

## 4. Обеспечение Qos при управлении потоком в узлах сети

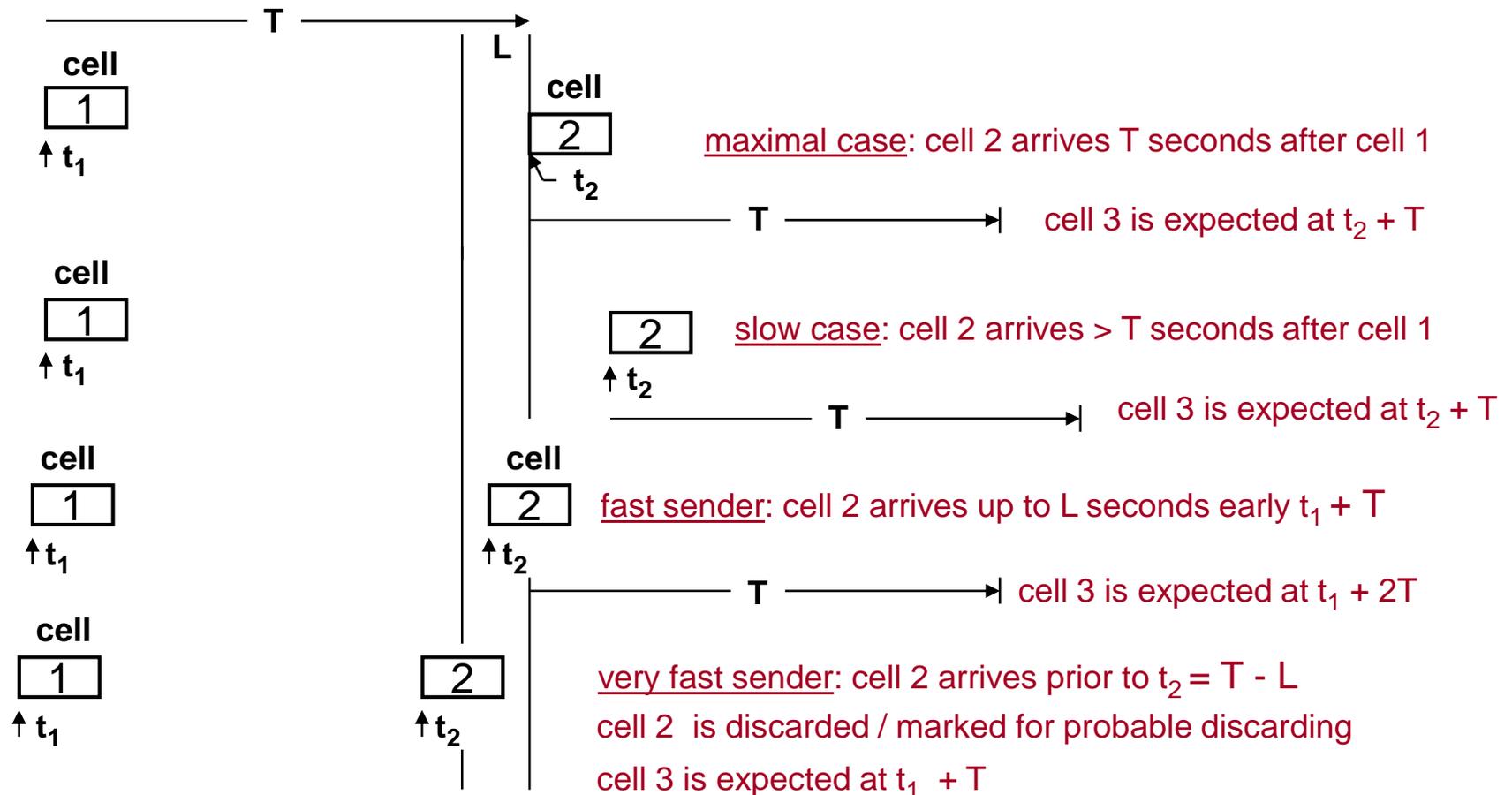
### 4.1 Профилирование

Выполняется:

анализ определенных параметров входящего потока,  
приведение этих параметров к условиям SLA.

Пример: ATM GFCA ( Generic Flow Control Algorithm).

Параметр трафика – PCR (Peak Cell Rate), параметр QoS - CDVT (Cell Delay Deviation Time),  
введенные параметры:  $T=1 / \text{PCR}$ ,  $L=\text{CDVT}$ .



## 4.2 Обслуживание очередей

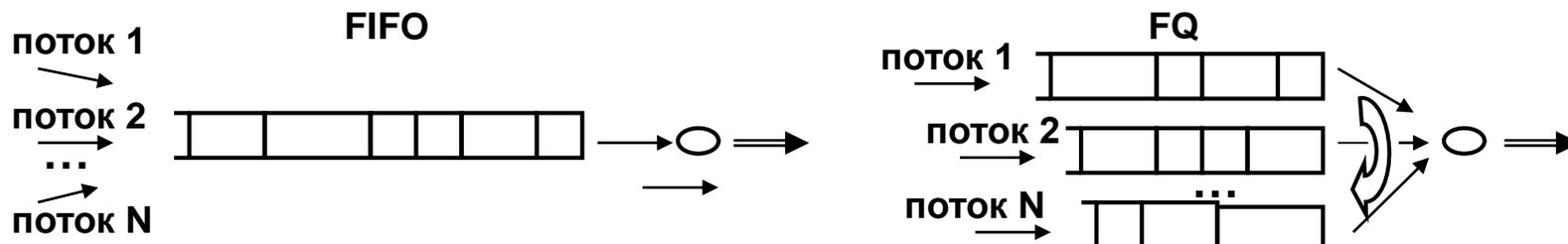
### 4.2.1 Дисциплины

Базовая проблема сетевого узла - передача нескольких потоков пакетов (разные источники, параметры SLA, важность и проч.) в одну выходную линию, связанная с обслуживанием очередей. Требование справедливости обслуживания по определенному критерию.

FIFO, общая очередь – простота, широкое применение, но

большая вероятность нарушения справедливости по большинству критериев (например, отличие задержки пакетов в очереди при вариантах «малые за большим», «большой за малыми»).

Справедливая очередность (Fair Queuing, FQ) - создание нескольких очередей соответствующих разным потокам. Критерий – обслуживание каждого потока в общем цикле.



Несправедливость – большой пакет получает большую долю пропускной способности линии.

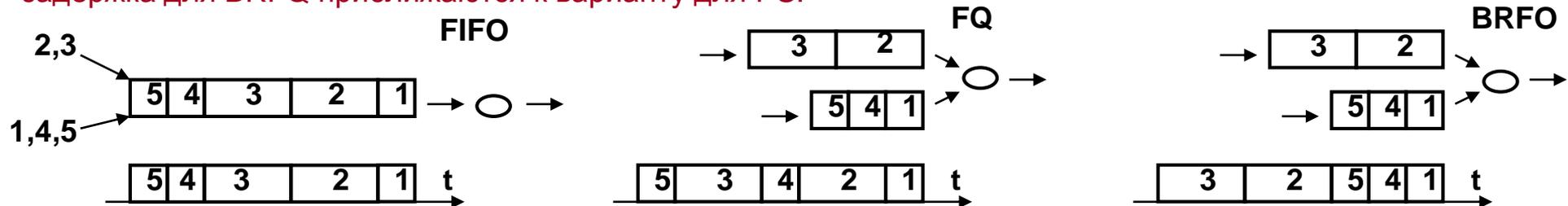
Справедливая очередность с по-битовым циклом (Bit-Round Fair Queuing, BRFQ)

Критерий – равная доля пропускной способности выходной линии для каждого потока.

*Идеальный подход (Processor Sharing, PS)* : в каждом цикле из каждой непустой очереди передается один бит головного пакета; если при по-битной передаче головного из потока приходит новый пакет, он включается в очередь

В варианте BRFQ из очередей выбираются не биты, а пакеты - когда заканчивается передача пакета, следующим выбирается тот, для которого сумма времени его прибытия в очередь и времени передачи минимальна, то есть, при выборке приоритет у застоявшихся коротких пакетов (короткие пакеты критичны к задержке)

Теоретически показано, что с увеличением времени наблюдения средняя пропускная способность и задержка для BRFQ приближаются к варианту для PS.



Взвешенная справедливая очередность (Weighted Fair Queuing, WFQ)

Критерий – обслуживание очередей согласно заданному приоритету.

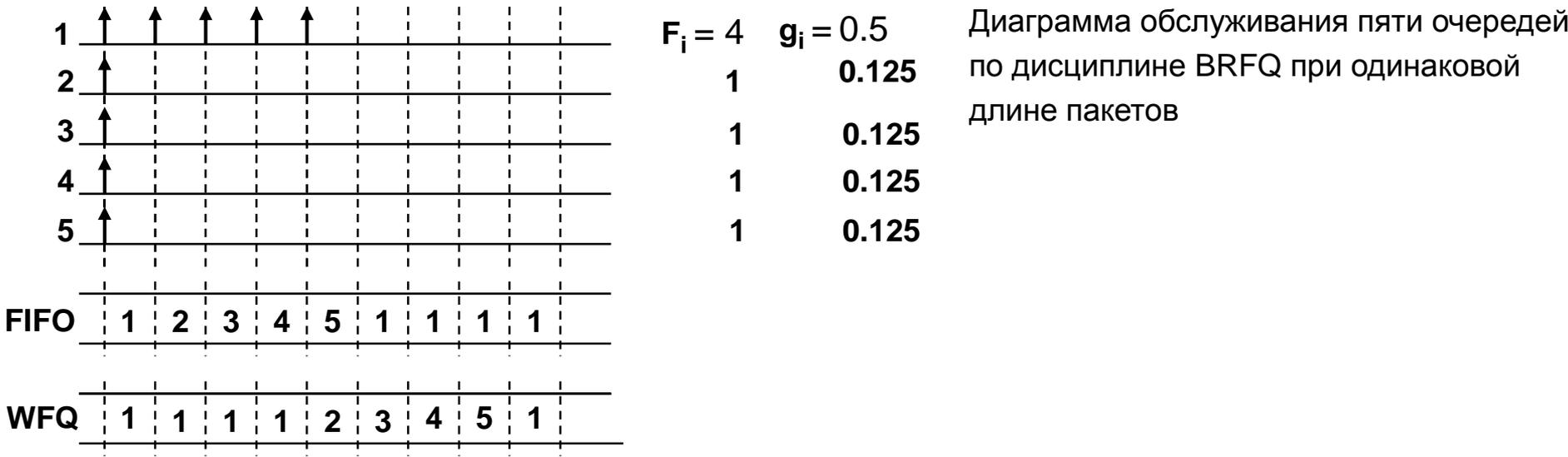
Идеальный подход: в каждом цикле из непустой очереди I выбирается  $F_i$  битов – это вес потока.

Таким образом, скорость обслуживания потока I равна

$$g_i = (F_i / \sum_j F_j) \times C, \text{ где } C \text{ – пропускная способность выхода.}$$

Принцип обобщенного разделения процессора (Generalized Processor Sharing, GPS) заключается в том, что, если для потока требуется обслуживание со скоростью  $g_i$ , то при наличии ресурсов для него назначается соответствующий вес.

В варианте WFQ эмулируется GPS, но вместо битов передаются пакеты. Выбор из очередей выполняется аналогично BRFQ, но при оценке времени передачи берется в расчет скорость  $g_i$ .



Приоритетная очередность

Очередь с высшим приоритетом получает гарантированную часть пропускной способности, остальные – по остаточному принципу.



## 4.3 Формирование потока

### 4.3.1 Математическое описание обслуживания потока заявок



Принципиальная схема процессов, определяющих уровень QoS в сетевом узле коммутации: поступление потока заявок (пакетов) на обработку обслуживающим устройством (процедурой узла) с вероятностью образования входной очереди. Интервалы поступления и обслуживания характеризуются случайными величинами

#### Параметры процессов

#### Функция распределения вероятности

$F(x) = \Pr [X \leq x]$   $F(-\infty) = 0$ ,  $F(\infty) = 1$ , где  $X$  - измеряемая величина,  $x$  – число,

#### Функция плотности вероятности

$f(x) = F'(x)$

#### Среднее (ожидаемое) значение

$$E[X] = \mu_x = \int_{-\infty}^{\infty} x f(x) dx$$

#### Дисперсия

$$\text{Var}[X] = E [ | X - \mu_x |^2 ]$$

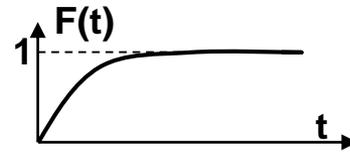
#### Среднеквадратичное отклонение

$$\sigma_x = \sqrt{\text{Var}[X]}$$

## Описание процесса в нотации Кенделла M/M/1

Нотация Кенделла –  $X / Y / N$ , где  $X$  – закон распределения интервалов между поступлением заявок,  
 $Y$  – закон распределения интервалов обслуживания,  
 $N$  – количество серверов.

Типично для процессов поступления и обслуживания заявок в коммуникационных схемах:  
экспоненциальная зависимость вероятности события от времени (M), один сервер, нотация M/M/1.



### *Поток заявок*

$$F(t) = \Pr [t_a \leq t] = 1 - e^{-\lambda t}, \text{ где}$$

$t_a$  – интервал между заявками,

$\lambda$  – интенсивность поступления заявок

(заявок / с), равная  $1 / T_a$ ,

где  $T_a$  - средний (ожидаемый) интервал  
поступления заявок,

$\sigma_a$  - среднее квадратичное отклонение  
интервала поступления.

При нотации M/M/1 справедливо  $\sigma_a = T_a$

$\rho = \lambda / \mu$  - коэффициент использования  
обслуживающего устройства,

$T_w = (\rho T_s) / (1 - \rho)$  – среднее время ожидания в очереди,

$T_r = T_w + T_s$  - среднее время ожидания в системе.

### *Процесс обслуживание*

$$F(t) = \Pr [t_s \leq t] = 1 - e^{-\mu t}, \text{ где}$$

$t_s$  – интервал обслуживания,

$\mu$  - интенсивность обслуживания,

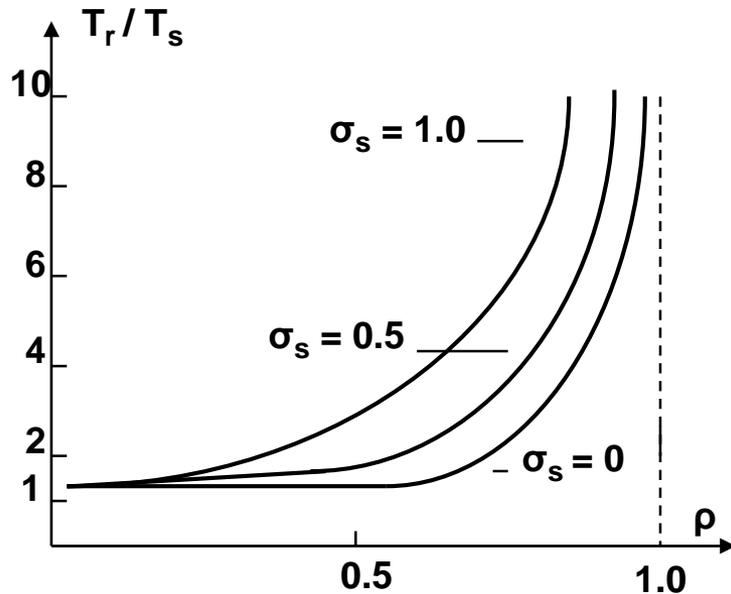
(обслуживаний/с), равная  $1 / T_s$ ,

где  $T_s$  - среднее (ожидаемое) время  
обслуживания,

$\sigma_s$  - среднее квадратичное отклонение  
интервала обслуживания.

При нотации M/M/1 справедливо  $\sigma_s = T_s$

## Критические закономерности



Эффект резкого возрастания задержки при  $\rho = 1$  ( $T_a = T_s$ ).

Ввиду вероятностного характера процесса, интервал поступления заявок может уменьшаться относительно ожидаемого значения, а интервал обслуживания – увеличиваться. Результат – образование очереди и возрастание задержки.

При  $\rho \sim 0.5$  запас по возможности обслуживания позволяет сгладить этот эффект.

## Выводы:

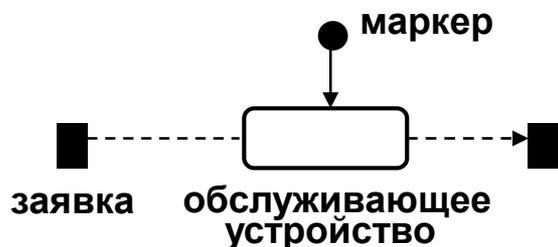
Минимизация задержки обеспечивается при:

- фиксированном времени обслуживания ( $\sigma_s = 0$ ).
- существенно большем интервале поступления заявок по сравнению с интервалом обслуживания ( $\rho \leq 0.5$ ).

Это означает, что обслуживающее устройство в состоянии обработать всплеск нагрузки, обусловленный вероятностным характером процесса.

### 4.3.2 Абстрактная модель процесса

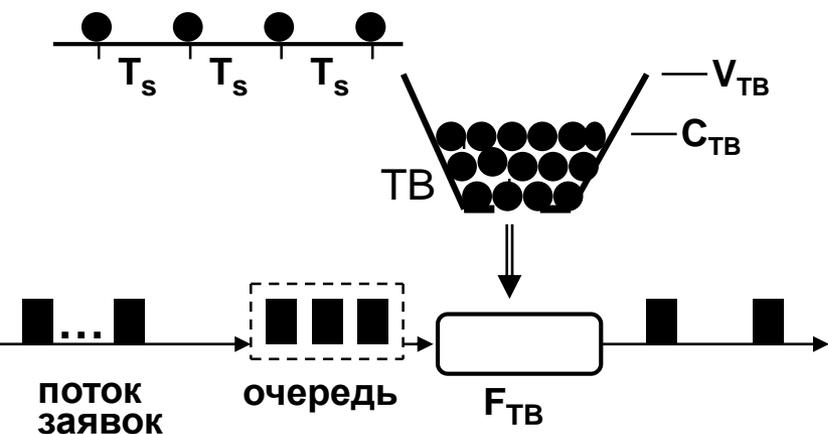
#### Использование маркеров



Обслуживающее устройство по отношению к заявке работает как клапан, открывающийся при поступлении маркера.

При  $\rho < 1$  частота поступления маркеров выше частоты поступления заявок, поэтому необходимо ведро (bucket), сохраняющее неиспользованные маркеры.

#### Модель token bucket, TB



TB – (token bucket) – маркерное ведро

$F_{TB}$  – сервер маркерного ведра

$V_{TB}$  – объем ведра (в маркерах)

$C_{TB}$  – уровень наполнения ведра (в маркерах)

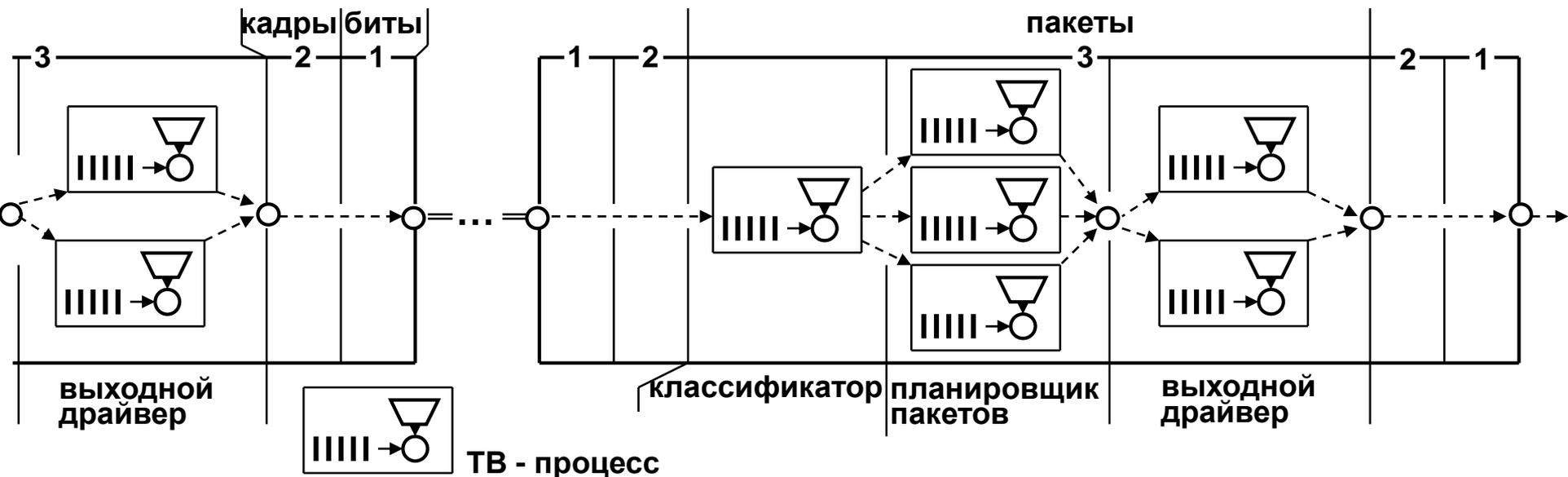
$T_s$  – интервал обслуживания при  $\sigma_s = 0$

Выбор интервала  $T_s$  позволяет задать нужное значение  $\rho$  при известном  $T_a$ .

Маркер в ведре – неиспользованная возможность обслуживания одной заявки.

Уровень наполнения ведра  $C_{TB}$  определяет возможность  $F_{TB}$  обслуживать группу заявок, это запас по обработке всплесков нагрузки. Максимальному запасу соответствует значение  $V_{TB}$ .

## 4.4 Транзит пакетов в узлах коммутации



**Функциональные уровни в узлах – физический (1), звена данных (2), сетевой (3).**

**Классификатор – разделение общего входного трафика на потоки разных классов, заданных в трафик- контрактах.**

**Планировщик пакетов – обслуживание пакетов в потоках, выделенных классификатором: профилирование / маршрутизация / передача в выходной драйвер.**

**Выходной драйвер – передача пакетов из очередей в выходной интерфейс узла при наличии необходимой пропускной способности линии.**

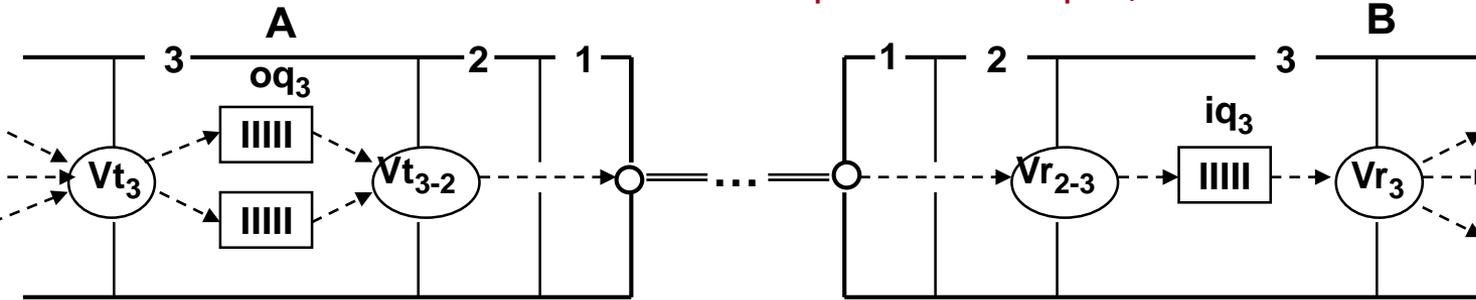
**Маркер в ТВ-процессах узла – часть ресурсов (производительность, память процессора или пропускная способность линии), необходимая для продвижения одного пакета.**

**Проблема реализации: как поддерживать необходимый для данного потока объем ведра и уровень наполнения ( в терминах ресурсов).**

**Решения: резервирование ресурсов для нового трафика и динамическое «наполнение ведра» за счет ресурсов, освобождающихся после обработки пакетов.**

## 5. Перегрузка (congestion)

### 5.1 Проявление процесса



Входная очередь  $iq_3$  в узле B образуется, если скорость приема  $Vr_{2-3}$  выше скорости  $Vr_3$ , то есть, процессы в зоне 3 не успевают обрабатывать входной поток.

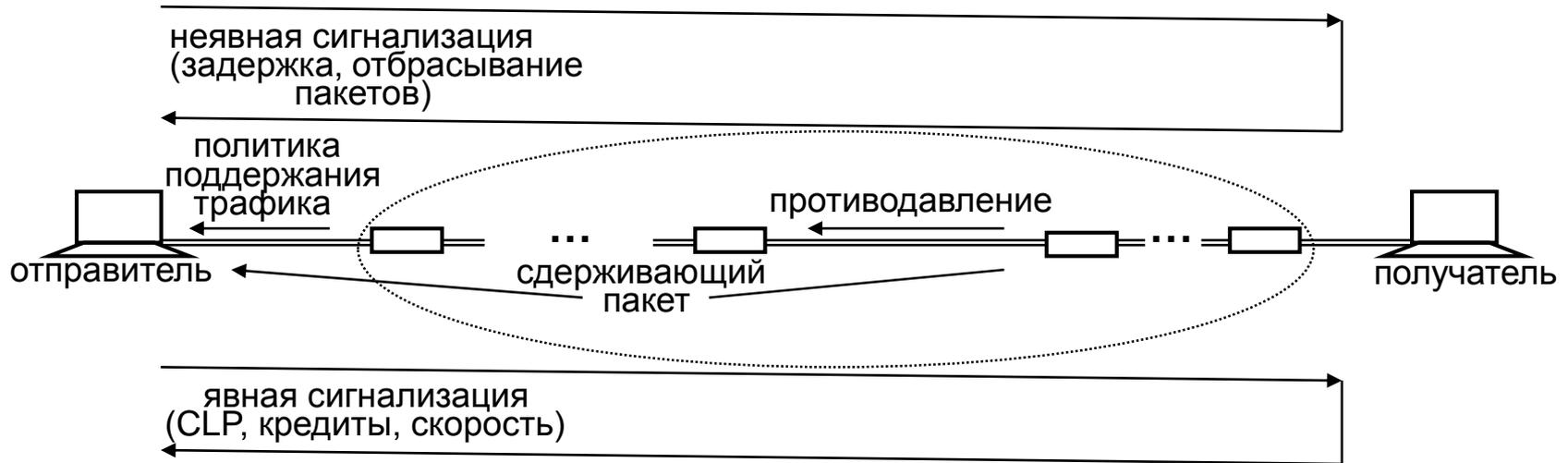
Если условие кратковременно, то очередь исчезает, но при ее критическом увеличении, ведущем к образованию «пробки» в B, следует сократить или остановить входной поток, например, уменьшить кредит, выдаваемый от СЕ в предыдущем примере.

При этом возможно, что суммарная скорость  $Vt_3$  превысит суммарную скорость  $Vt_{3-2}$  и начнет расти очередь  $oq_3$  в узле A. Если это условие сохраняется, то пробка перемещается в A (при сохранении пробки в B).

Увеличение допустимой длины очереди смягчает процесс, но приводит к задержке и фиксации тайм-аута, когда отправитель должен послать дубликат пакета. Это может принять лавинный характер, распространяющий эффект перегрузки на тракты, фрагменты сети или всю сеть.



## 5.2 Методы борьбы с перегрузкой



### Противодавление (backpressure)

Ограничение потока между соседними узлами

(сети, допускающие управление потоком в ретрансляционных участках, например, X.25).

### Сдерживающий пакет (choke packet)

Узел сети посылает его отправителю как знак отбрасывания очередного пакета или переполнения очередей – при этом прием пакетов продолжается.

В Internet это одна из функций протокола ICMP (Internet Control Message Protocol).

### Неявная сигнализация о перегрузке

Увеличение времени задержки пакетов или потеря пакетов, сообщаемые приемником отправителю.

( В Internet выполняется на уровне TCP).

### Политика поддержания трафика

Регулирование интервала передачи подтверждений (на уровне TCP).

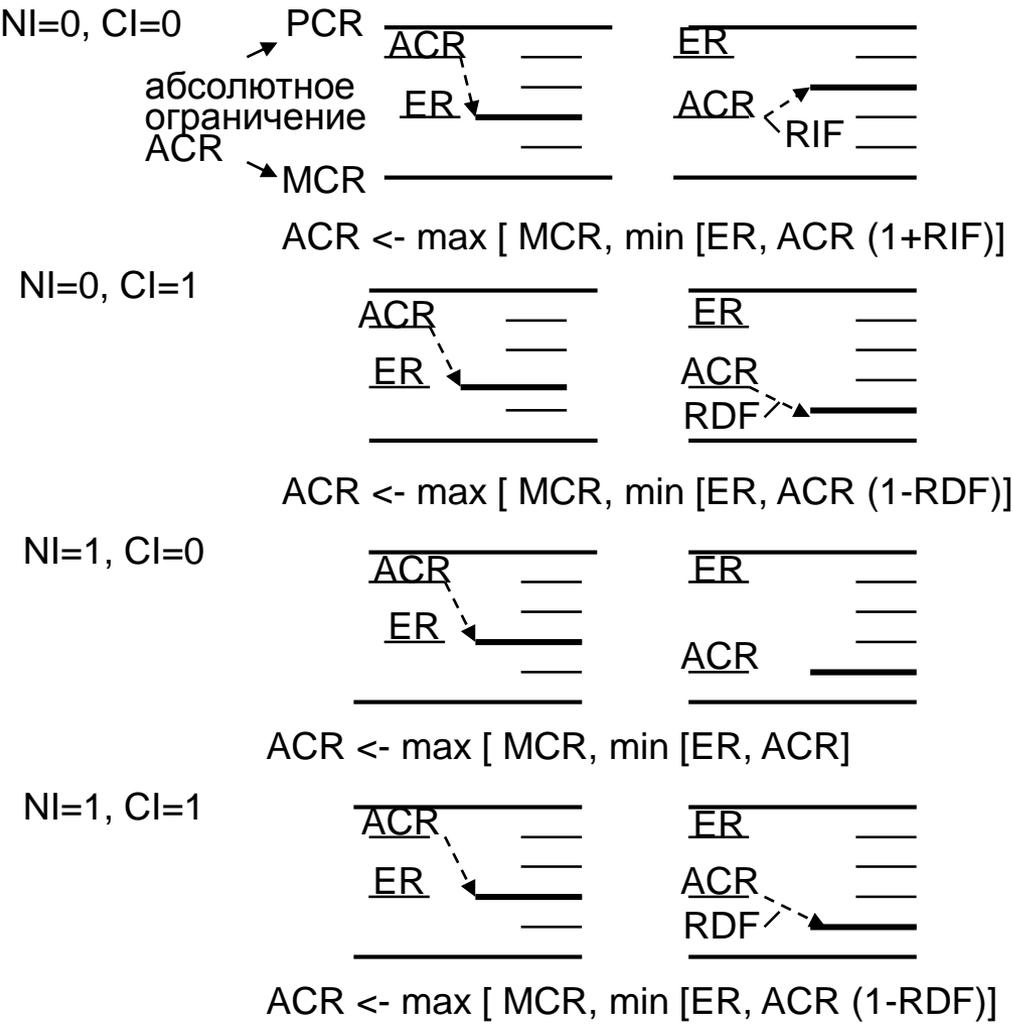
### Явная сигнализация о перегрузке

Сеть предупреждает оконечные системы о растущей нагрузке, они принимают меры для ее снижения.

Технологии: CLP / кредитование / forward (.сбщ. к получателю) и backward (сбщ. к отправителю).

## 5.3 Пример управления трафиком в режиме ABR (технология ATM)

### Процедура в Terminal Equipment (TE) при получении ячейки Resource Management (RM)



### Параметры трафика в TE:

Allowed Cell Rate, ACR – текущая скорость, с которой источник может передавать ячейки

Minimum Cell Rate, MCR – минимальное значение, которое может принять ACR (меньшая скорость допустима).

Peak Cell Rate, PCR – макс. значение, которое может принять ACR.

Initial Cell Rate, ICR – нач. значение ACR.

Rate Increase Factor, RIF – коэффициент увеличения скорости.

Rate Decrease Factor, RDF – коэффициент уменьшения скорости.

### Управление ресурсами передатчика через ячейки Resource Management (RM), поля RM:

Congestion Indication, CI

No Increase, NI

Explicit Cell Rate, ER – запрос источника

