

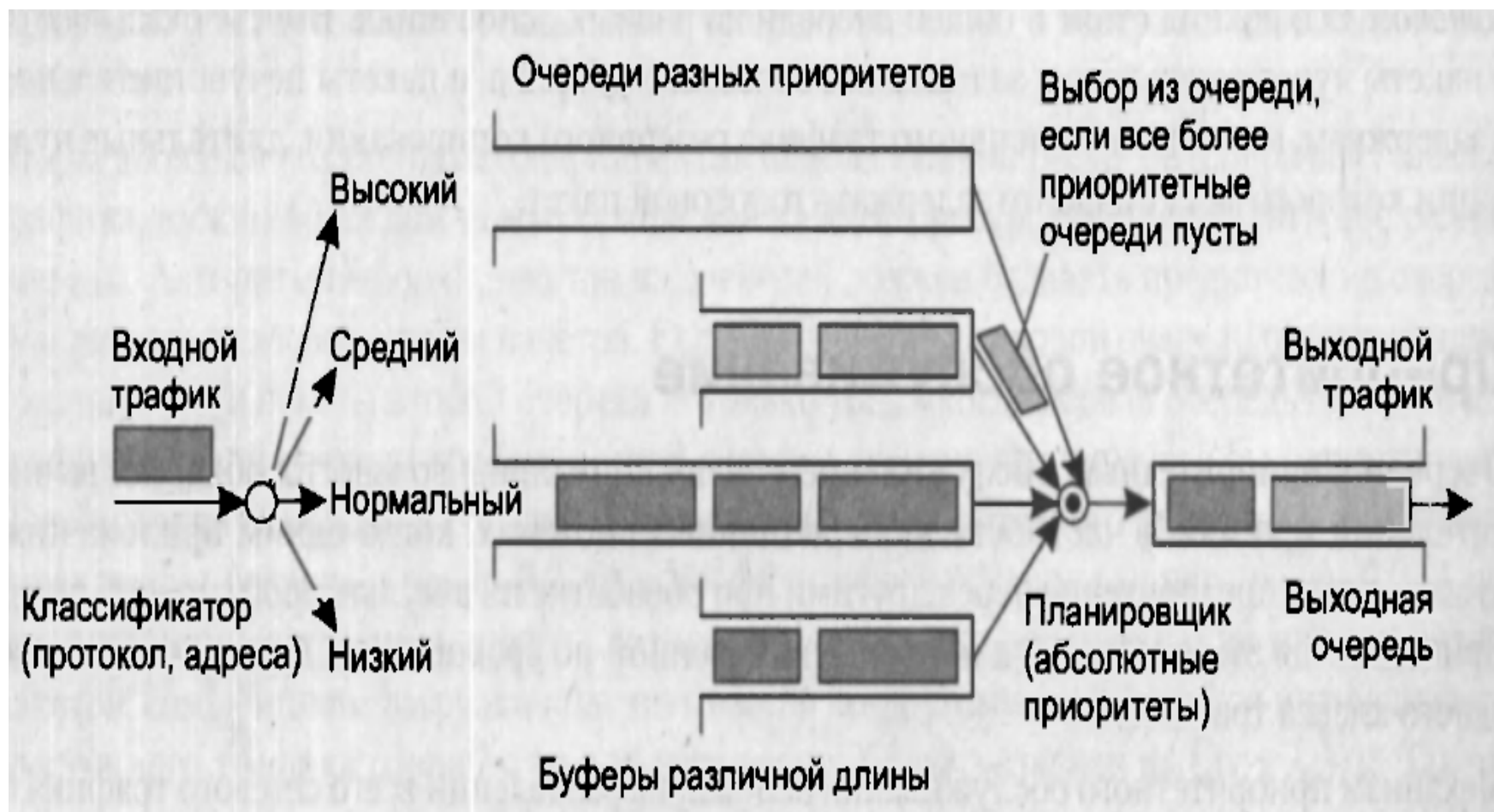
# Техника управления очередями

# Очередь FIFO

- В очереди FIFO в случае перегрузки все пакеты помещаются в одну общую очередь и выбираются из нее в том порядке, в котором поступили. Во всех устройствах с коммутацией пакетов алгоритм FIFO используется по умолчанию, так что такая очередь также обычно называется очередью «по умолчанию». Достоинствами - **простота реализации** и **отсутствие потребности в конфигурировании**.
- коренной недостаток — **невозможность дифференцированной обработки пакетов различных потоков**. Все пакеты стоят в общей очереди на равных основаниях.

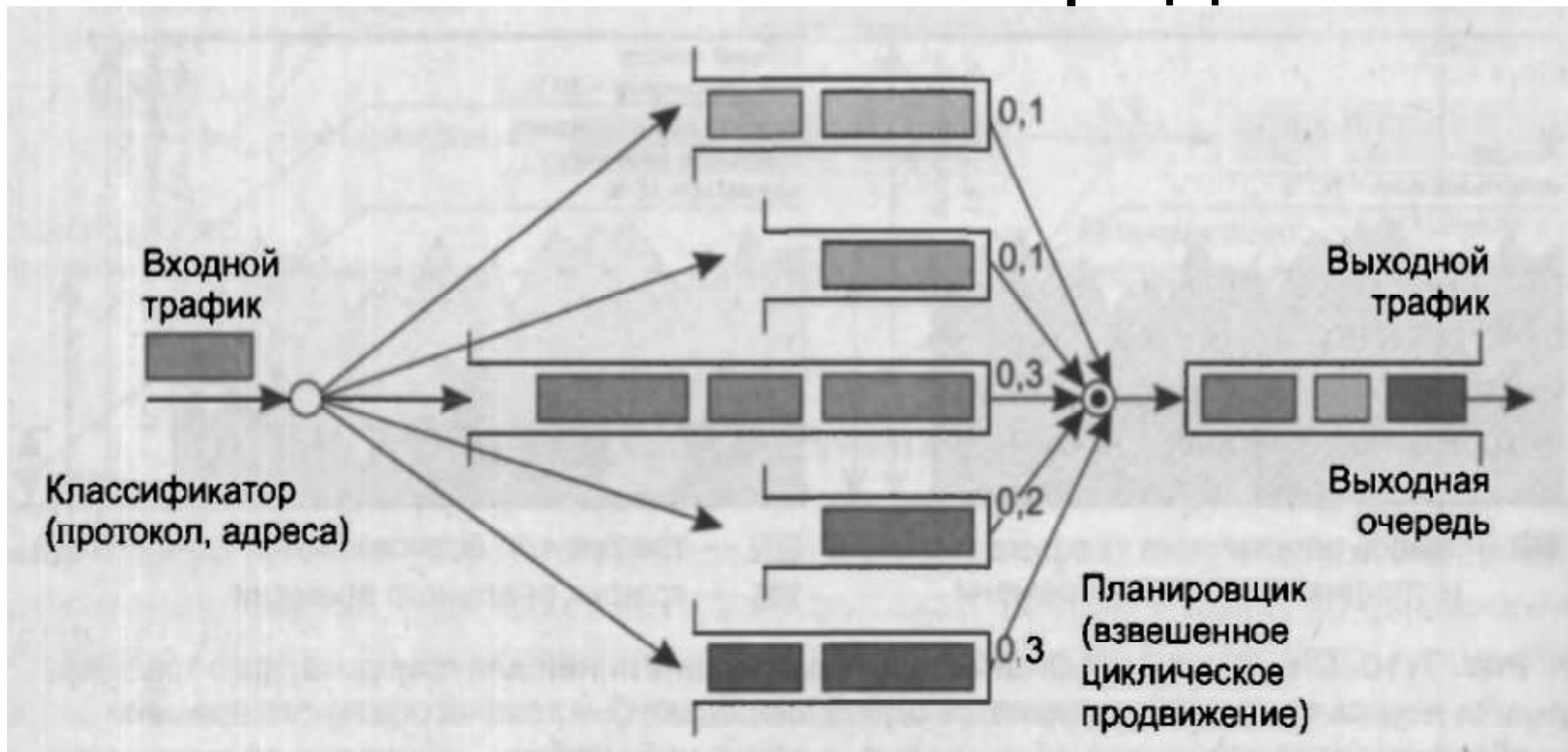
# Приоритетное обслуживание

- В сетевом устройстве, поддерживающем приоритетное обслуживание, **имеется несколько очередей** (буферов) — по одной для каждого приоритетного класса. Пакет, поступивший в период перегрузок, помещается в очередь, соответствующую его приоритетному классу.
- До тех пор пока из более приоритетной очереди не будут выбраны все имеющиеся в ней пакеты, устройство не переходит к обработке следующей менее приоритетной очереди.



- Трафик всех остальных приоритетных классов почти прозрачен для пакетов высшего приоритета. Слово «почти» относится к ситуации, когда высокоприоритетный пакет вынужден ждать завершения обслуживания низкоприоритетного пакета, если его приход совпадает по времени с началом продвижения низкоприоритетного пакета на выходной интерфейс. Если высокоприоритетный трафик высокоинтенсивен, то низкоприоритетные трафики практически замораживаются. Предполагается что высокоприоритетный трафик низкоинтенсивный. (например голос – 8-16 кбит/с)

# Взвешенные очереди



Каждый раз выбирается количество бит из очереди в соответствии с весом циклически по всем очередям. Например при скорости 100мбит/с, за сек из четвертой очереди будет получено 20мбит. Но обычно передача идет пакетами, потому необходимо получать пакеты, но тогда нужно оптимально выбрать время получения пакетов из одной очереди. Если слишком малое время, то плохо учитывается вес, если слишком большое то возникают простои.

# Обслуживание взвешенной очереди

- При росте коэффициента использования – отношения входного трафика к пропускной способности, **очередь растет для каждого вида трафика отдельно.**
- Как и для приоритетного обслуживания, при взвешенном обслуживании администратор может **назначать разным классам очередей буферы разных размеров.** Уменьшение размеров буферов для очередей ведет к росту числа потерь пакетов при перегрузках, но зато **снижает время ожидания для тех пакетов, которые не были отброшены и попали в очередь.**

# Справедливое обслуживание

- Еще одним вариантом взвешенного обслуживания является взвешенное справедливое обслуживание (Weighted Fair Queuing, WFQ). В случае подобного обслуживания пропускная способность ресурса делится между всеми потоками поровну, то есть «справедливо».
- Еще один вариант выбирать пакет из очереди с заданной вероятностью разыгрывая номер очереди. (почему-то метод не отмечен..)

# Смешанный режим с одной приоритетной очередью

- Существуют комбинированные алгоритмы обслуживания очередей. В наиболее популярном алгоритме подобного рода поддерживается одна приоритетная очередь, а остальные очереди обслуживаются в соответствии со взвешенным алгоритмом. Обычно приоритетная очередь используется для чувствительного к задержкам трафика, а остальные — для эластичного трафика нескольких классов. Каждый класс эластичного трафика получает некоторый минимум пропускной способности при перегрузках. Этот минимум вычисляется как процент от пропускной способности, оставшейся от приоритетного трафика.



# Механизмы кондиционирования трафика

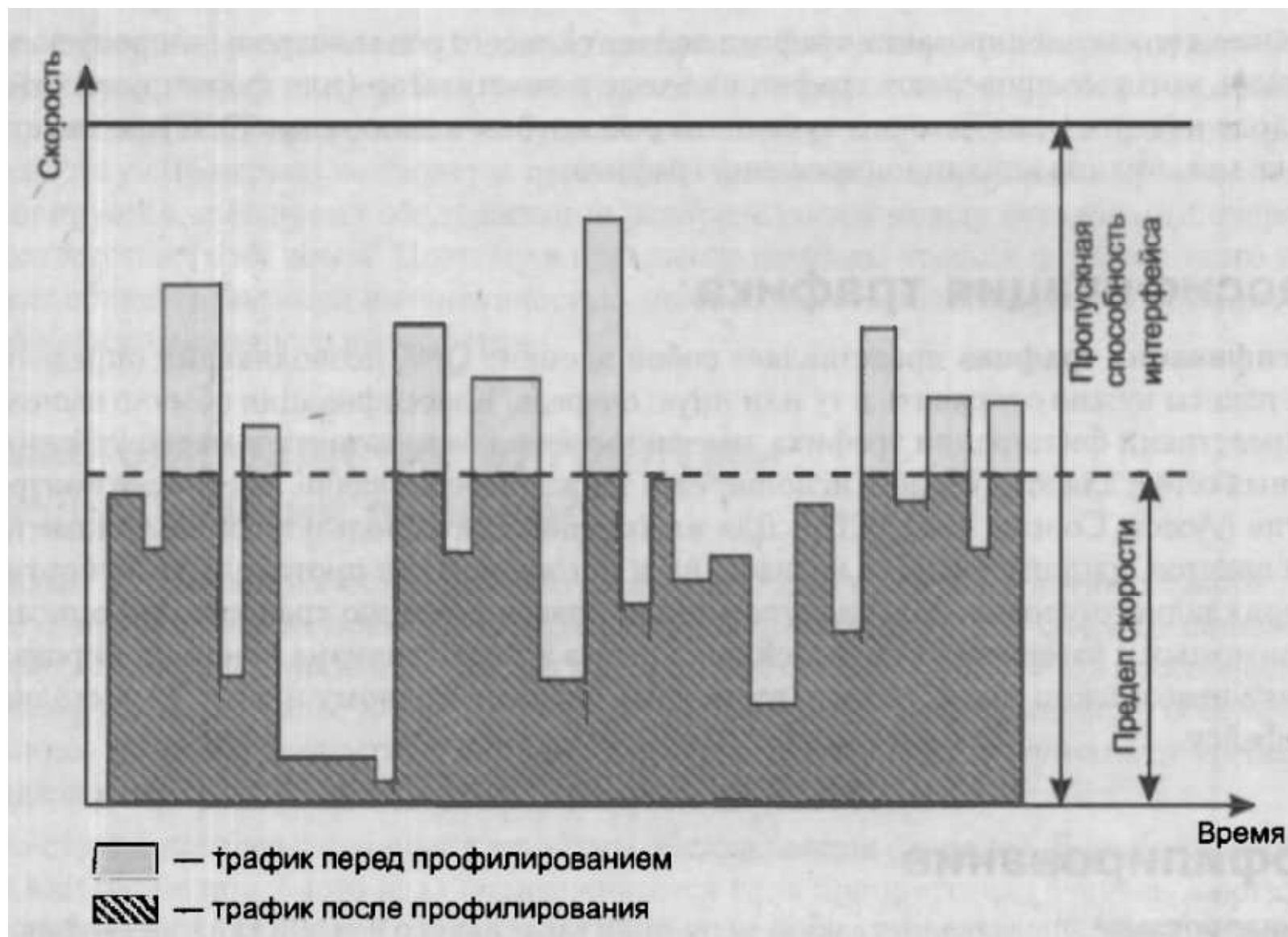
- **контролируют** текущие параметры потоков трафика, такие как его **средняя скорость** и **пульсация**.
- Очереди с различными алгоритмами обслуживания выделяют определенную долю пропускной способности некоторому потоку пакетов. Однако еще необходимо удержание скорости потока в определенных пределах с целью обеспечить желаемый коэффициент использования пропускной способности, которая выделена потоку с помощью некоторой очереди. Если коэф использ большой то и качество обслуживания низкое.
- Механизмы кондиционирования трафика являются своего рода контрольно-пропускными пунктами, которые проверяют трафик на входе в коммутатор (или формируют трафик на выходе из него). Существует несколько механизмов кондиционирования трафика (будут рассмотрены далее).

**Классификация трафика** представляет собой элемент QoS, позволяющий определить, какие пакеты нужно отправить в ту или иную очередь. Классификация обычно выполняется средствами фильтрации трафика, имеющимися в коммутаторах и маршрутизаторах.

- Для классификации используются различные **признаки пакетов**, например адреса назначения и источника, тип протокола транспортного или прикладного уровня.

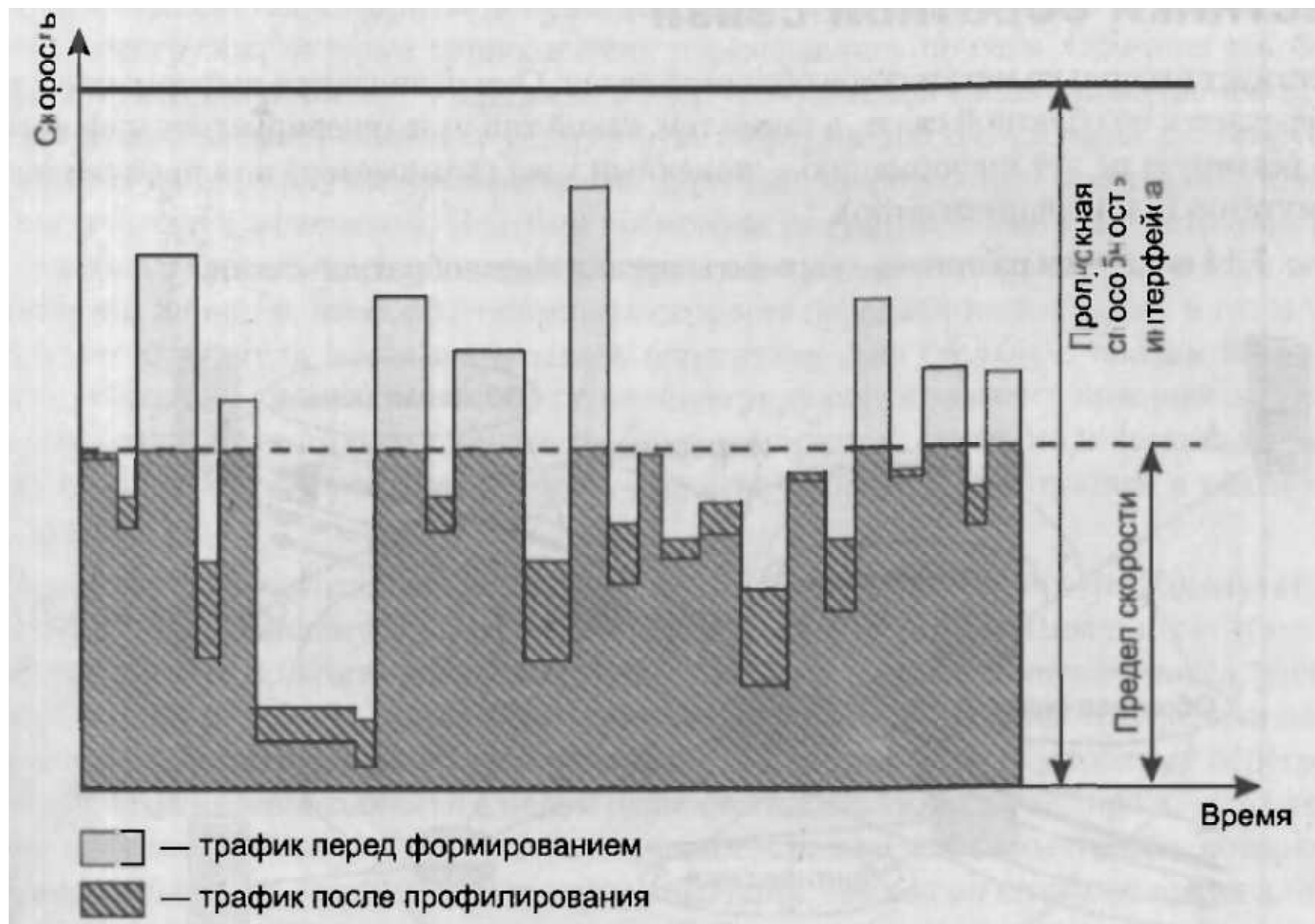
# Профилирование

- Профилирование представляет собой меру принудительного воздействия на трафик, которая служит для ограничения скорости потока пакетов. Профилирование обеспечивает соответствие потока пакетов заданному скоростному профилю; в качестве параметров профиля обычно выбирается средняя скорость потока пакета, измеренная на определенном интервале времени. **Пакеты, которые не укладываются в заданный профиль, либо отбрасываются, либо деквалифицируются**, то есть помещаются в класс обслуживания с более низкими привилегиями, например переводятся из приоритетного класса в стандартный.
- Профилирование обычно применяют для ограничения трафика, поступающего в приоритетную очередь, так как этот механизм является единственно возможным средством предотвращения ситуации вытеснения всего остального трафика приоритетным трафиком.



Эффект профилирования — отбрасывание избыточного трафика

- Формирование трафика — это **процесс намеренной задержки некоторых пакетов** из общего потока в целях удержания средней скорости трафика в некоторых заданных пределах.
- Формирование трафика напоминает профилирование, так как имеет схожую цель — ограничение скорости трафика, но достигается эта цель другим способом. Вместо того чтобы отбрасывать избыточные пакеты (то есть те, передача которых могла бы привести к превышению лимита скорости), механизм формирования трафика задерживает пакеты-нарушители так, что результирующая скорость оказывается в заданных пределах.



Эффект формирования трафика — сглаживание

# Обратная связь

- Механизм предотвращения перегрузки
- Коммутатор шлет узлам или коммутаторам обратно сигнал с просьбой уменьшить трафик если он перегружен или сигнал что он уже не перегружен, разрешение повысить трафик. Например, ICMP (подавление трафика).



Обратная связь между коммутаторами, узлом источником и приемником, коммутатором и узлом источником, узлами источником и приемником с промежуточным снижением скорости на коммутаторах, от промежуточного коммутатора до узла назначения и обратно до узла источника



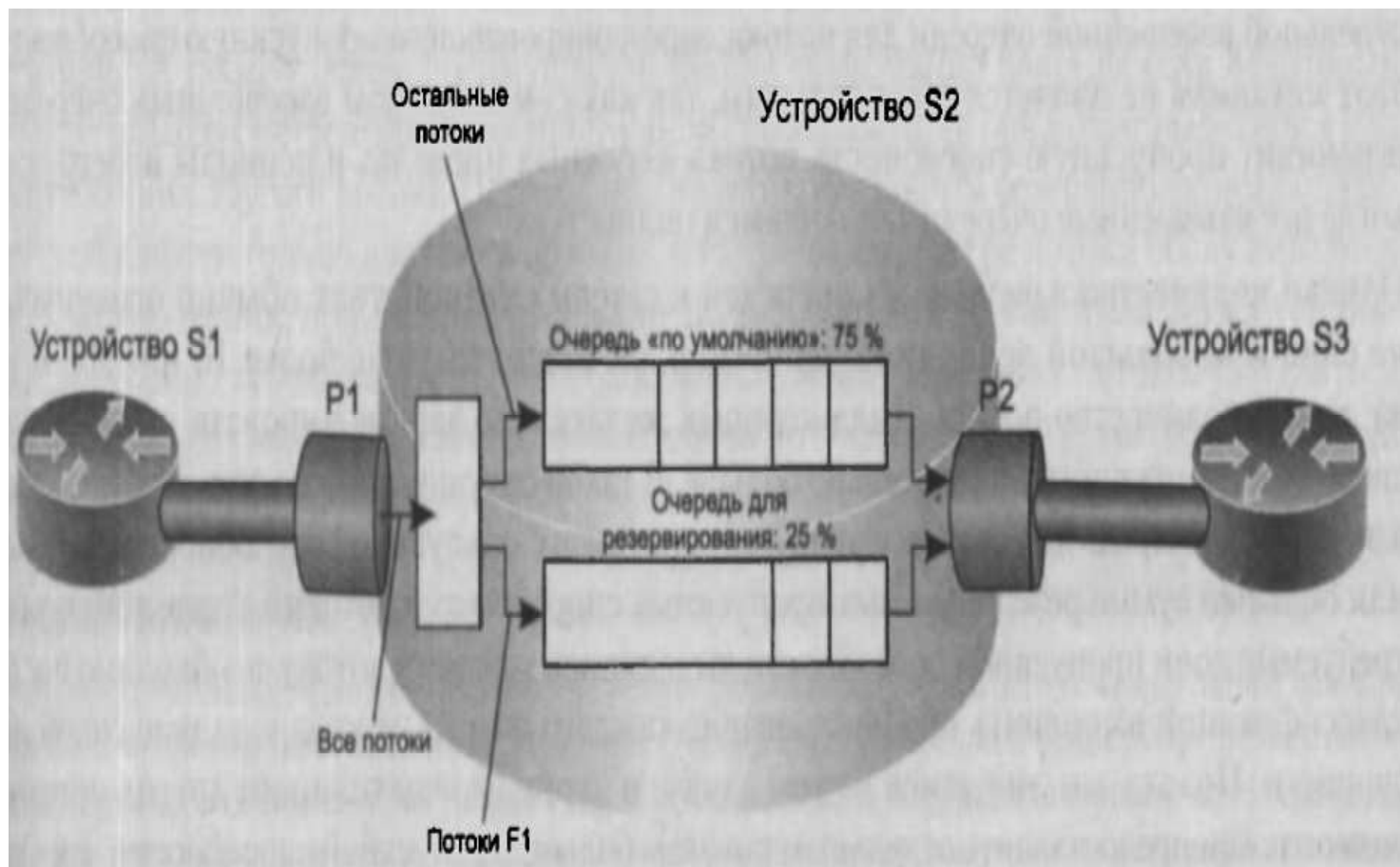
# Информация обратной связи

- В применяемых сегодня методах обратной связи используются следующие основные типы сообщений:
  - признак перегрузки (istr);
  - максимальная скорость передачи;
  - максимальный объем данных (кредит. скользящее окно (tcr));
  - косвенные признаки (признак – потеря пакета, tcr, управление перегрузкой за счет алгоритма медленного старта).

# Резервирование ресурсов, резервирование пропускной способности

Основная идея процедуры состоит в том, что каждому узлу сети вдоль маршрута следования потока задается вопрос, может ли этот узел обслужить некоторый новый поток с заданными характеристиками QoS, если известны предельные характеристики скорости потока, такие как средняя и пиковая скорости? Каждый узел при ответе на этот вопрос должен оценить свои возможности, то есть проверить, достаточно ли у него свободных ресурсов, чтобы принять на обслуживание новый поток и обслуживать его качественно. При положительном ответе узел должен некоторым образом резервировать часть своих ресурсов для данного потока

- Резервирование пропускной способности в пакетной сети «из конца в конец» начинается с операции, называемой **контролем допуска** в сеть потока, который просит зарезервировать для своего обслуживания некоторую пропускную способность сети между ее двумя конечными узлами. Эта операция состоит в проверке наличия доступной (то есть незарезервированной для других потоков) пропускной способности в каждом из узлов сети на протяжении всего маршрута следования потока. Очевидно, что максимальная средняя скорость потока должна быть меньше, чем запрашиваемая пропускная способность, иначе поток будет обслужен с очень плохим качеством даже несмотря на то, что ему была зарезервирована некоторая пропускная способность.



Выделение зарезервированной пропускной способности

Существуют три модели для реализации QoS:

Best-Effort (Негарантированная доставка);  
IntServ (Интегрированное обслуживание);  
DiffServ (Дифференцированные услуги).

**Best-Effort** — это модель сервиса по умолчанию, которая применяется к различным сетевым приложениям, таким как протокол передачи файлов (FTP) и электронная почта. Приложение может отправлять любое количество пакетов в любое время без уведомления сети. Затем сеть пытается передать пакеты. Модель **Best-Effort** подходит для услуг, которые предъявляют минимальные требования к задержке и скорости потери пакетов.

В модели **IntServ** приложение использует протокол сигнализации для уведомления сети о параметрах трафика и применения уровня QoS перед отправкой пакетов. Сеть резервирует ресурсы для приложения на основе параметров трафика. После того, как приложение получает подтверждение о зарезервированных ресурсах, оно начинает отправлять пакеты в пределах указанного диапазона.

Согласно RFC 1633, модель интегрированного обслуживания обеспечивает сквозное (End-to-End) качество обслуживания, гарантируя необходимую пропускную способность. **IntServ** использует для своих целей протокол резервирования сетевых ресурсов RSVP, который обеспечивает выполнение требований ко всем промежуточным узлам. В отношении IntServ часто используется термин «резервирование ресурсов» (Resource reservation).

**DiffServ** — часто используемая модель QoS, классифицирует пакеты в сети и выполняет установленные действия для каждого класса. Когда происходит перегрузка сети, пакеты классов обрабатываются на основе приоритетов, что приводит к неодинаковой частоте потери пакетов, неоднородным задержкам и дрожанию. Пакеты одного и того же класса объединяются и отправляются одним блоком.

Описана в RFC 2474 и RFC 2475. Обеспечивает QoS на основе распределения ресурсов в ядре сети и определенных классификаторов и ограничений на границе сети, комбинируемых с целью предоставления требуемых услуг. В этой модели вводится разделение трафика по классам, для каждого из которых определяется свой уровень QoS. **DiffServ** состоит из управления формированием трафика (классификация пакетов, маркировка, управление интенсивностью) и управления политикой (распределение ресурсов, политика отбрасывания пакетов). DiffServ является наиболее подходящим примером «умного» управления приоритетом трафика.

В отличие от IntServ, модель DiffServ не требует протокола сигнализации. В этой модели приложение не запрашивает сетевые ресурсы перед отправкой пакетов. Вместо этого приложение устанавливает параметры QoS в пакетах, через которые сеть может узнать требования QoS приложения. Сеть предоставляет дифференцированные услуги на основе параметров QoS каждого потока данных.



Другим методом оценки качества, особенно в IP-телефонии и IP-TV, является метрический метод, который отражает или предсказывает **субъективно опытное качество**. Используются субъективные оценки и показатели типа «**воспринимаемая пользователем производительность**», «**степень удовлетворения пользователя**», «**число счастливых клиентов**». Метод носит название Mean Opinion Score (MOS, «абсолютно субъективное понятие»), или Quality of Experience (QoE, «метод эмпирических оценок»).

В этом контексте QoS — совокупный эффект от удовлетворения заказчика сервисов, затрагивающих все виды обслуживания. Это определение оценивает соответствующее приложение в виде субъективной оценки и корректируется соответствующими коэффициентами  $\alpha$ : времени ответа, наличии и величине провалов, шумов, слышимых посторонних разговоров, уровня громкости, частотной характеристики, заметного эха и т. д., и также включает субъективную оценку обслуживания (служба поддержки).

Протоколы, которые предоставляют услугу QoS

IP Differentiated services ([DiffServ](#))

IP Integrated services ([IntServ](#))

Resource reSerVation Protocol (RSVP (протокол))

Multiprotocol Label Switching (MPLS)

RSVP-TE

Frame relay

X.25

Asynchronous Transfer Mode (ATM)

IEEE 802.1p

IEEE 802.1Q

IEEE 802.11e

IEEE 802.11p

HomePNA