

Диаграммы потоков данных (DFD)

Лекция 9 (25)

Понятие потока данных

Овчинников П.Е.

МГТУ «СТАНКИН»,

ст.преподаватель кафедры ИС

Терминология: информация и данные

Федеральный закон "Об информации, информационных технологиях и о защите информации" от 27.07.2006 [N 149-ФЗ](#) (последняя редакция)

информация

сведения (сообщения, данные) независимо от формы их представления

информационные технологии (ИТ)

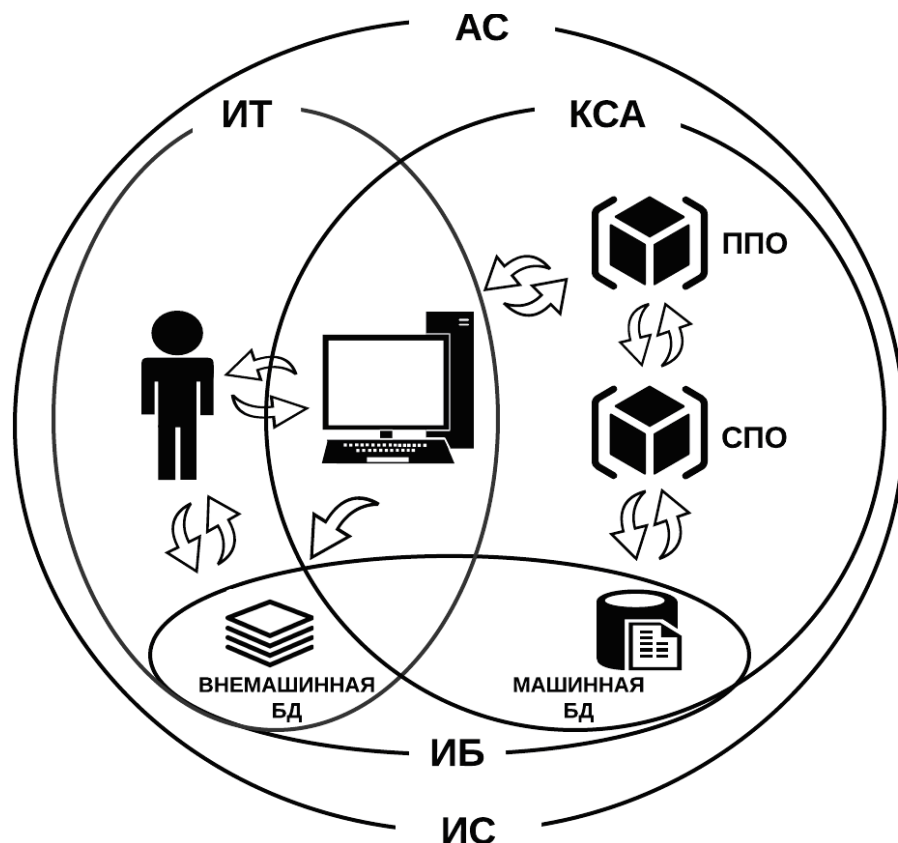
процессы, методы **поиска, сбора, хранения, обработки, предоставления, распространения** информации и способы осуществления таких процессов и методов

информационная система (ИС)

совокупность содержащейся **в базах данных** информации и обеспечивающих ее обработку информационных технологий и технических средств

Обозначения

АС - автоматизированная система
БД - база данных
ИБ - информационная база
ИС - информационная система
ИТ - информационные технологии
КСА - комплекс средств автоматизации
ППО - прикладное программное обеспечение
СПО - системное программное обеспечение



Терминология: данные и программы

ГОСТ 34.321-96 Информационные технологии (ИТ). Система стандартов по базам данных. Эталонная модель управления данными
данные (data)

информация, представленная **в формализованном виде**, пригодном для **передачи, интерпретации или обработки** с участием человека или автоматическими средствами

ГОСТ 19781-90 Обеспечение систем обработки информации программное.
Термины и определения
программа (Program)

данные, предназначенные для управления конкретными компонентами системы обработки информации в целях реализации определенного **алгоритма**

программное обеспечение (Software)

совокупность программ системы обработки информации и **программных документов**, необходимых для эксплуатации этих программ

программирование (Programming)

научная и практическая деятельность по созданию программ

Терминология: поток (программирование)

Поток данных (англ. *stream*) в программировании — абстракция, используемая для чтения или записи файлов, сокетов и т. п. в единой манере.

Потоки являются удобным унифицированным программным интерфейсом для чтения или записи файлов (в том числе специальных и, в частности, связанных с устройствами), сокетов и передачи данных между процессами

Пото́к выполне́ния (тред; от англ. *thread* — нить) — наименьшая единица обработки, исполнение которой может быть назначено ядром операционной системы.

Несколько потоков выполнения могут существовать в рамках одного и того же процесса и совместно использовать ресурсы, такие как память, тогда как процессы не разделяют этих ресурсов.

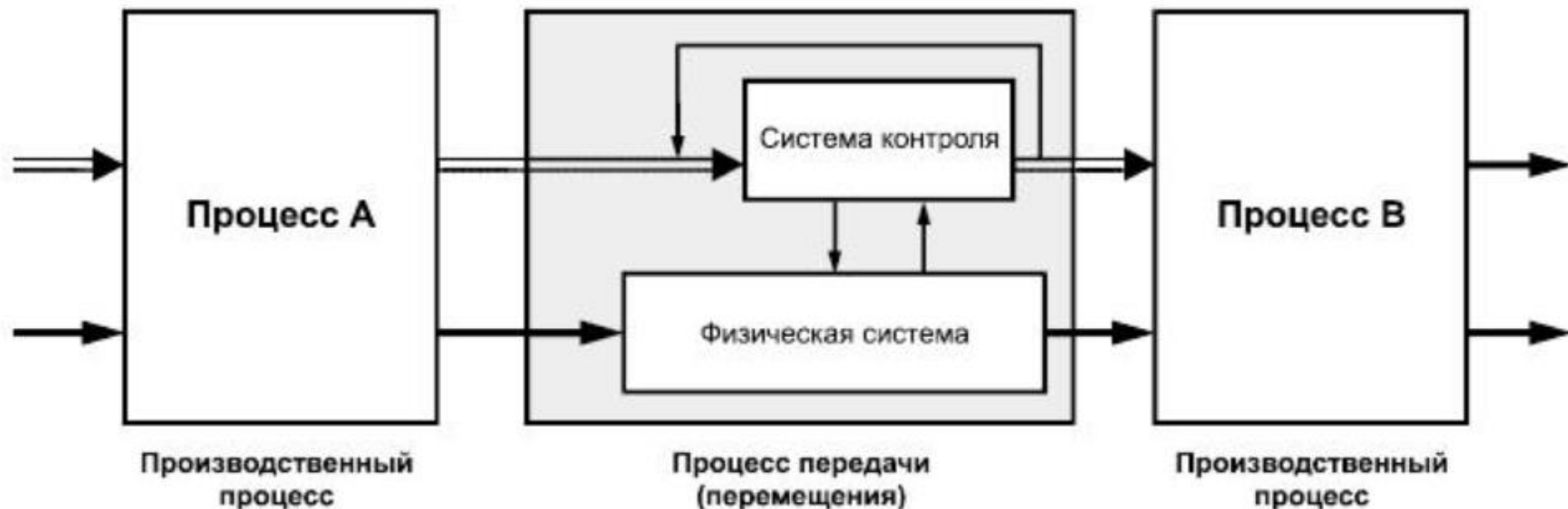
В частности, потоки выполнения разделяют инструкции процесса (его код) и его контекст (значения переменных, которые они имеют в любой момент времени). В качестве аналогии потоки выполнения процесса можно уподобить нескольким вместе работающим поварам. Все они готовят одно блюдо, читают одну и ту же кулинарную книгу с одним и тем же рецептом и следуют его указаниям, причём не обязательно все они читают на одной и той же странице.

Терминология: поток (производство)

ГОСТ Р ИСО 15531-43-2011 Системы промышленной автоматизации и интеграция. Данные по управлению промышленным производством. Часть 43. Информация для управления производственными потоками. Модель данных для мониторинга и обмена производственной информацией

поток (flow):

движение множества физических или информационных объектов в пространстве и времени



Терминология: схема данных

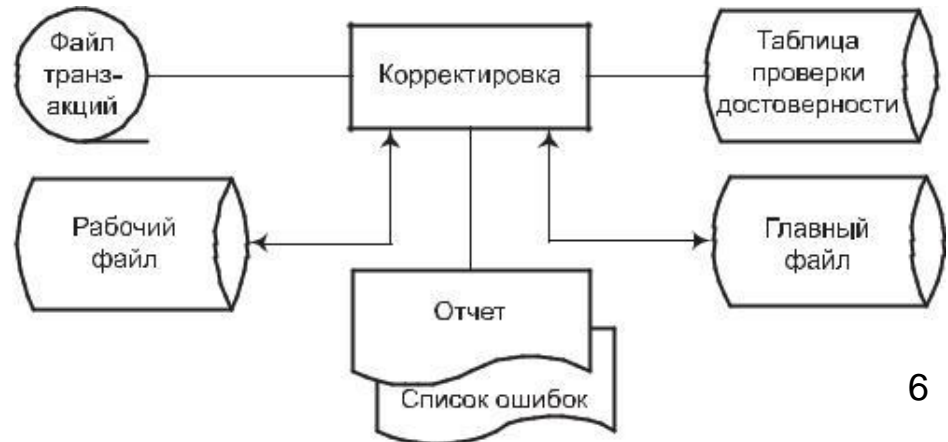
ГОСТ 19.701-90 (ИСО 5807-85) Единая система программной документации (ЕСПД). Схемы алгоритмов, программ, данных и систем. Обозначения условные и правила выполнения

Схемы данных отображают путь данных при решении задач и определяют этапы обработки, а также различные применяемые носители данных. Схема данных состоит из:

- 1) символов **данных** (символы данных могут также указывать вид носителя данных)
- 2) символов **процесса**, который следует выполнить над данными (символы процесса могут также указывать функции, выполняемые вычислительной машиной)
- 3) символов **линий**, указывающих **потоки данных** между процессами и (или) носителями данных

- 4) специальных символов, используемых для облегчения написания и чтения схемы

Символы данных предшествуют и следуют за символами процесса. Схема данных начинается и заканчивается символами данных



Терминология: измерение потоков

Поток событий

последовательность событий, которые наступают в случайные моменты времени

Интенсивность потока

среднее **число событий**, которые появляются в **единицу времени**:

100/сек

10/час

1/год

Если постоянная интенсивность потока известна, то вероятность появления k событий простейшего потока за время длительностью t определяется формулой Пуассона:

$$P_t(k) = \frac{(\lambda t)^k e^{-\lambda t}}{k!}$$

Марковский процесс

случайный процесс, эволюция которого после любого заданного значения временного параметра t **не зависит** от эволюции, предшествовавшей t , при условии, что значение процесса в этот момент фиксировано («будущее» процесса не зависит от «прошлого» при известном «настоящем»); другая трактовка (Вентцель): «будущее» процесса зависит от «прошлого» лишь через «настоящее»).

Терминология: измерение потоков

Событие в объектно-ориентированном программировании

сообщение, которое возникает в различных точках исполняемого кода при выполнении определённых условий

События предназначены для того, чтобы иметь возможность предусмотреть реакцию программного обеспечения

Для решения поставленной задачи создаются **обработчики событий**: как только программа попадает в заданное состояние, происходит событие, посылается сообщение, а обработчик перехватывает это сообщение. В общем случае в обработчик не передаётся ничего, либо передаётся ссылка на объект, инициировавший (породивший) обрабатываемое событие.

В особых случаях в **обработчик передаются значения некоторых переменных** или ссылки на какие-то другие объекты, чтобы обработка данного события могла учесть контекст возникновения события

Самое простое событие — это событие, сообщающее о начале или о завершении некоторой процедуры. Событие, по сути, сообщает об **изменении состояния** некоторого объекта. Наиболее наглядно события представлены в пользовательском интерфейсе, когда каждое **действие пользователя** порождает **цепочку событий**, которые затем обрабатываются в приложении

Терминология: измерение потоков

Транзакция (англ. *transaction*)

группа последовательных операций с базой данных, которая представляет собой **логическую единицу работы с данными**

Транзакция может быть выполнена либо целиком и успешно, соблюдая целостность данных и независимо от параллельно идущих других транзакций, либо не выполнена вообще, и тогда она не должна произвести никакого эффекта.

Транзакции обрабатываются транзакционными системами, в процессе работы которых создаётся история транзакций.

Различают последовательные (обычные), параллельные и распределённые транзакции.

Распределённые транзакции подразумевают использование более чем одной транзакционной системы и требуют намного более сложной логики (например, two-phase commit — двухфазный протокол фиксации транзакции).

Также в некоторых системах реализованы автономные транзакции, или подтранзакции, которые являются автономной частью родительской транзакции.

Терминология: измерение потоков

Ка́дровая частота́, частота кадровмен ([англ.](#) *Frames per Second (FPS)*, *Frame rate*, *Frame frequency*)

количество сменяемых кадров за единицу времени в компьютерных играх, [телевидении](#) и [кинематографе](#)

Общепринятая единица измерения — кадры в секунду

Транзакции в секунду

единица измерения, определяющая количество произведённых операций по пересылке данных в секунду в каком-либо [канале связи](#)

Бод ([англ.](#) *baud*) в [связи](#) и [электронике](#) — единица измерения [символьной скорости](#), количество изменений информационного параметра несущего периодического сигнала в [секунду](#)

Битре́йт (от [англ.](#) *bitrate*) — количество бит, используемых для передачи/обработки данных в единицу времени. Битрейт принято использовать при измерении эффективной скорости передачи потока данных по каналу, то есть минимального размера канала, который сможет пропустить этот поток без задержек.

Битрейт выражается [битами в секунду](#) (бит/с, *bps*), а также производными величинами с [приставками](#) кило- (кбит/с, *kbps*), мега- (Мбит/с, *Mbps*) и т. д.

Терминология: измерение производительности

APDEX (Application Performance Index)

индекс производительности приложений. Открытый международный стандарт, разработанный с целью формирования объективной оценки показателей производительности корпоративных информационных систем

APDEX является числовой мерой **удовлетворенности пользователей производительностью** приложений.

Для расчета APDEX собирается множество статистических данных о **времени исполнения** операций приложением. Далее производится компоновка данных, с учетом факторов влияния, и математическое преобразование, которое приводит все эти данные к значению в интервале от «0» до «1» («0» = все пользователи разочарованы, никто не может дождаться ответа от приложения, «1» = все пользователи довольны, приложение у всех работает очень быстро).

В зависимости от вариации алгоритма APDEX, важным фактором влияния на расчет может быть приоритет операций с точки зрения бизнеса. В большинстве алгоритмов приоритет операции используется для ранжирования операций, но в некоторых алгоритмах приоритет операции может в значительной степени влиять на результирующий APDEX по приложению в целом.

Терминология: измерение производительности

Пропускная способность

метрическая характеристика, показывающая соотношение предельного количества проходящих **единиц** (информации, предметов, объёма) **в единицу времени** через канал, систему, узел

Теория массового обслуживания, или **очередей** (англ. *queueing theory*), — раздел теории вероятностей, целью исследований которого является рациональный выбор **структуры** системы обслуживания и **процесса** обслуживания на основе изучения **потоков требований** на обслуживание, поступающих в систему и выходящих из неё, **длительности ожидания** и **длины очередей**.

В теории массового обслуживания используются методы теории вероятностей и математической статистики

Очередь — абстрактный тип данных с дисциплиной доступа к элементам «первый пришёл — первый вышел» (FIFO, англ. *first in, first out*). Добавление элемента (принято обозначать словом enqueue — поставить в очередь) возможно лишь в конец очереди, выборка — только из начала очереди (что принято называть словом dequeue — убрать из очереди), при этом выбранный элемент из очереди удаляется

Терминология: измерение качества

ГОСТ Р ИСО 9000-2015 Системы менеджмента качества. Основные положения и словарь

качество (quality):

степень соответствия совокупности присущих **характеристик объекта требованиям**

требование (requirement):

потребность или **ожидание**, которое установлено, обычно предполагается или является обязательным.

Примечания

- 1 Слова "обычно предполагается" означают, что это общепринятая практика *организации и заинтересованных сторон*, что рассматриваемые потребности или ожидания предполагаются.
- 2 Установленным является такое требование, которое определено, например, в *документированной информации*

объект (object), сущность (entity), элемент (item):

что-либо воспринимаемое или воображаемое.

Примечание - Объекты могут быть материальными (например, двигатель, лист бумаги, алмаз), нематериальными (например, коэффициент конверсии, план проекта) или воображаемыми (например, будущее положение организации).

Терминология: измерение качества

ГОСТ Р ИСО 9000-2015 Системы менеджмента качества. Основные положения и словарь

характеристика (characteristic):
отличительное свойство.

Характеристика может быть **присущей** или **присвоенной**

Характеристика может быть **качественной** или **количественной**

Существуют различные классы характеристик, такие как:

- физические (например, механические, электрические, химические или биологические характеристики)
- органолептические (например, связанные с запахом, осязанием, вкусом, зрением, слухом)
- этические (например, вежливость, честность, правдивость)
- **характеристики, связанные со временем** (например, пунктуальность, безотказность, доступность, непрерывность)
- эргономические (например, физиологические характеристики или связанные с безопасностью человека)
- **функциональные** (например, максимальная скорость самолета).

характеристика качества (quality characteristic):

присущая объекту **характеристика**, относящаяся к **требованию**

Терминология: измерение качества

В курсовом проекте:

Определение числовых **показателей** для поставленной **цели моделирования**

- **уточнение вопросов**, на которые должна ответить модель
- формализация вопросов и **определение показателей**
- **определение единиц измерения** для каждого показателя

Определение числовых показателей для **цели** потенциального **проекта автоматизации**

- формализация **показателей качества** для автоматизируемых процессов
- определение **зависимости** показателей качества от **степени автоматизированности** процессов
- расчет **потенциального эффекта** от выполнения проекта автоматизации

DFD – диаграммы потоков данных

DFD

общепринятое сокращение от англ. *data flow diagrams* — диаграммы потоков данных

Так называется методология графического структурного анализа, описывающая:

1. **внешние** по отношению к системе **источники** и **адресаты** данных
2. логические **функции**
3. **потоки** данных и
4. **хранилища** данных, к которым осуществляется доступ

Диаграмма потоков данных (data flow diagram, DFD) — один из основных инструментов структурного анализа и проектирования информационных систем, существовавших до **широкого распространения** UML

Несмотря на имеющее место в современных условиях смещение акцентов от структурного к объектно-ориентированному подходу к анализу и проектированию систем, «старинные» структурные нотации по-прежнему широко и эффективно используются как в бизнес-анализе, так и в анализе информационных систем

DFD – диаграмма потоков данных

Информационная система **принимает извне** и **передает наружу** потоки данных. Для обозначения элементов внешней среды функционирования системы используется понятие внешней сущности.

Внутри системы существуют **процессы преобразования информации**, порождающие новые потоки данных. Потоки данных могут:

- поступать **на вход** к процессам, в т.ч. от внешних сущностей
- **помещаться** в накопители данных и **извлекаться** из них
- **передаваться** к другим процессам и внешним сущностям

Модель DFD, как и большинство других структурных моделей — иерархическая модель. Каждый процесс может быть подвергнут декомпозиции, то есть разбиению на структурные составляющие, отношения между которыми в той же нотации могут быть показаны на отдельной диаграмме.

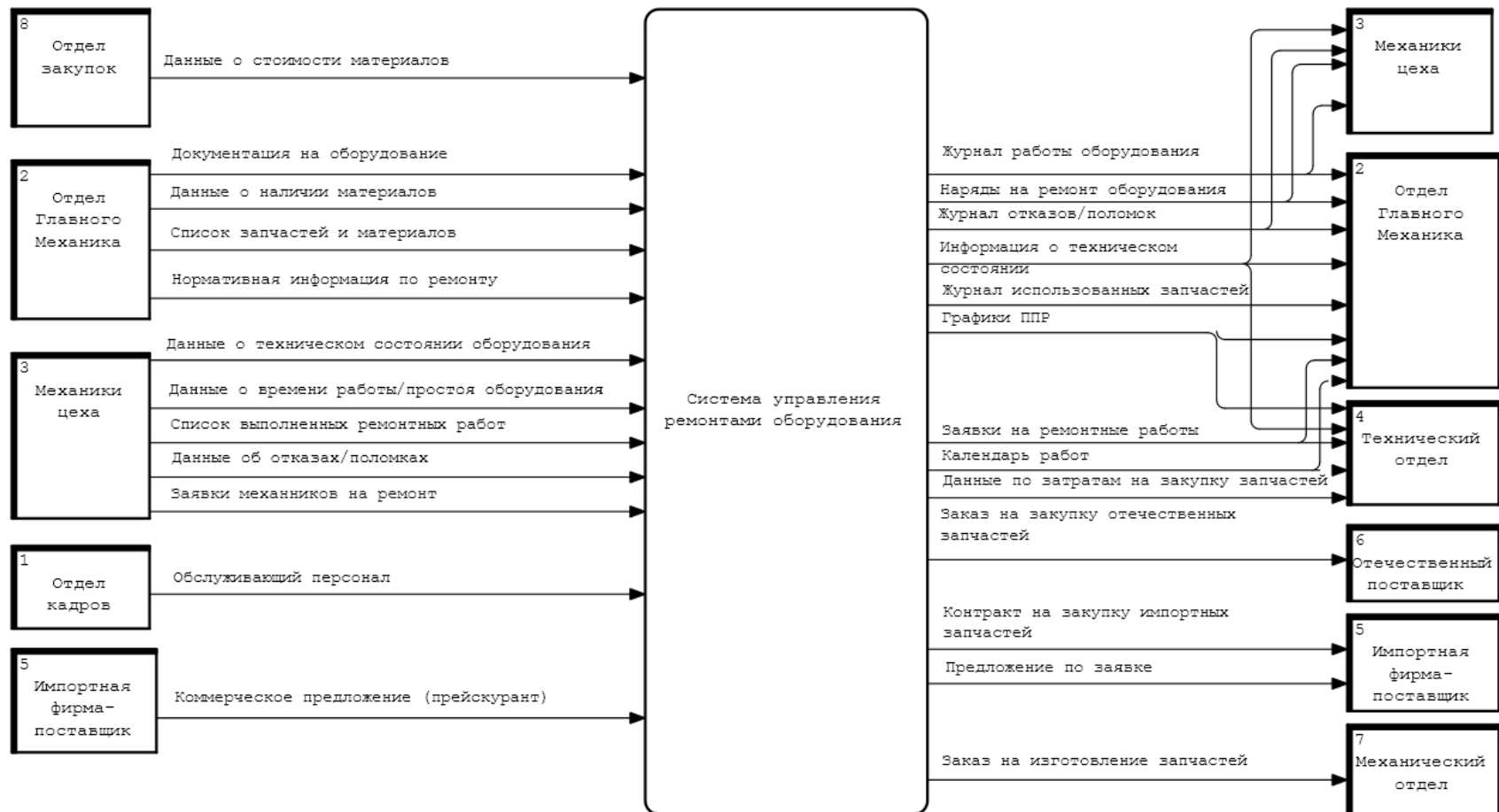
Когда достигнута требуемая глубина декомпозиции — процесс нижнего уровня сопровождается мини-спецификацией (текстовым описанием).

Кроме того, нотация DFD поддерживает понятие **подсистемы** — структурного компонента разрабатываемой системы.

DFD: Контекстная диаграмма

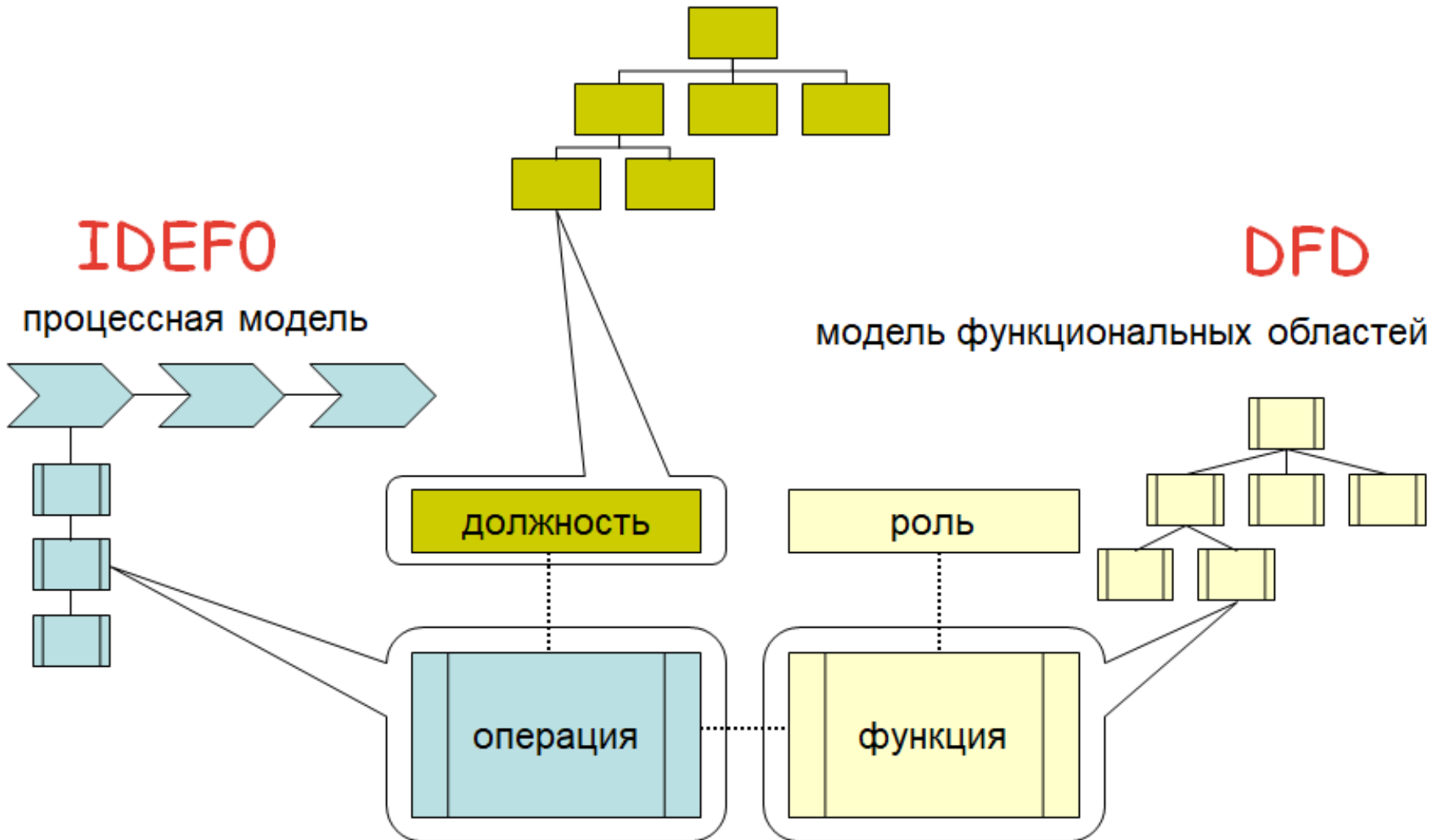
Нотация DFD — удобное средство для формирования контекстной диаграммы, то есть диаграммы, показывающей **разрабатываемую ИС** в **коммуникации с внешней средой**.

Её назначение — **ограничить рамки** системы, определить, где заканчивается разрабатываемая система и начинается среда.



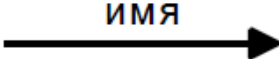
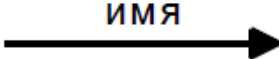






DFD: Комплексные модели

функциональная модель предприятия





Нотации DFD

Исторически сложилось так, что для описания диаграмм [DFD](#) используются две нотации - **Йодана** (Yourdon) и **Гейна-Сарсона** (Gane-Sarson), отличающиеся синтаксисом.

Компонента	Нотация Йордана	Нотация Гейна-Сарсона
поток данных		
процесс		
хранилище		
внешняя сущность		

Нотация Йордона

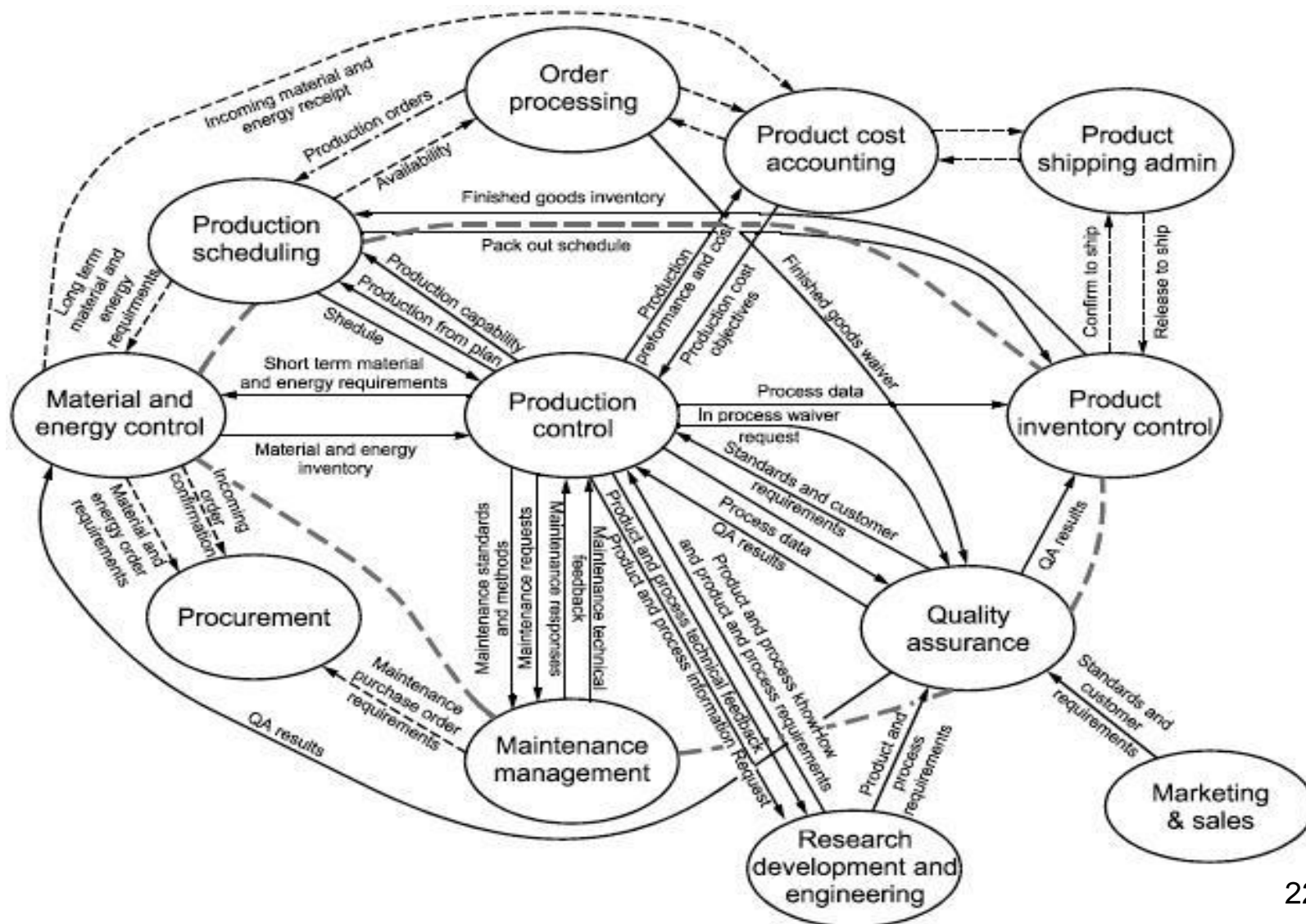
ГОСТ Р МЭК 62264-1-2014 Интеграция систем управления предприятием. Часть 1. Модели и терминология

Символ	Определение
	<p>Функцию представляют в виде помеченного эллипса. Каждую функцию можно в дальнейшем разбить на более мелкие функции, с большим уровнем дробления</p>
	<p>Сплошная стрелка представляет собой группу данных, передаваемую между функциями. Все сплошные линии имеют имя потока данных.</p> <p>Поток данных на одном уровне функциональной иерархии может представляться одним или несколькими потоками на более низком иерархическом уровне</p>
	<p>Пунктирная стрелка представляет собой группу данных, передаваемую между функциями. Данные не сохраняются на границах системы управления предприятием и приводятся для иллюстрации контекста функций. Все сплошные линии имеют имя потока данных.</p> <p>Потоки данных без имен, обозначенные пунктирными стрелками, в данной модели не идентифицированы</p>

Нотация Йордона

ГОСТ Р МЭК 62264-1-2014 Интеграция систем управления предприятием.

Часть 1. Модели и терминология



Нотация Гейна-Сарсона



Нотация Гейна-Сарсона



Нотация Гейна-Сарсона

расширение для систем реального времени

Квазинепрерывный поток (лат. quasi – как будто, якобы) – поток данных, непрерывный во времени. Отображается линией с двумя стрелками на конце.

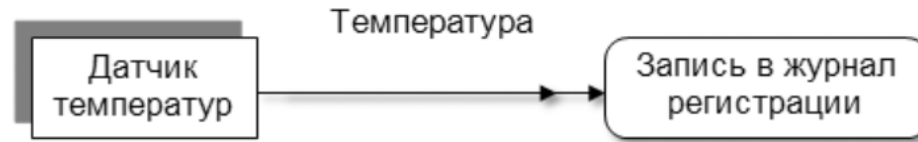


Рис. 6.25. Квазинепрерывный поток

Управляющий процесс – процесс, формирующий сигналы управления на выходе.

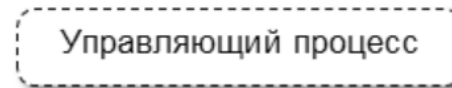


Рис. 6.26. Управляющий процесс

Управляющий поток – управляющая информация, запускающая процесс (подсистему) или изменяющая ход его выполнения.

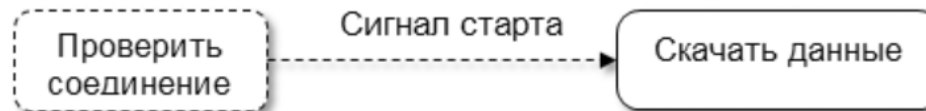
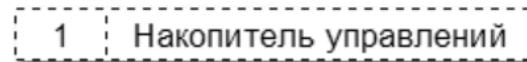


Рис. 6.27. Управляющий поток

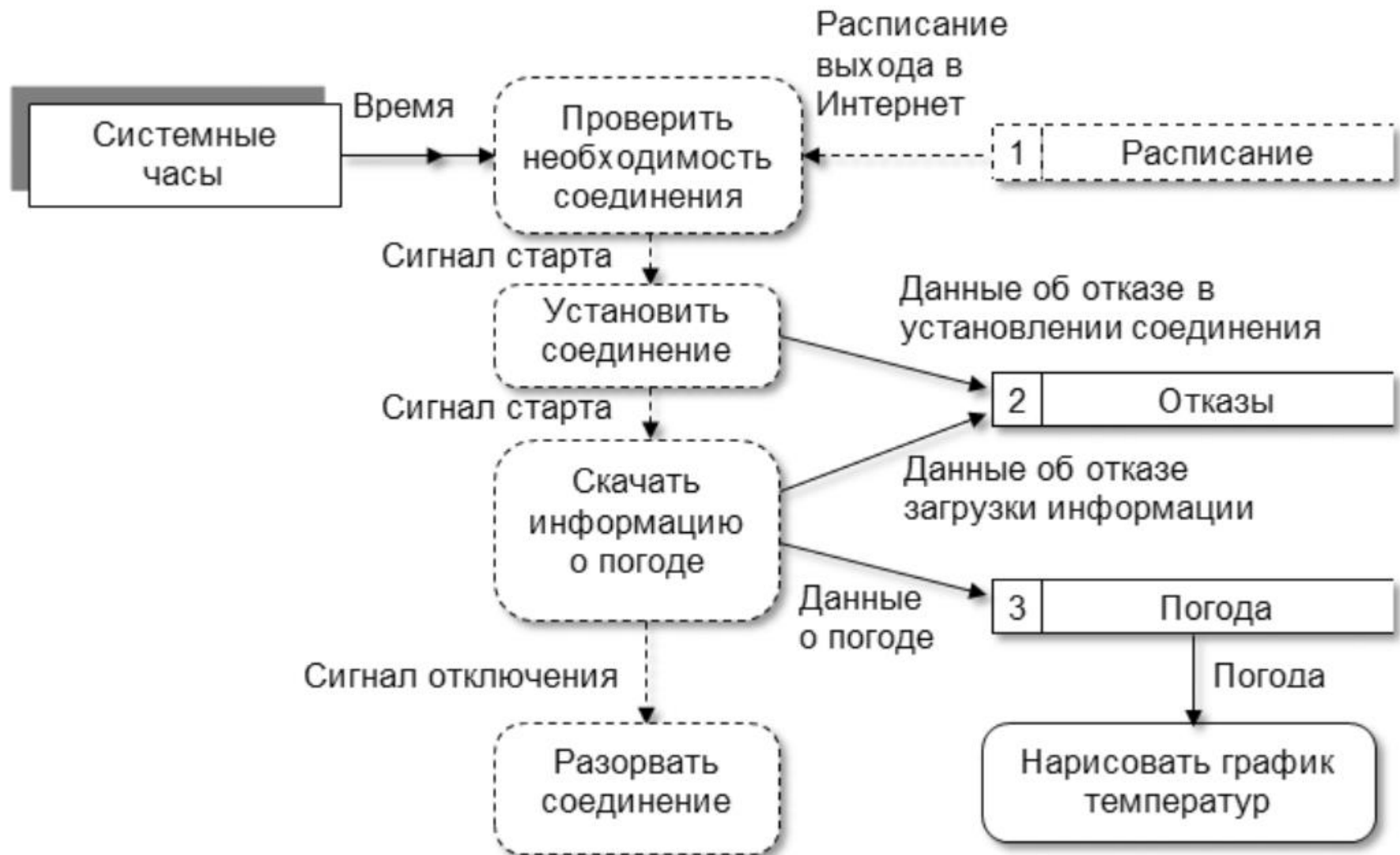
Использование управляющих потоков позволяет отделить управляющую информацию от обрабатываемой, как это делается на диаграммах IDEF0.

Накопитель управлений – накопитель управляющих потоков.



Нотация Гейна-Сарсона

расширение для систем реального времени



Нотация Гейна-Сарсона

использование для материальных потоков



Сравнение: схемы данных и работы системы

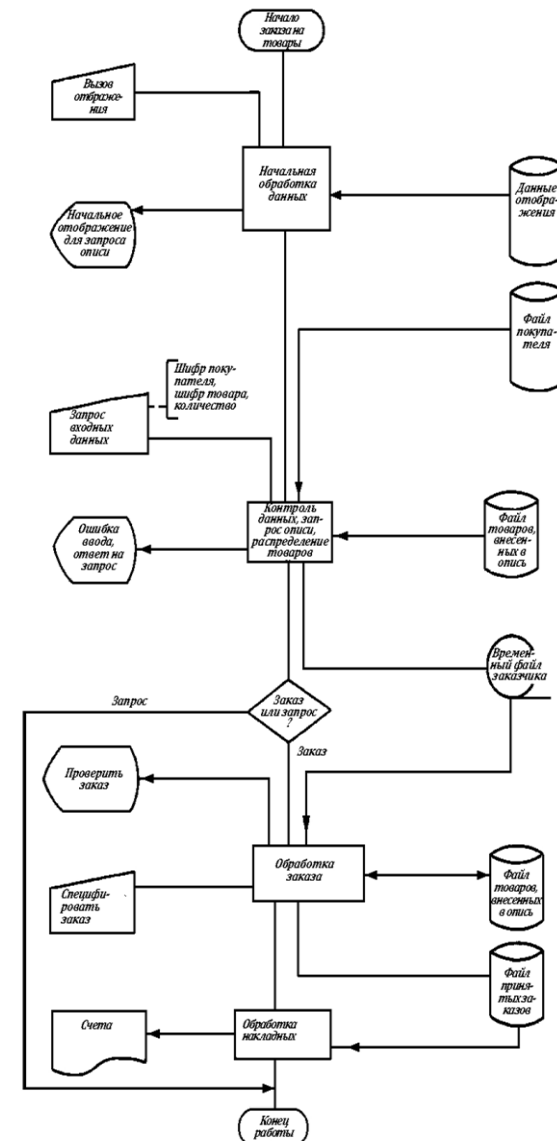
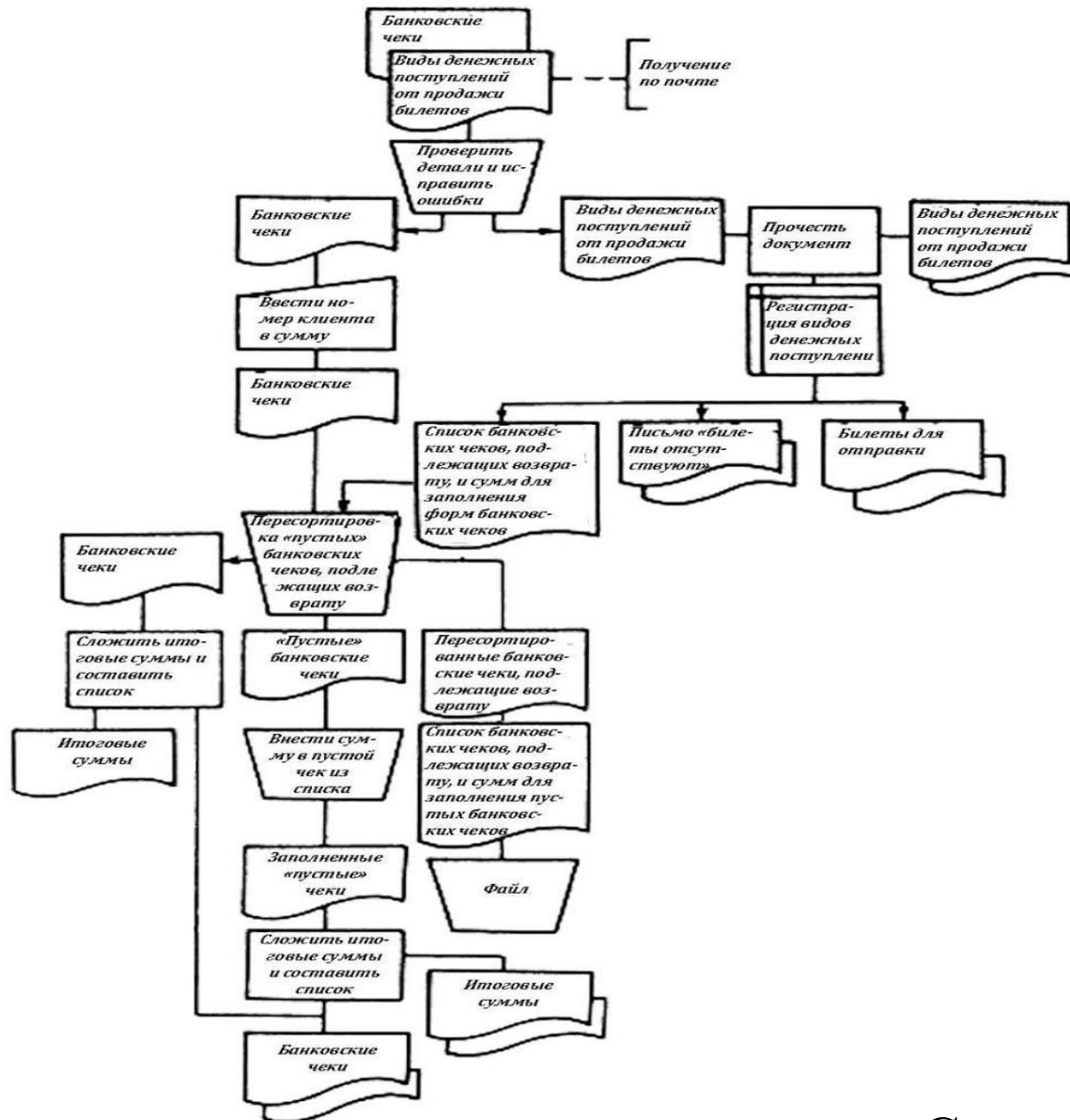
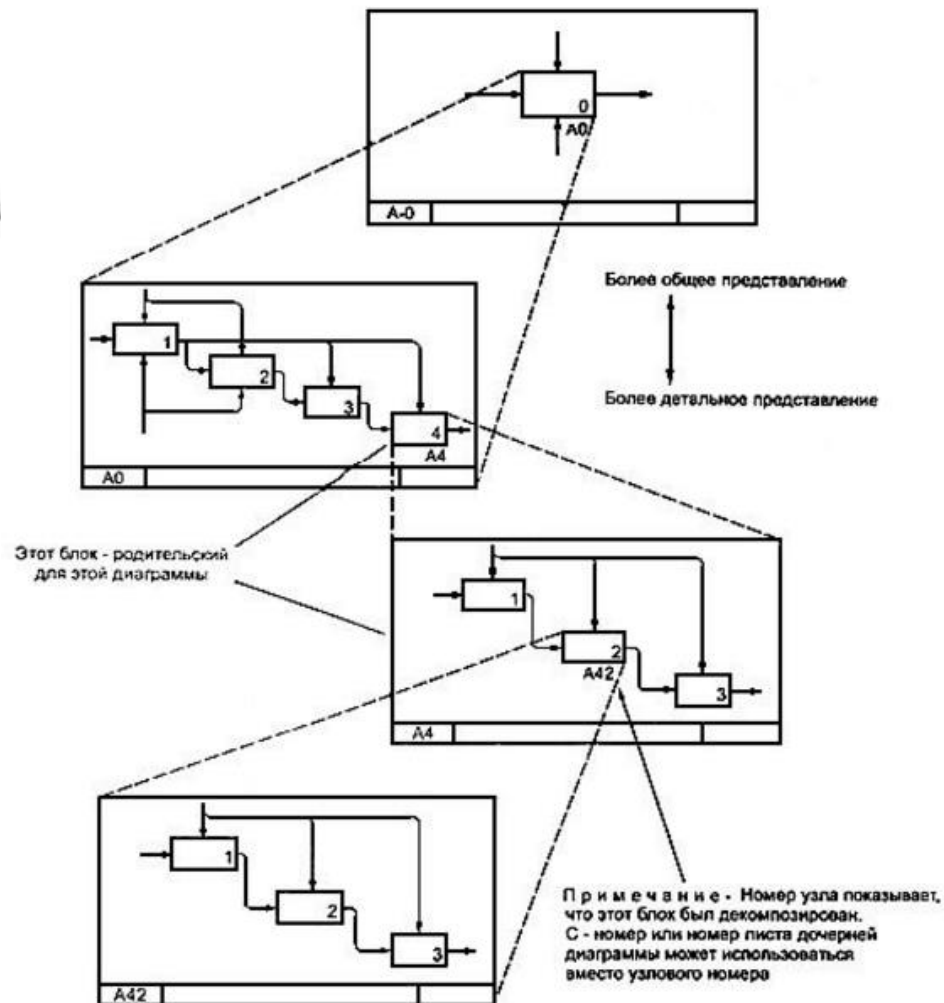


Схема работы системы

Сравнение: SADT

[Р 50.1.028-2001](#) Информационные технологии поддержки жизненного цикла продукции. Методология функционального моделирования



Сравнение: неформальные процессные модели

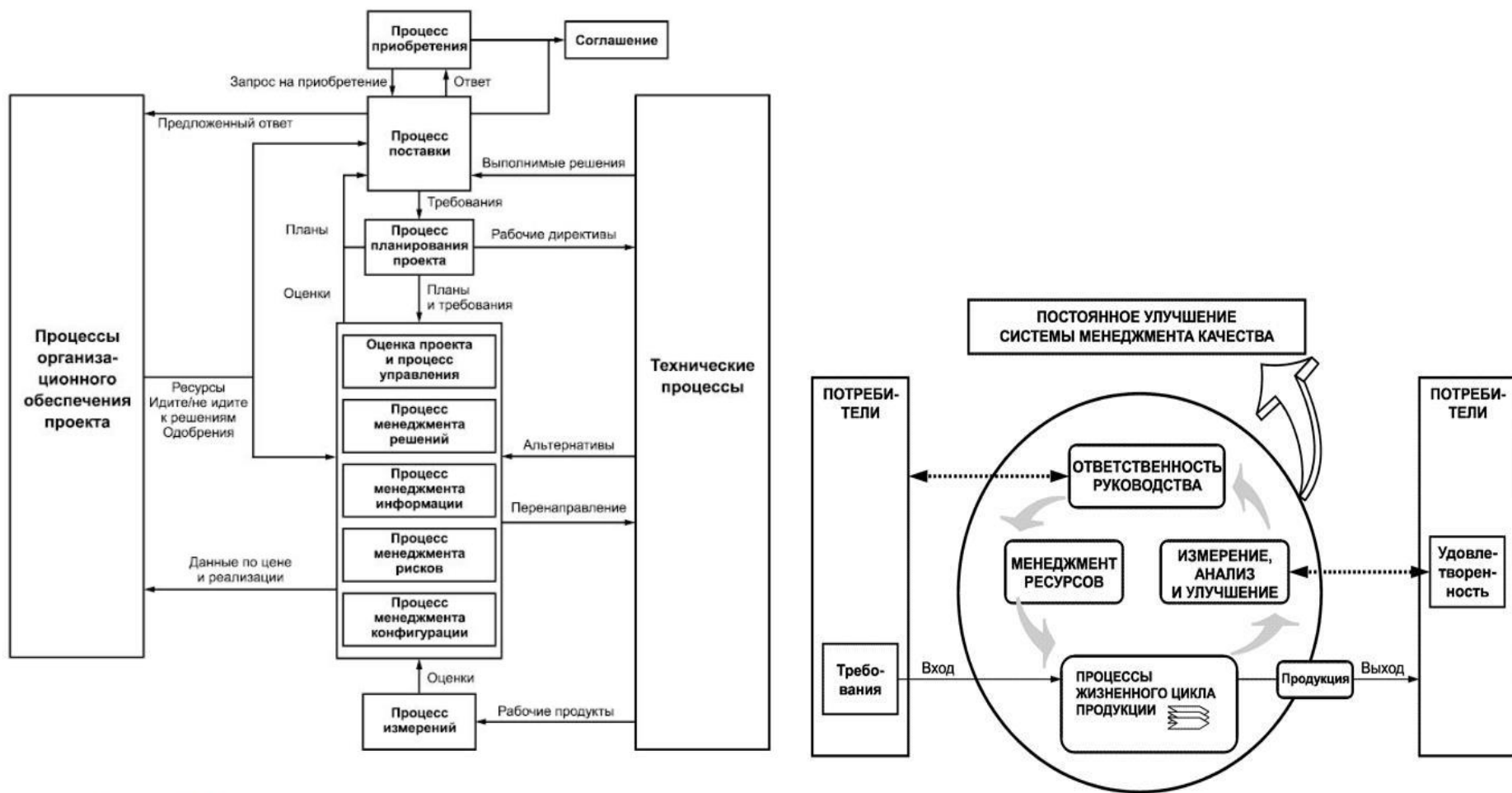
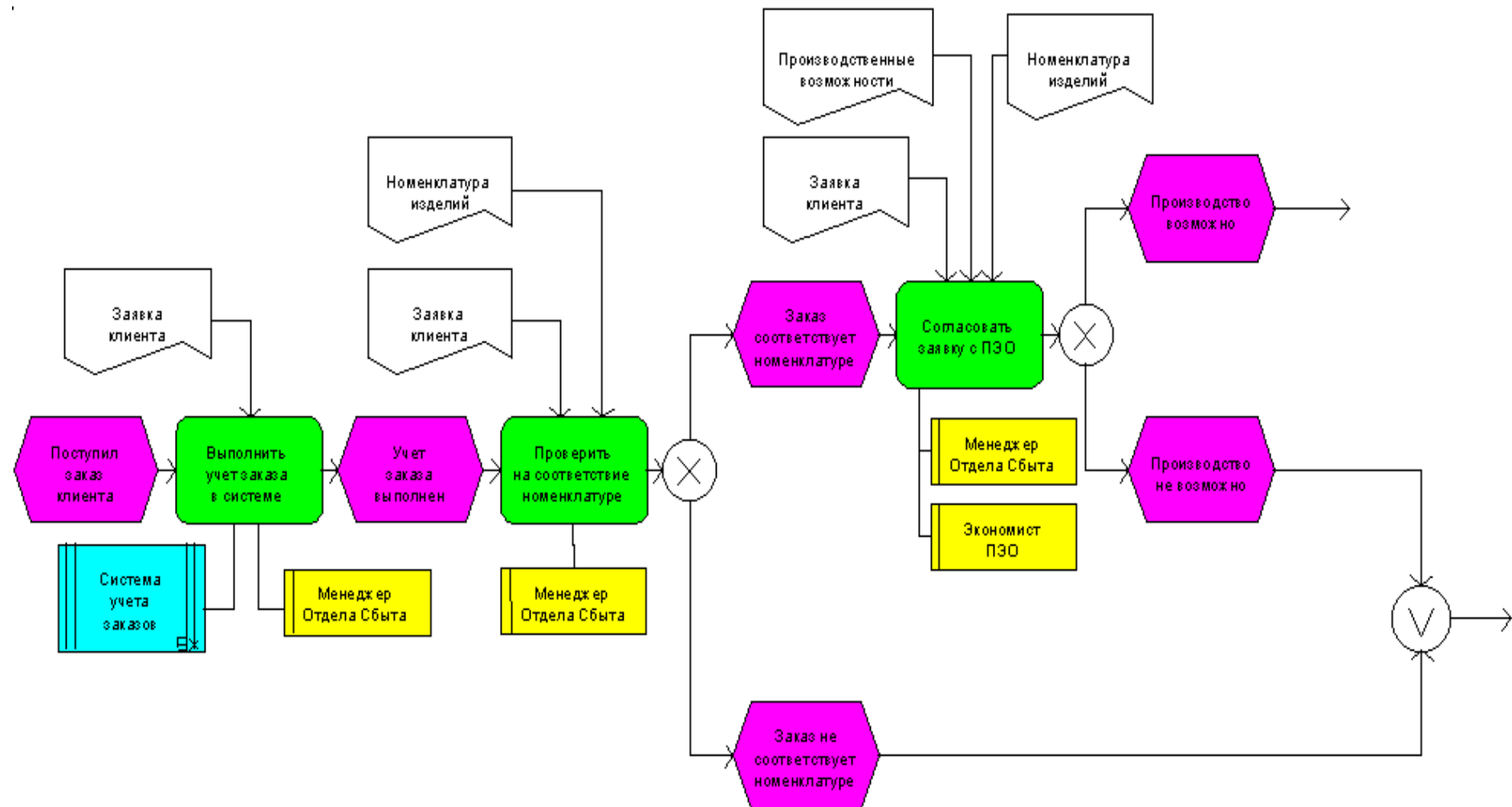


Рисунок 17 - Применение процесса к форме формального соглашения

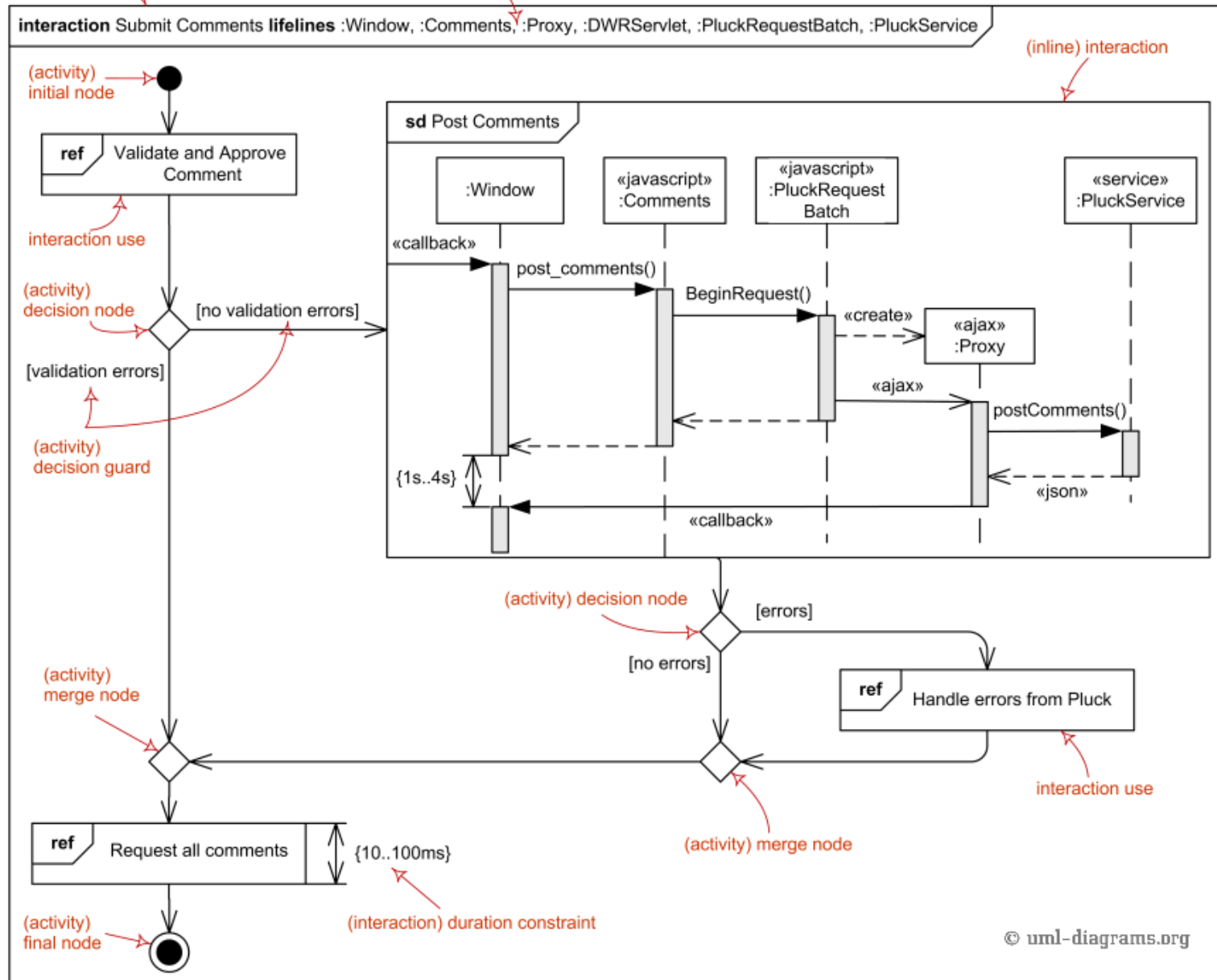
Сравнение: ARIS eEPC



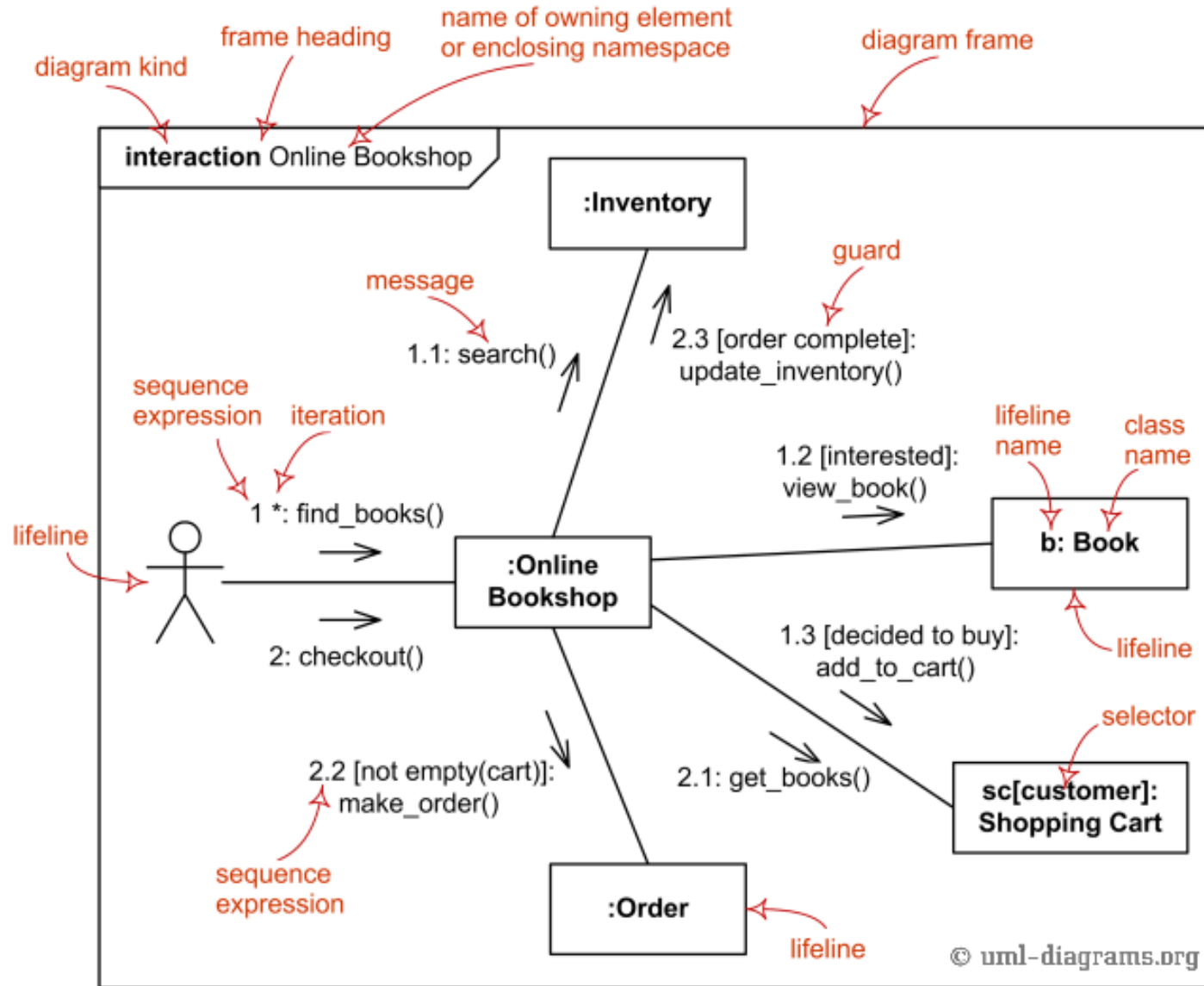
Сравнение: UML обзор взаимодействия

interaction diagram frame

lifelines participating in interaction



Сравнение: UML диаграмма коммуникаций



Сравнение: UML диаграмма потоков информации (диаграмма деятельности)

