Traffic Policing/Marking

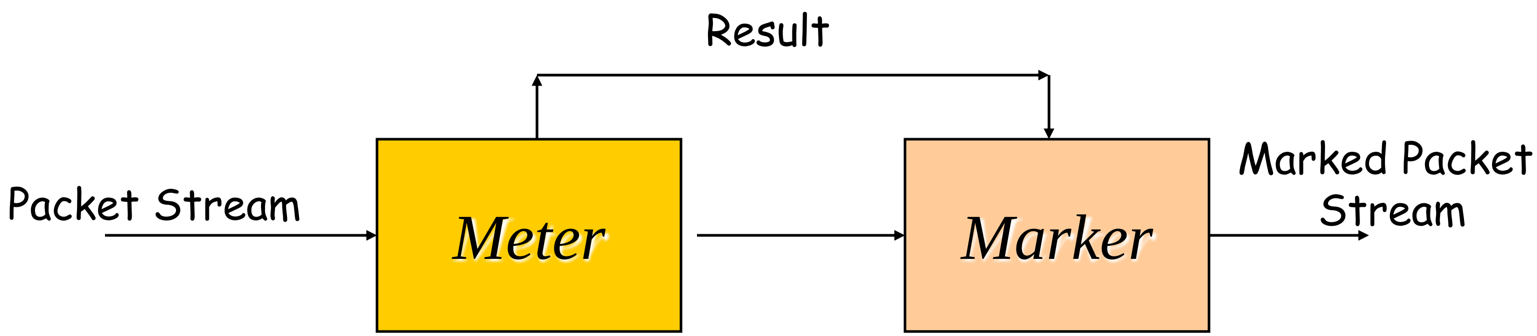
Outline

- What is scheduling
- Why we need it
- Requirements of a scheduling discipline
- Fundamental choices
- Scheduling best effort connections
- Scheduling guaranteed-service connections
- Packet drop strategies
- Traffic Policing

Traffic Policing/Traffic Marking

- Traffic Policing/Marking is used to check for the parameter conformance
  - Traffic is admitted based on the negotiated contract
  - Edge switches/routers check whether the user is sending the traffic as per the traffic contract or exceeding by policers/markers
  - Policing/Marking function tags packets as conforming or violating
- IETF proposed rate and color markers
  - RFC 2697 A Single Rate Three Color Marker. J. Heinanen, R. Guerin. September 1999.
  - RFC 2698 A Two Rate Three Color Marker. J. Heinanen, R. Guerin. September 1999.
- ATM uses Generic Cell Rate Algorithm (GCRA) to police traffic
  - Based on Leaky Bucket
  - Can be implemented using a Virtual Scheduling algorithm

IETF Markers



- Two Markers (Policers) proposed by IETF
  - RFC 2697 A Single Rate Three Color Marker (srTCM)
  - RFC 2698 A Two Rate Three Color Marker (trTCM)
- Meters operate in one of the two modes:
  - *Color-Blind mode:* Meter assumes that the packet stream is uncolored.
  - *Color-Aware mode:* Meter assumes that some preceding entity has pre-colored the incoming packet stream so that each packet is either green, yellow, or red
- Marker (re)colors an IP packet according to the results of the Meter.
- The color is coded in the DS field of the packet in a PHB specific manner

IETF Markers (srTCM)

- RFC 2697: srTCM (Single Rate Three Color Marker)
- The Single Rate Three Color Marker (srTCM) meters an IP packet stream and marks its packets either green, yellow, or red.
- Marking is based on a Committed Information Rate (CIR) and two associated burst sizes, a Committed Burst Size (CBS) and an Excess Burst Size (EBS).
- A packet is marked green if it doesn't exceed the CBS, yellow if it does exceed the CBS, but not the EBS, and red otherwise.
- The srTCM is useful, for example, for ingress policing of a service, where only the length, not the peak rate, of the burst determines service eligibility.
- The CIR is measured in bytes of IP packets per second, i.e., it includes the IP header, but not link specific headers.
- The CBS and the EBS are measured in bytes.

srTCM (continued)

- The behavior of the meter is specified in terms of its mode and two token buckets, C and E, which both share the common rate CIR. The maximum size of the token bucket C is CBS and the maximum size of the token bucket E is EBS.
- The token buckets C and E are initially (at time 0) full, i.e., the token count Tc(0) = CBS and the token count Te(0) = EBS. *Thereafter, the token counts Tc and Te are updated CIR times per second as follows*:
  - **If** Tc is less than CBS, Tc is incremented by one, **else**
  - **if** Te is less then EBS, Te is incremented by one, **else**
  - neither Tc nor Te is incremented.

srTCM (continued ..)

- When a packet of size B bytes arrives at time t and if the srTCM is configured to operate in the *Color-Blind mode*:
  - If Tc(t)-B >= 0, the packet is **green** and Tc is decremented by B down to the minimum value of 0, else
  - if Te(t)-B >= 0, the packets is **yellow** and Te is decremented by B down to the minimum value of 0, else
  - the packet is **red** and neither Tc nor Te is decremented.

srTCM (continued ..)

- When a packet of size B bytes arrives at time t and if the srTCM is configured to operate in the *Color-Aware mode*:
  - **If** the packet has been **precolored as green** and Tc(t)-B >= 0, the packet is **green** and Tc is decremented by B down to the minimum value of 0, **else**
  - **If** the packet has been **precolored as green or yellow** and if Te(t)-B >= 0, the packets is **yellow** and Te is decremented by B down to the minimum value of 0, **else**
  - the packet is **red** and neither Tc nor Te is decremented.
  - Note that according to the above rules, marking of a packet with a given color requires that there be enough tokens of that color to accommodate the entire packet.

IETF Markers (trTCM)

- RFC 2698: trTCM (Two Rate Three Color Marker)
- The Two Rate Three Color Marker (trTCM) meters an IP packet stream and marks its packets either green, yellow, or red based on two rates:
  - **Peak Information Rate (PIR)** and **Committed Information Rate (CIR),** and their associated burst sizes
- A packet is marked **red** if it exceeds the Peak Information Rate (PIR).
- Otherwise it is marked either **yellow** or **green** depending on whether it exceeds or doesn't exceed the Committed Information Rate (CIR).
- The trTCM is useful, for ingress policing of a service, where a peak rate needs to be enforced separately from a committed rate

trTCM (continued)

- The trTCM is configured by setting its mode and by assigning values to four traffic parameters:
  - a Peak Information Rate (PIR) and its associated Peak Burst Size (PBS) and
  - a Committed Information Rate (CIR) and its associated Committed Burst Size (CBS).
- The PIR and CIR are measured in bytes of IP packets per second, i.e., it includes the IP header, but not link specific headers.
- The PIR must be equal to or greater than the CIR.
- The PBS and the CBS and are measured in bytes and both of them must be configured to be greater than 0.
- It is recommended that they be configured to be equal to or greater than the size of the largest possible IP packet in the stream.

trTCM (continued ..)

- The behavior of the Meter is specified in terms of its mode and two token buckets, P and C, with rates PIR and CIR, respectively.
- The maximum size of the token bucket P is PBS and the maximum size of the token bucket C is CBS.
- The token buckets P and C are initially (at time 0) full, i.e., the token count Tp(0) = PBS and the token count Tc(0) = CBS.
- Thereafter, the token count Tp is incremented by one PIR times per second up to PBS and the token count Tc is incremented by one CIR times per second up to CBS.

trTCM (continued ..)

- When a packet of size B bytes arrives at time t, the following happens if the trTCM is configured to operate in the *Color-Blind mode*:
  - If Tp(t)-B < 0, the packet is red, else
  - if Tc(t)-B < 0, the packet is yellow and Tp is decremented by B, else
  - the packet is green and both Tp and Tc are decremented by B.

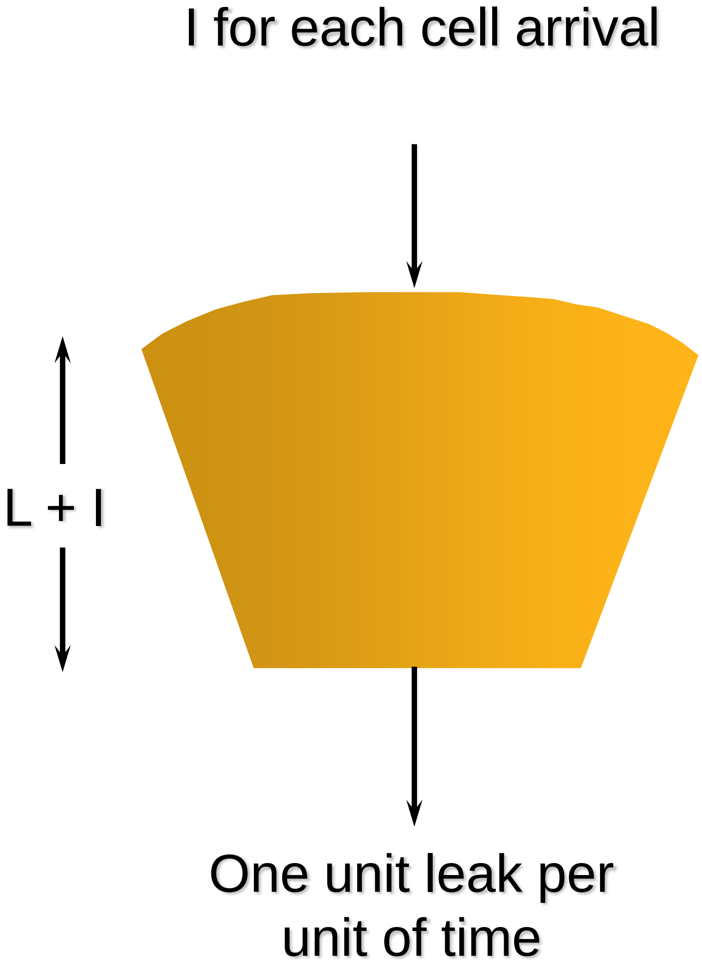
trTCM (continued ..)

- When a packet of size B bytes arrives at time t, the following happens if the trTCM is configured to operate in the *Color-Aware mode*:
  - If the packet has been precolored as red or if Tp(t)-B < 0, the packet is red, else
  - if the packet has been precolored as yellow or if Tc(t)-B < 0, the packet is yellow and Tp is decremented by B, else
  - the packet is green and both Tp and Tc are decremented by B.

Generic Cell Rate Algorithm (GCRA)

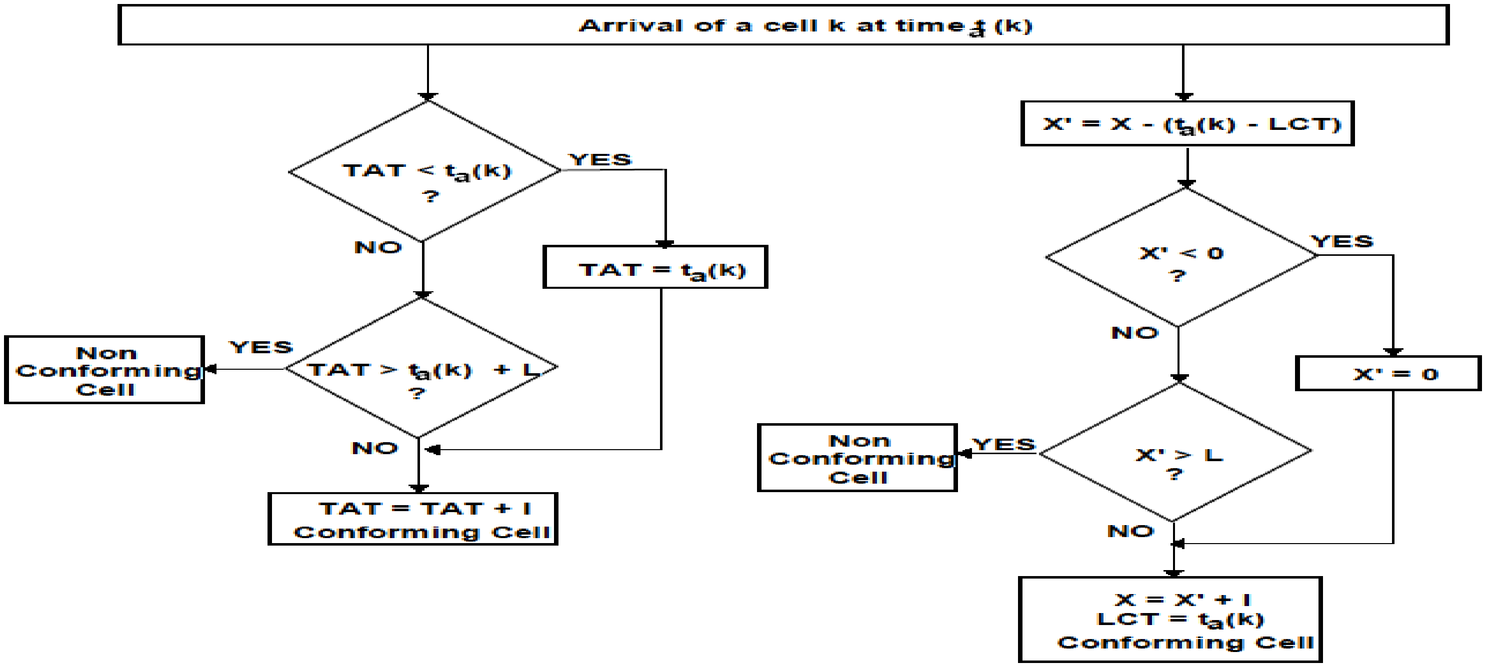
- GCRA is a continuous Leaky Bucket Algorithm. As each cell arrives, GCRA checks its conformance to an agreed rate.
- The GCRA has two parameters:
  - I is the Increment per time unit
  - L is the Limit or capacity
- When checking a rate with GCRA:
  - I is the reciprocal of the rate
  - L is the tolerance
- Use a single bucket for one rate (e.g., peak rate) policing
- Use two cascaded buckets for two rate (e.g., peak and average rate) policing
  - Conforming cells from peak rate policer pass through the second policer for average rate check up!!

Generic Cell Rate Algorithm



- For a sequence of cell arrival times, $\{t_k\}$, GCRA determines which cells conform to the traffic contract
- A counter scheme based on two parameters denoted GCRA(I,L)
  - Increment parameter: I
    - Defines minimum cell spacing
  - Limit parameter: L
    - Allows cell bursts and CDV
- "Leaky bucket"
  - *A cell that would cause the bucket to overflow is non-conforming*

Virtual Scheduling



VIRTUAL SCHEDULING ALGORITHM

CONTINUOUS-STATE LEAKY BUCKET ALGORITHM

TAT	Theoretical Arrival Time
t_a(k)	Time of arrival of a cell

X	Value of the Leaky Bucket counter
X'	auxiliary variable
LCT	Last Compliance Time

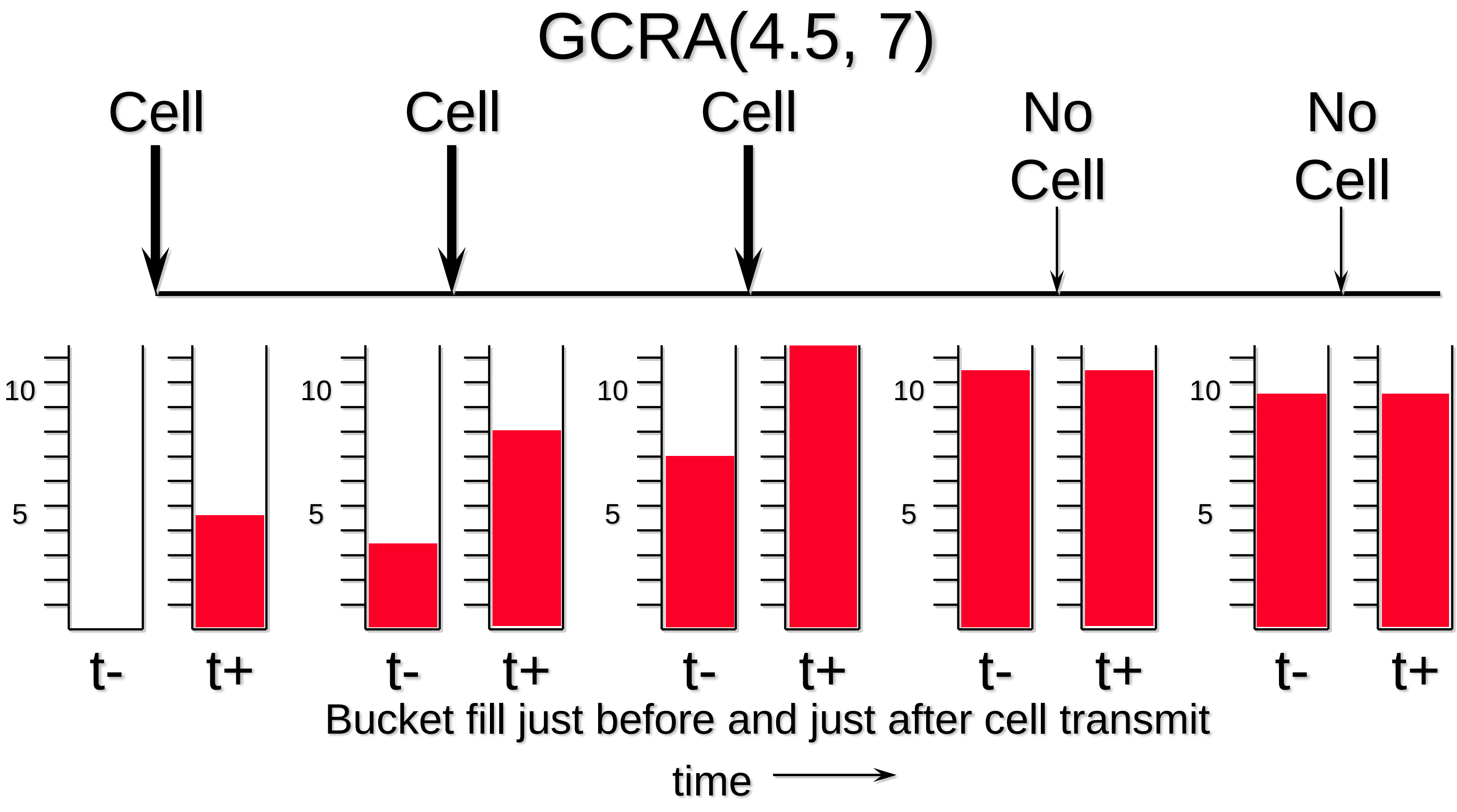
I	Increment
L	Limit

At the time of arrival t_a of the first cell of the connection, TAT = t_a(1)

At the time of arrival t_a of the first cell of the connection, X = 0 and LCT = t_a(k)

Figure 4-1: Equivalent versions of the Generic Cell Rate Algorithm

Bursty Traffic



Traffic Contracts – GCRA (I,L)

- Conformance Tests
- Typically Policed by the Network
  - CBR
    - Must conform to GCRA (1/PCR, CDVT); I=1/PCR; L=CDVT
  - VBR
    - Must conform to both:
      - GCRA (1/PCR, CDVT)
      - GCRA (1/SCR, CDVT+BT); BT = (MBS-1)(1/SCR-1/PCR)
  - ABR
    - Must conform to DGCRA (1/ACR , t) where:
      - ACR is the network specified rate (PCR>ACR>MCR)
      - t is a dynamically variable burst tolerance
  - UBR
    - Whether UBR must conform to GCRA (1/PCR, CDVT) is network specific

CDVT is the Cell Variance Delay Tolerance and is specific to the network, not signalled
BT is the Burst Tolerance for SCR

Traffic Conformance

Table 3.1. Conformance definitions per service categories.

Name	Service category	PCR flow	SCR flow	MCR flow	Non-conforming action	CLR	Max-CTD P2P-CDV
CBR.1	CBR	0+1	N/A	N/A	Discard	0+1	0+1
VBR.1	rt-VBR, nrt-VBR	0+1	0+1	N/A	Discard	0+1	0+1 (rt)
VBR.2	rt-VBR, nrt-VBR	0+1	0	N/A	Discard	0	0 (rt)
VBR.3	rt-VBR, nrt-VBR	0+1	0	N/A	Tag	0	0 (rt)
ABR.1	ABR	0	N/A	0	Discard	0	N/A
GFR.1	GFR	0+1	N/A	0	Discard	0	N/A
GFR.2	GFR	0+1	N/A	0	Tag	0	N/A
UBR.1	UBR	0+1	N/A	N/A	Discard	N/A	N/A
UBR.2	UBR	0+1	N/A	N/A	Tag	N/A	N/A

Policing CBR

Figure 3.3. Leaky bucket algorithm for CBR (PCR of aggregate flow).

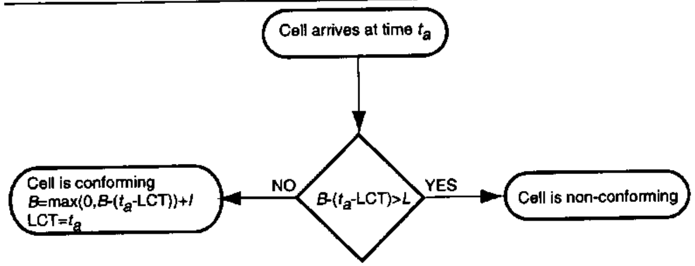
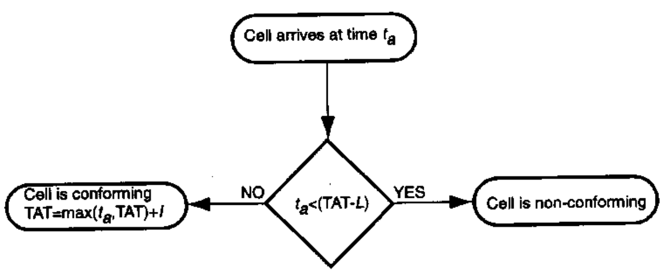


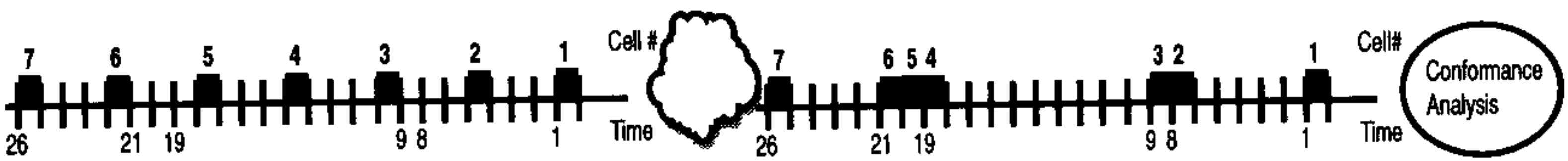
Figure 3.4. Virtual scheduling algorithm for CBR (PCR of aggregate flow).



- Bucket *B* fills with *I* units every time a conforming cell arrives and continuously leaks one unit every unit of time
- Bucket has a capacity of *L*
- Cell is conforming if bucket capacity *<=L* when a cell arrives
- Bucket overflows if cells arrive faster than drain rate
- Overflowing cells are non conforming
- Cell arrives at t_a, bucket drains by $(t_a\text{-}LCT)$, LCT is the last conforming time
- If cell is conforming, bucket is filled by I
- Can also be performed as Virtual Scheduling algorithm
- TAT is theoretical arrival time

Example

Figure 3.5. Cell stream originally equally spaced every 4 time units and jittered.

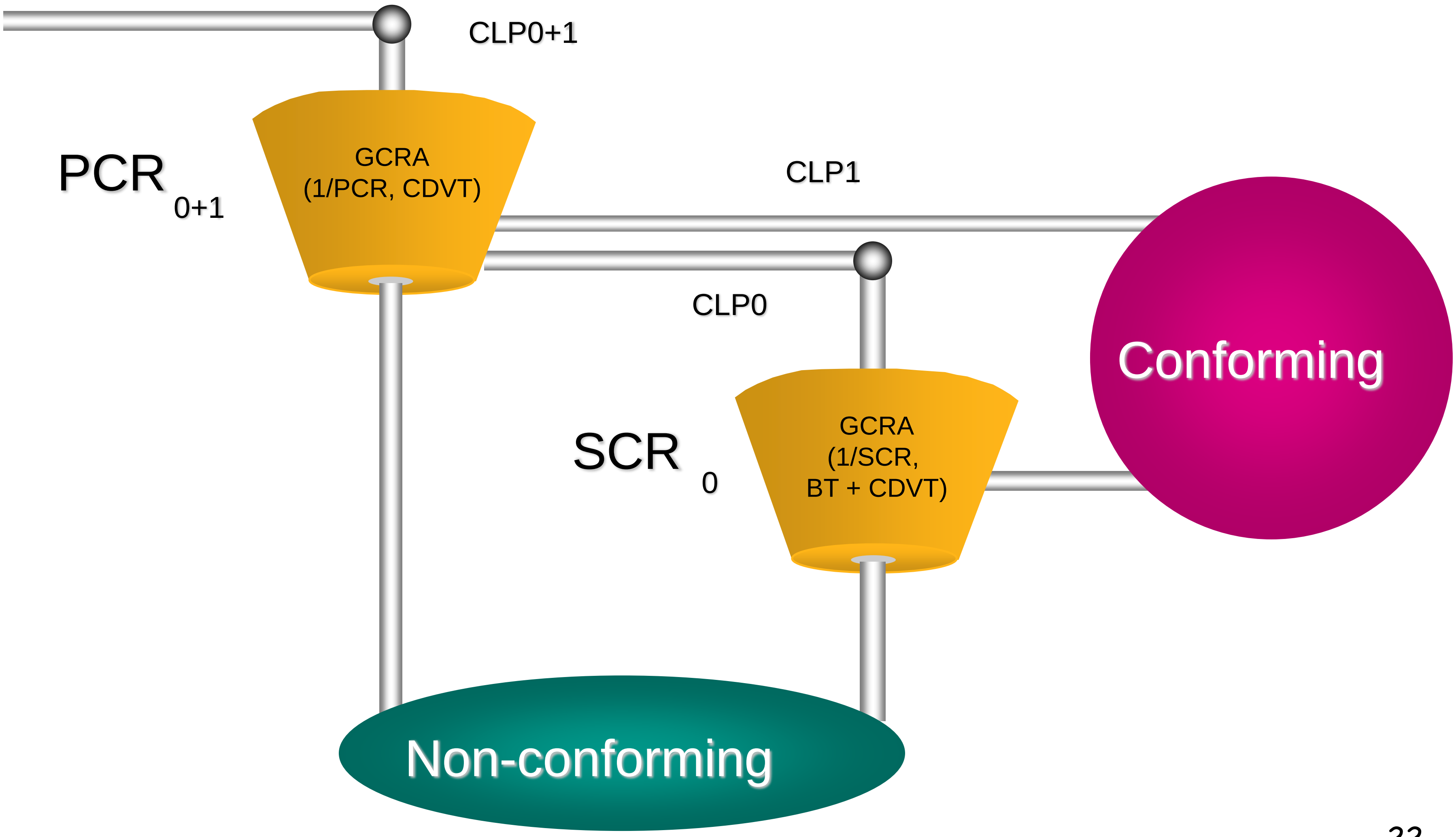


PCR = ¼ Line Rate; CDVT = 1 time unit

Table 3.2. Evolution of the leaky bucket and virtual scheduling on the cell stream of Fig. 3.5. $I = 4$, $L = 2$.

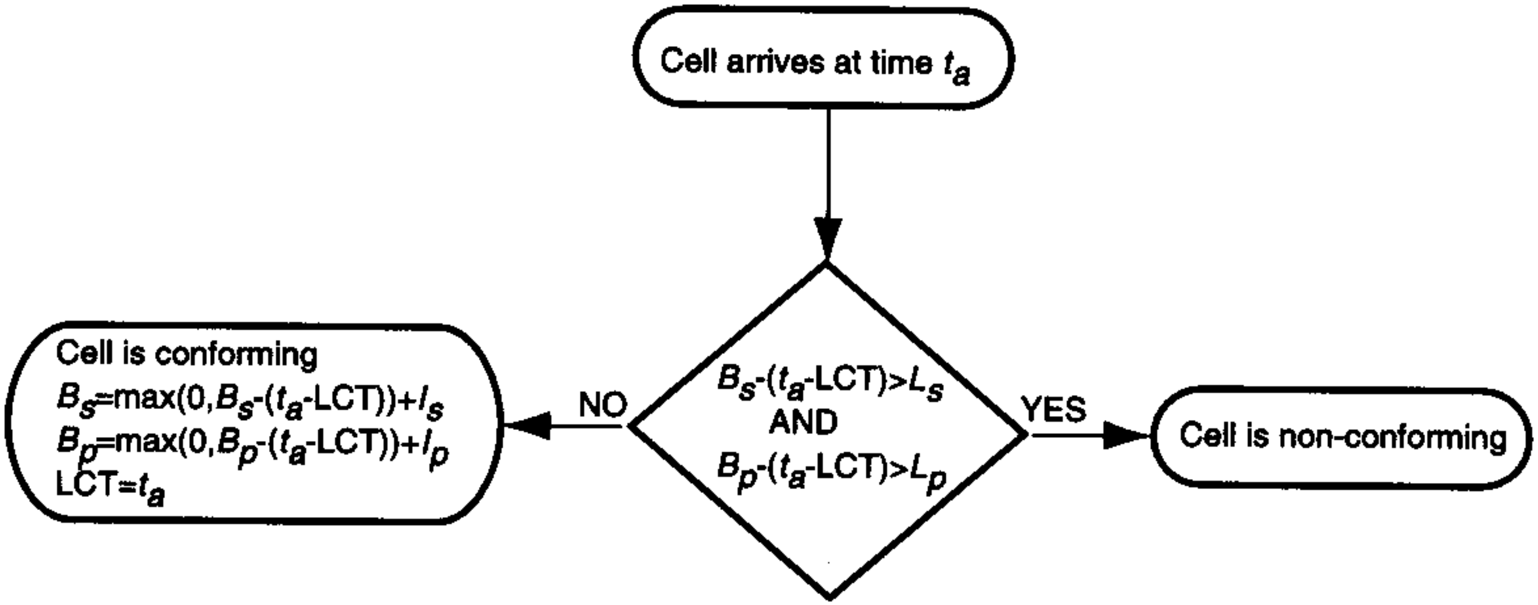
Cell#	t_a	TAT	TAT-t_a	B	LCT	B-(t_a – LCT)	Conformance	*TAT-L*
1	1	1	0	0	1	0	Yes	*0*
2	8	5	–3	4	1	–3	Yes	*3*
3	9	12	3	4	8	3	No	*10*
4	19	12	–7	4	8	–7	Yes	*10*
5	20	23	3	4	19	3	No	*21*
6	21	23	2	4	19	2	Yes	*21*

Example Conformance Definition for VBR



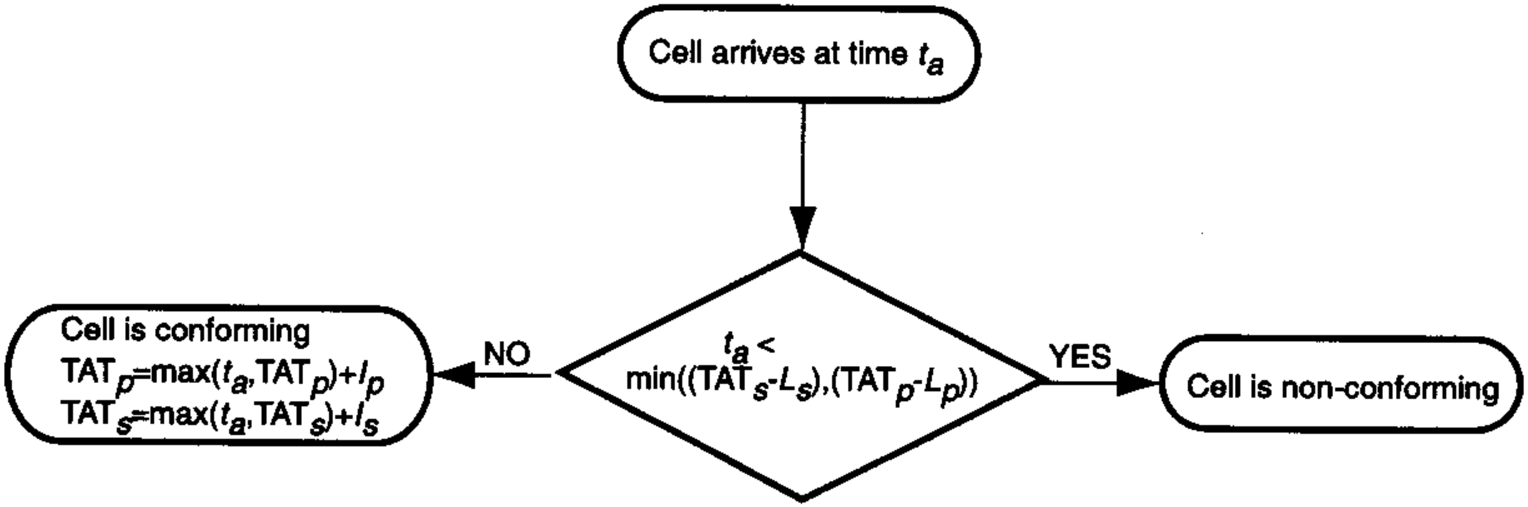
VBR Policing

Figure 3.8. Dual leaky bucket algorithm for VBR.1.



VBR Policing

Figure 3.10. Dual virtual-scheduling algorithm for VBR.1.



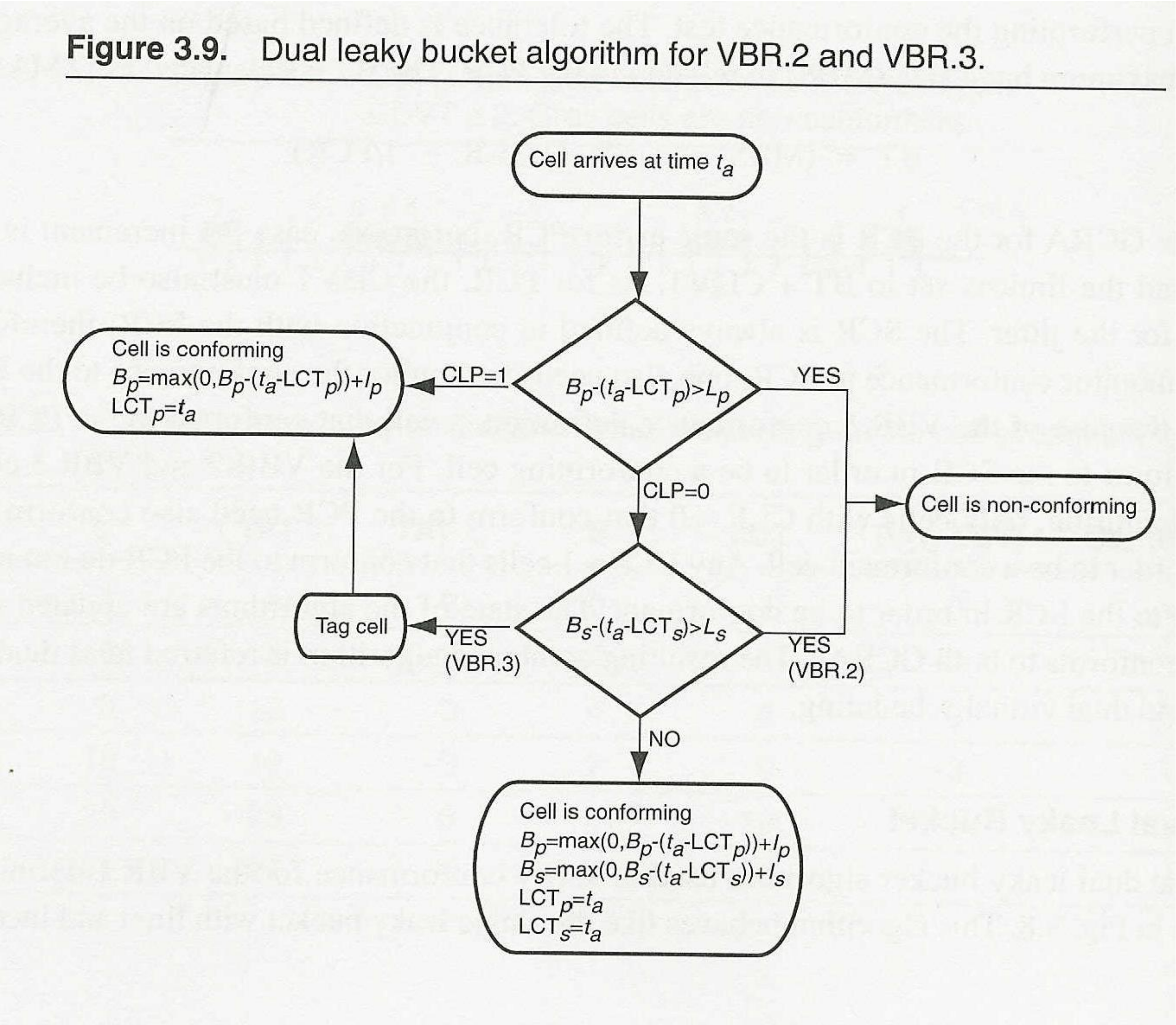
VBR Policing

Cell#	t_a	TAT_s	TAT_p	B_s	B_p	LCT	TAT_s-t_a = B_s-(t_a-LCT)	TAT_p-t_a = B_p-(t_a-LCT)	Conformance
1	1	1	1	0	0	1	0	0	Yes
2	2	5	3	4	2	1	3	1	Yes
3	3	9	5	7	3	2	6	2	No
4	4	9	5	7	3	2	5	1	Yes
5	5	13	7	9	3	3	8	2	No
6	6	13	7	9	3	3	7	1	Yes
7	7	17	9	11	3	6	10	2	No
8	8	17	9	11	3	6	9	1	No
9	9	17	9	11	3	6	8	0	No
10	10	17	9	11	3	6	7	–1	Yes

PCR = ½ Line Rate
SCR= ¼ Line Rate
MBS = 4 Cells
CDVT = 1 time unit

BT=6
$I_s=4;\ I_p=2$
Ls=7; Lp=2

VBR.2 and VBR.3 Policing



ATM Traffic Shaping

- Traffic shaping function modifies the traffic characteristics to conform to the traffic contract
- Traffic Shaping can be performed at the network ingress, intermediate points or at the egress
- The shaping function requires the traffic to be queued per-VC basis and allow the cells to be transmitted as per conformance
- Use reverse leaky bucket or reverse virtual scheduling algorithm to send a conforming cell
- A cell is transmitted if the bucket (B) is empty. When the cell is transmitted, the bucket is filled by I=1/PCR units and continuously leaks one unit every time until emptied.
- For two rate VBR shaping, a cell is scheduled if the PCR bucket is empty and if the SCR bucket fill is lower than BT. When the cell is transmitted, the PCR bucket fills by 1/PCR and SCR bucket by 1/SCR. Both buckets leak at 1 every unit time.