Generic Cell Rate Algorithm – GCRA

By: Anders Johansson, Dragos Ilie and Ola Nilsson

Contents Page

1 Abstracts

In order to make traffic controls over the network the ATM Forum has defined a range of functions to maintain the QoS of ATM connections. One traffic control function is the Usage Parameter Control (UPC). The UPC can use one of two algorithms in determing whether the user is complying with the traffic contract. This algorithm is the *Generic Cell Rate Algorithm* that (1) serves as an operational definition of the relationship between peak cell rate and CDVT, and (2) can be used for usage parameter control to monitor compliance with the traffic contract. This algorithm can also be used by the Network Parameter Control (NPC) but this paper will only describe it for the UPC case.

The algorithm is reffered to as the GCRA because it is used for the sustainable cell rate. There are two equivalent versions of the algorithm:

- Virtual scheduling algorithm
- Continuous-State Leaky Bucket algorithm

The algorithm takes two arguments, an increment *I* and a limit *L*, and is expressed as GCRA(I,L).

To get a conceptual understanding about the purpose of and some terms for the GCRA there is a section about traffic control in the next chapter.

2 Traffic control

ITU-T and the ATM Forum have defined a range of traffic control functions to maintain the QoS of ATM connections. ATM traffic control function refers to the set of action taken by the network to avoid congestion conditions or to minimize congestion effects. The following functions have been defined:

- Resource management using virtual paths
- Connection admission control
- Usage parameter control
- Selective cell discard
- Traffic shaping

In this paper the focus will be on the Usage Parameter Control and its algorithm to determine whether the user is complying with the traffic contract.

2.1 Usage Parameter Control

Once a connection has been accepted by the Connection Admission Control function, the Usage Parameter Control (UPC) function of the network monitors the connection to determine whether the traffic conforms to the traffic contract. The main purpose of UPC is to protect network resources from an overload on one connection that would adversely affect the QoS on other connections by detecting violations of assigned parameters and taking appropriate actions. The simplest strategy is that compliant cells are passed along and non-compliant cells are discarded at the point of the UPC action.

There are two separate functions encompassed by UPC:

- Control of peak cell rate and the associated CDVT
- Control of sustainable cell rate and the associated burst tolerance.

2.2 Traffic parameters and descriptors

Traffic parameters describe traffic characteristics of a source. For a given connection, traffic parameters are grouped into a source traffic descriptor, which in turn is a component of a connection traffic descriptor.

2.2.1 Traffic parameters

A traffic parameter describes an inherent characteristic of a traffic source. It may be quantitative or qualitative. Traffic parameters defined in this specification include Peak Cell Rate (PCR), Sustainable Cell Rate (SCR), Maximum Burst Size (MBS) and Minimum Cell Rate (MCR).

2.2.2 Source traffic descriptor

A source traffic descriptor is the set of traffic parameters of the ATM source. It is used during the connection establishment to capture the intrinsic traffic characteristics of the connection requested by a particular source.

2.2.3 Cell conformance and connection compliance

Conformance applies to cells as they pass the UNI and are, in principle, tested according to some algoritm. The first cell of the connection initializes the algorithm and from then on each cell is either conforming or non-conforming.

2.2.4 Peak cell rate conformance

The Peak Cell Rate (PCR) traffic parameter specifies an upper bound on the rate at which traffic can be submitted on an ATM connection. Enforcement of this bound by the UPC allows the network to allocate sufficient resources to ensure that the network performance objectives (e.g. for Cell Loss Ratio) can be achieved. In the signaling message, the PCR is coded in cells per second.

2.2.5 Sustainable cell rate and burst tolerance

The SCR is an upper bound on the average rate of the conforming cells of an ATM connection, over time scales, which are long relative to those for which the PCR is defined. Enforcement of this bound by the UPC could allow the network to allocate sufficient resources, but less than those based on the PCR, and still ensure that the performance objectives (e.g. for Cell Loss Ratio) can be achieved.

2.2.6 Cell Delay Variation Tolerance (CDVT) for PCR and SCR

ATM layer functions (e.g. cell multiplexing) may alter the traffic characteristics of connections by introducing Cell Delay Variation. When cells from two or more connections are multiplexed, cells of a given connection may be delayed while cells of another connection are being inserted at the output of the multiplexer. Similarly, some cells may be delayed while physical layer overhead or OAM cells are inserted. Consequently with reference to the peak emission interval T (i.e., the inverse of the contracted PCR), some randomness may affect the inter-arrival time between consecutive cells of a connection (i.e. the inverse of the contracted PCR) as monitored at the UNI (public or private). The upper bound on the “clumping” measure is the CDVT.

3 Generic Cell Rate Algorithm

Generic Cell Rate Algoritm (GCRA) is a method for describing the traffic flowing through a VCC. GCRA defines the five service categories: CBR, VBR-rt, VBR-nrt, ABR and UBR.
Up until now the service classes (or bearer capabilities in ITU-T terms) specified or standardised are:

- constant bit rate (CBR) traffic based on the specification of a peak bit rate and some tolerance value together with a so-called Generic Cell Rate Algorithm (GCRA);
- variable bit rate (VBR) traffic based on a specification which combines two GCRAs to describe a peak and a so-called sustainable cell rate. It has to be noted that this traffic class was specified with mainly frame relay traffic in mind;
- available bit rate (ABR) traffic as a type of flow-controlled specification with a maximum peak and a guaranteed minimum bit rate.

The GCRA algorithm was specified in ATM Forum Traffic Management Specification Version 4.0 in 1996. GCRA is used to provide the formal definition of traffic conformance to the negotiated traffic parameters. GCRA is used to define, in an operational manner, the relationship between PCR and the CDVT, and the relationship between SCR and the Burst Tolerance (BT). GCRA is not a required implementation for UPC or NPC; rather the network provider may use any UPC/NPC whose performance is consistent with the QOS objectives for the VCCs.

In contrast to a traffic shaping leaky bucket which controls the flow of complient cells, the GCRA leaky bucket simply monitors the traffic and rejects or discards noncompliant cells.

3.1 Virtual scheduling algorithm

The algorithm (figure 1) is initialized with the arrival of the first cell on a connection at time $t_a(1)$. The algorithm updates a theoretical arrival time (TAT), which is a target time for the next cell arrival. If the cell arrives later than the TAT, then it is compliant and the TAT is updated to the arrival time plus T. If the cell arrives earlier than TAT but within τ time units of TAT, then the cell is still considered compliant and TAT is incremented by T. In this latter case, it is permissible for the cell to arrive ealy because it does so within the CDVT. Finally, if the cell arrives too ealy (before TAT - τ) then it is outside the CDVT bound and is declared noncomlpiant. In this case TAT remains unchanged.

3.2 Continuous-state leaky bucket algorithm

The algorithm (figure 1) maintains a running count of the cumulative amount of data sent in a counter X. The counter is decremented at a constant rate of one unit per time unit to a minimum value of zero. This is equivalent to a bucket that leaks at a rate of 1. The counter is incremented by I.

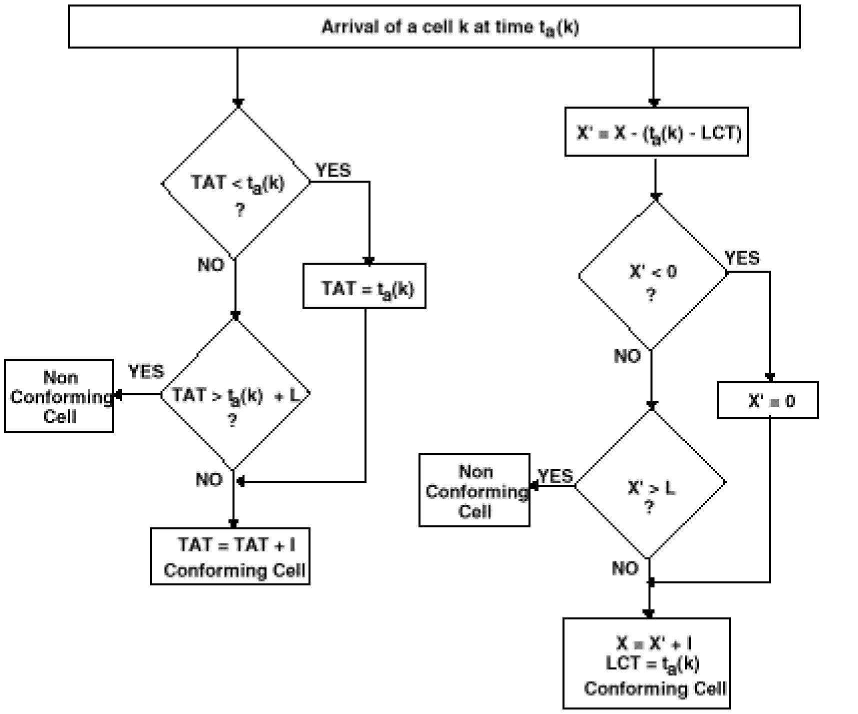
For each arriving cell, subject to the restriction that the maximum counter value is $I + L$. Any arriving cell that would cause the counter to exceed its maximum is defined as nonconforming. This is equivalent to a bucket with a capacity of $I + L$. The algorithm defines a finite-

capacity bucket that drains at a continuous rate of 1 unit per time unit and whose content is increased by T for each comliant cell. The total capacity of the bucket is $T + \tau$. After the arrival of the *k*th cell, at $t_a(k)$, the algorithm checks to see if the bucket has overflowed. If so, the cell is nonconforming. If not, the bucket is incremented. The amount of the increment depends on whether the bucket was fully drained between cell arrivals.

3.3 Sustainable cell rate algorithm

The sustainable cell rate algorithm (1) serves as an operational definition of the ralationship between sustainable cell rate and burst tolerance, and (2) can be used for usage parameter control to monitor compliance with the traffic contract.

The same algorithm that is used to define peak cell rate monitoring is also used to define sustainable cell rate monitoring. In this case, for a sustainable cell rate R_s, $T_s = 1/R_s$ is the interarrival time between cells at that rate if there is no burstiness. The burst tolerance is represented as τ_s. Thus the sustainable cell rate algorithm is expressed as GCRA(T_s,τ_s). Unlike the CDVT, the burst tolerance is not selected directly. Rather, it is derived from an understanding of the burstiness of the traffic stream. In particular, let T be the time between cells at the peak rate.



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 = $t_a(k)$

4 Glossary of acronyms and terms

ABR Available Bit Rate
ACR Allowed Cell Rate
BT Burst Tolerance
CBR Constant Bit Rate
CCR Current Cell Rate
CDV Cell Delay Variation
CDVT CDV Tolerance
CER Cell Error Ratio
CLP Cell Loss Priority
CLR Cell Loss Ratio
CTD Cell Transfer Delay
DGCRA Dynamic GCRA
EPRCA Enhanced Proportional Rate Control Algorithm
GCRA Generic Cell Rate Algorithm
IBT Intrinsic Burst Tolerance
ICR Initial Cell Rate
LCT Last Conformance Time
MACR Mean Allowed Cell Rate
maxCTD Maximum Cell Transfer Delay
MBS Maximum Burst Size
MCR Minimum Cell Rate
MCRmin Minimum acceptable MCR
NNI Network to node interface
NPC Network Parameter Control
Nrm Maximum number of cells between RM-cell generation
nrt-VBR Non-Real-Time VBR
PCR Peak Cell Rate
PDU Protocol Data Unit
QoS Quality of Service
rt-VBR Real-Time VBR
SCR Sustainable Cell Rate
SDU Service Data Unit
STD Source Traffic Descriptor
SVC Switched Virtual Connection
SW Switch
TAT Theoretical Arrival Time
UBR Unspecified Bit Rate
UDP User Datagram Protocol
UNI User Network Interface
UPC Usage Parameter Control
VBR Variable Bit Rate
VC Virtual Connection
VCC Virtual Channel Connection
VCI Virtual Channel Identifier
VPC Virtual Path Connection
VPI Virtual Path Identifier

VP-SW Virtual Path Switching Function

References

[ATM96] ATM Forum Traffic Management Specification, Version 4.0, 1996.

[ ]