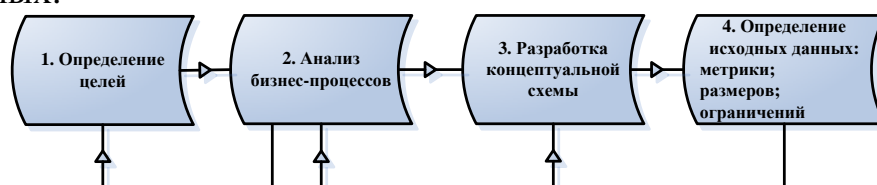
- 1) удовлетворение потребностей покупателя/потребителя в услугах компании;
- 2) нахождение конкурентных преимуществ посредством определения новых партнеров или пересмотра предложений от уже действующих.

Вспомогательные процессы, такие как *бизнес-планирование*, *формирование штат сотрудников* или *накопление библиотек знаний*, поможет реализовать цели основных процессов путем предоставления специфических услуг.

В дальнейшем, чтобы описать основные идеи данного подхода сосредоточимся на упрощенной форме всей модели бизнес-процесса, основанной на процессах связанных с предоставлением услуг телематики.

### 3.2 Процесс определения исходных данных для построения ХД

Введем в нашу модель процесс определения исходных параметров хранилища данных.



**Рис. 3.** Основные этапы процесса определения исходных параметров ХД

Подход включает в себя четыре основных этапа (рис. 3). Если первые три этапа соответствуют в значительной степени семантической модели объектов методологии и являются основополагающими для определения характеристик хранилища данных, то четвертый этап позволяет сгладить определение структуры хранилища данных.

После определения целей системы (1 этап), анализируется сеть бизнес-процессов, чтобы получить полное представление о системе (2 этап). Далее разрабатывается так называемая концептуальная схема (3 этап), в которой описывается структура данных, необходимых для выполнения выбранного бизнес-процесса. Вследствие этого можно определить исходную структуру хранилища данных (4 этап).

Четвертый этап более детально описывается в разделе 4.

#### 3.2.1 Определение цели

Вначале необходимо определить цели и соответствующие службы компании, которые будут их достигать. В основном цели две:

- 1) определить товары/услуги, которые будут предоставляться;
- 2) количество этих товаров/услуг.

Поиск целей это сложный и творческий процессом. При их определении мы должны учитывать все интересы как внутренней, так и внешней среды. Как известно на нормальное функционирование компании влияет целый ряд факторов, таких как правовая и политическая стабильность, вход в отрасль “новичков” и т.д. В связи, с чем мы введем следующее упрощение:

- достижение цели возможно несколькими способами;
- внешняя среда не влияет на достижение цели.

Отсюда мы получим цели для выбранной компании:

1. Предоставление услуг телематики
  2. Обучение работников (повышение их квалификации, переаттестация)
  3. Расширение границ влияния на рынке в связи с изучением опыта аналогичных компаний и компаний заменителей
  4. Расширение (в высоту/ширину) функциональных возможностей Call-центра
- В связи с нашей моделью (рисунок 2) основными процессами компании являются:

1. Предоставление услуг телематики
2. Определение конкурентных преимуществ

Отсюда можно сделать вывод, что предоставление услуг телематики (как цель компании), является одним из основных процессов компании.

### 3.2.2 Анализ бизнес-процессов

Подробно анализируя бизнес-процесс, его, как правило, максимально детализируем, что дает более глубокое понимание этого процесса и представление о разделах будущего хранилища данных. В ходе анализа описываются процедуры, которые должны реализовывать поставленные цели.

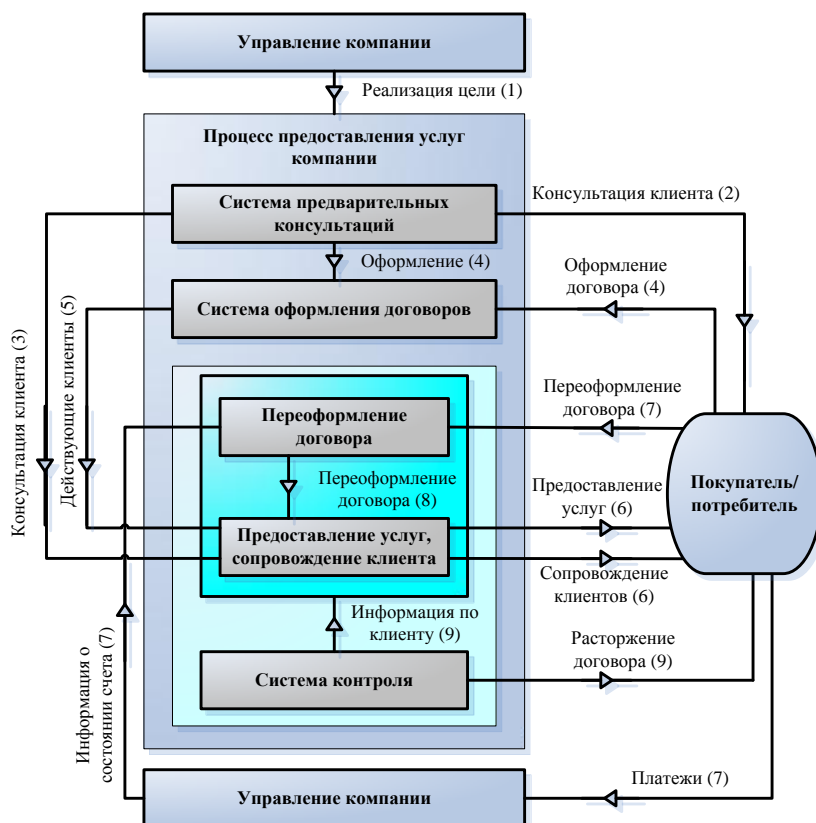


Рис. 4 Схема для анализа бизнес-процессов компании (более подробно, рассмотрена в Приложении А)

Мы выделим два вида процедур - управляющие и операционные. Управляющие процедуры вносят непосредственный вклад в реализацию поставленных целей, операционные помогают в реализации основных процедур. На рисунке 4 показано более детальное представление схемы, изображенной на рисунке 2, т.к. цель анализа - выявить подпроцессы, проходящие в ходе выполнения основных процессов.

Сформулированные цели осуществляются *управлением компании* (1). Для обеспечения более высокой эффективности предприятия проводятся

*предварительные консультации*, которые повышают информированность потенциальных покупателей/потребителей об услугах и сервисах компании (2). Данный подпроцесс осуществляет система консультаций, осуществляемых в до договорной период (3). Система оформления договоров получает заявку от покупателя/потребителя с целью заключения договора (4), что является подпроцессом в предоставлении услуги, сопровождения клиента (5). Мы не станем углубляться в этот под-процесс, скажу лишь, что он обеспечивает достижение основной цели - предоставление услуг (6). В зависимости от условий договора цикл подтверждения действия договора (мы определим его как ежемесячный цикл) определяется оплатой потребленных услуг (7). Заявки о переоформлении договора информируют компанию о количестве клиентов повторно (или на других условиях) заключивших договор (8). Система контроля сообщает о статусе клиента (9).

После краткого описания процесса предоставления услуг телематики, мы получим концептуальную схему.

### 3.2.3 Концептуальная схема

Основная цель во время перехода от второго к третьему этапу (рис. 3) заключается в передаче максимального количества информации, для уточнения нужных ресурсов (3 блок табл. 1).

Рассматривая общую схему организации, опустим такие моменты, как работа с персоналом, закупка оборудования и т.п.

Особенностью разработки концептуальной схемы является ее визуализация. На ри.5 слева расположены так называемые независимые объекты, а справа зависимые, эта зависимость показана пунктирными линиями. Следовательно, зависимость между объектами проходит слева направо. Например, *подключение клиента* зависит от *заключения договора на предоставление товаров/услуг*, а *заключение договора* зависит от *предоставленных консультаций* и самого клиента.

Для проработки 4 этапа выделим лишь ключевые характеристики блоков, которые наиболее существенны. Из этих соображений выделим три блока:

- 1) покупатель/продавец;
- 2) предварительные консультации;
- 3) блок заключение договора.

Для каждого из них обозначим присущие им основные параметры.

Клиент	Консультации	Заключение договора
- Наименование фирмы - Реквизиты	- Тематика консультаций - Длительность - Как часто проводятся	- Тип (первичный/ вторичный)

Табл. 2 Основные блоки и их основные параметры

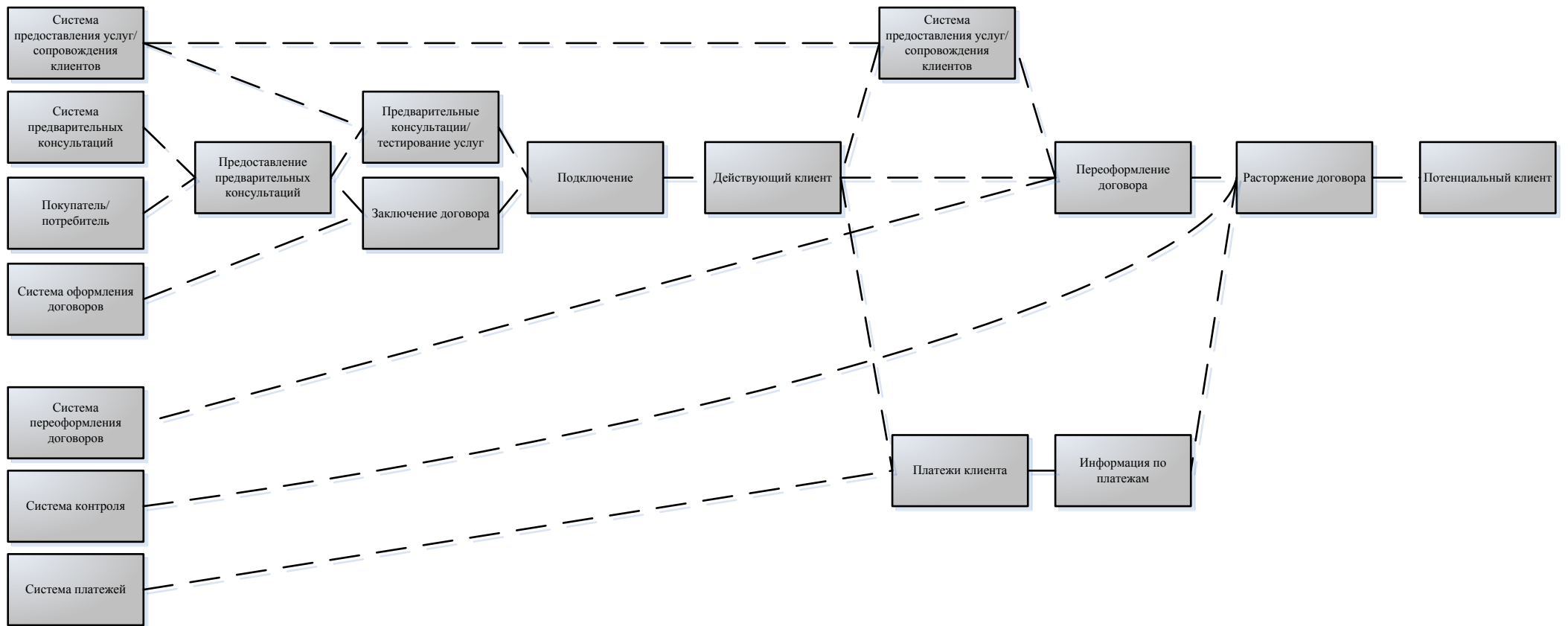


Рис. 5 Концептуальная схема

### **3.2.4 Определение структуры хранилища данных**

Исходя из данных, полученных при построении и описании концептуальной схемы можно определить структуру хранилища данных

Как говорилось ранее, построение концептуальной схемы - дело творческое, поэтому четких правил по ее построению нет. В связи с этим воспользуемся общим подходом к моделированию структуры хранилища данных.

Визуализация существующих зависимостей в концептуальной схеме позволяет судить о размерности и структурной иерархии хранилища данных. При их определении стоит рассматривать только те объекты, в которых возможно определение метрик.

Следует помнить, что при формировании статистической информации, может настать необходимость суммирования элементов нижнего уровня иерархии с более высокими уровнями. Вследствие чего метрики могут быть классифицированы как полу-аддитивные и неаддитивные на основе этих статистических функций, т.е. подводя итоги по количеству подключенных клиентов за год, мы получим бессмысленный результат.

Описанная модель действия содержит явные петли обратной связи к предыдущим этапам, на которых можно внести изменения в модель (рис.3). Между этапами 2, 3 и 4 существует прямая и обратная связь, а между этапами 3(4) и 1 существует лишь транзитивная связь. Причем после обратной связи необходимо повторно пройти все этапы для определения новых метрик и т.д. В связи с этим можно сделать вывод, что построение структуры хранилища данных - повторяющийся и развивающийся процесс.

### **4. Пример хранилища данных**

На практике построение хранилища данных в основном необходимо для решения задач прогнозирования, поэтому очень важно не столкнуться с проблемами нехватки информации и их недостоверности, т.к. в этом случае говорить о качественном анализе не приходится. К тому же в хранилище данных с определенной периодичностью должны поступать агрегированные данные. И эти данные не должны описывать только внутреннюю ситуацию в компании, они так же должны давать представление о влиянии внешних факторов на деятельность компании (сведения о ценах конкурентов, общее состояние рынка и т.д.). Пример хранилища данных приведен на рисунке 6 и 7 (рисунки основаны на данных из Приложения А).

Большинство информации о внешних факторах можно получить бесплатно или с минимальными вложениями. Перед внесением этих сведений, возможно, их предварительно придется обработать, но главное что информацию можно получить.



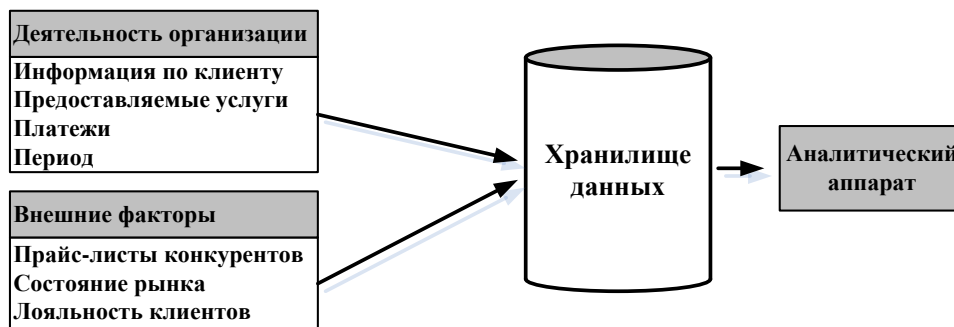


Рис. 6 Пример хранилища данных

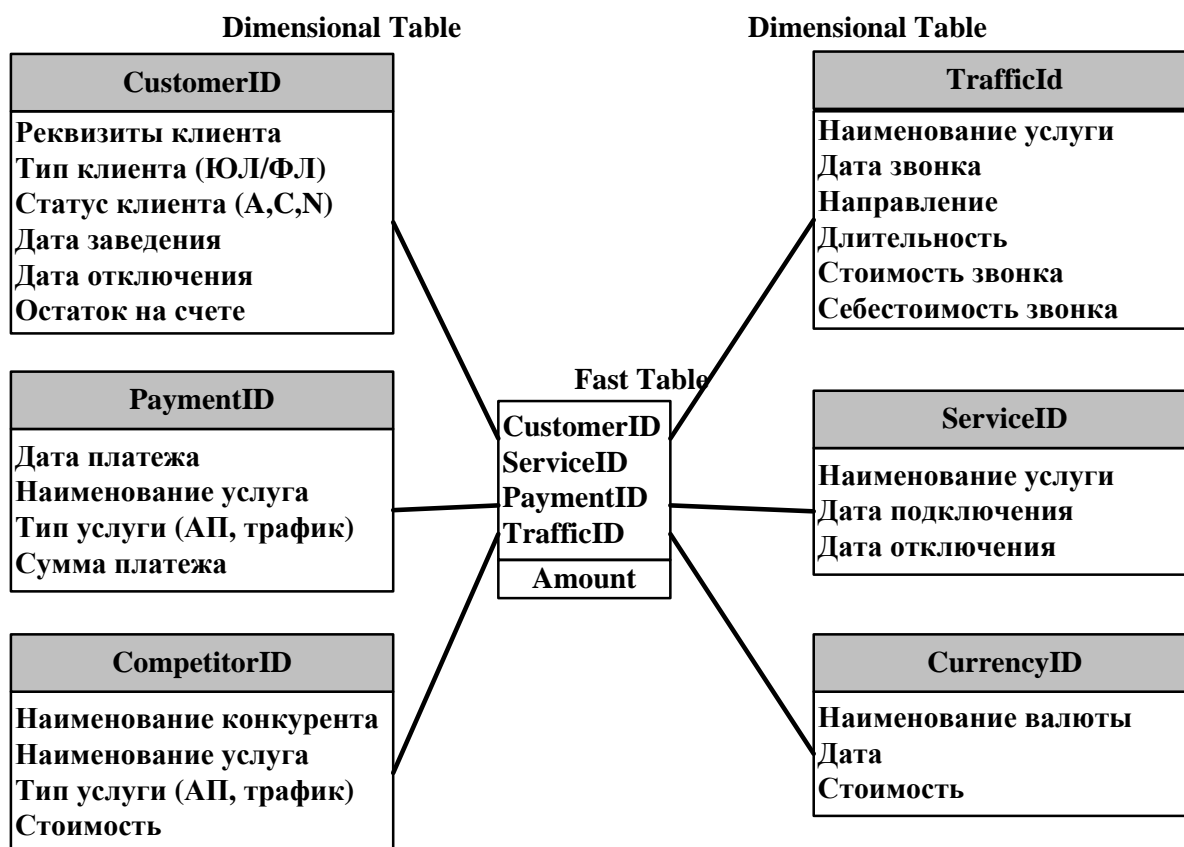


Рис. 7 Пример схемы звезда для хранилища данных

## 5. Итоги и перспективы

В этой статье я описал подход к построению хранилищ данных с ориентированием на бизнес-процесс. Данный подход наиболее актуален для руководителей подразделений/ организаций и наиболее удобен в определении основных параметров необходимых для построения хранилища данных.

# Приложение А. Блок-схема событий в ходе выполнения бизнес-процессов.

