Chapter 4: Congestion Control in ATM Networks

TOPICS
- Traffic characterization
- QoS parameters
- ATM categories of service
- Preventive congestion control
  - Call admission control
  - GCRA
- Reactive congestion control- ABR

Traffic characterization

- *peak cell rate* (PCR)
- *sustained cell rate* (SCR)
- *maximum burst size* (MBS)
- burstiness, and
- correlation of inter-arrival times
- Also
  - *cell delay variation tolerance* (CDVT)
  - *burst tolerance* (BT)

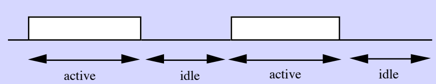
- Peak cell rate (PCR):
  - This is the maximum rate, expressed in cells per second, that can be submitted by a source to an ATM network.
  - Often, we use the *peak bit rate*, instead of the peak cell rate. One can be obtained from the other given that we know the specific AAL that is used.

- Sustained cell rate (SCR):
  - Compute the average number of cells submitted by the source over successive short periods T. The largest of all these averages is called the *sustained cell rate* (SCR)
  - T is not defined in the standards, but in the industry it is often taken to be equal to 1 second.
  - SCR is not to be confused with the average rate of cells submitted by a source. However, if we set T equal to the entire time that the source is transmitting over the ATM network, then the SCR becomes the average cell rate.

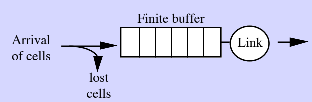
$$\boxed{\textit{Average cell rate} \leq \textit{SCR} \leq \textit{PCR}}$$

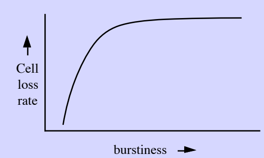
- Maximum burst size (MBS)
  - the maximum number of cells that can be submitted by a source back to back at peak cell rate.

- Burstiness
  - This is a *notion* related to how the cells transmitted by a source are clumped together



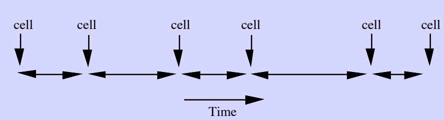
The burstiness of a source can significantly affect the performance of an ATM switch!!!





Cell loss rate vs burstiness

- Correlation
  - In an ATM environment it is highly likely that the inter-arrival times are correlated.
  - The correlation of a source can significantly affect the performance of an ATM switch



Standardized traffic descriptors

- The ATM Forum has standardized:
  - *peak cell rate (PCR),*
  - *sustained cell rate (SCR),*
  - *maximum burst size (MBS), and*
  - *cell delay variation tolerance (CDVT)*
- ITU-T has only standardized the *peak cell rate*.

Empirical Models

- Autoregression models are used for predicting the amount of traffic generated by an MPEG video coding algorithm for variable bit rate video.

- MPEG video compression:

  I B B B P B B B P B B B I …..

  - I-frame: compressed without any reference to the frames before it or after it.
  - B-frame: obtained by interpolation between the previous I-frame and the next P-frame, and a
  - P-frame: coded using motion-compensation prediction techniques with reference to the previous I-frame.
- The exact sequence of these frame types depends on the particular application.
- Autoregressive models predict size of I,B,P frames

An example of an autoregressive model

For the group of pictures

I B B P B B P B B P B B,

the following ARIMA model can be used to predict the number of bits S(i) of the ith frame:

$$S(i) = S(i-12) + e(i) – 0.69748\ e(i-3),$$

where e(i) is white noise and it follows the distribution $N(0, \sigma^2)$, with $\sigma^2 = 4849.5$ bits.

Probabilistic models

- Several models have been developed based on Markov modulated processes.
- In general, they are not realistic representations of arrival processes, but they capture the main characteristics of ATM traffic.
- Very useful in simulation and analytic models of ATM networks

Quality of service (QoS) parameters

- *cell loss rate* (CLR)
- *jitter*
- *cell transfer delay* (CTD)
- *Peak-to-peak cell delay variation* (CDV)
- *maximum cell transfer delay* (max CTD)
- *cell error rat*e (CER), and
- *cell misinsetion rate* (CMR)

Cell loss rate

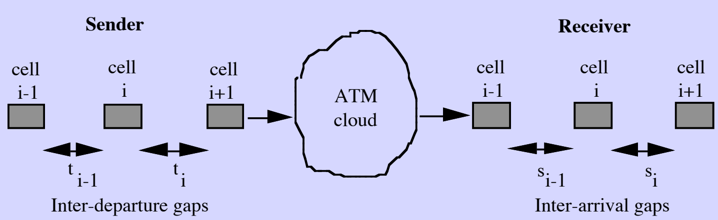
– This is a very popular QoS parameter and it was the first one to be used extensively in ATM networks.

– It is easy to quantify, as opposed to other QoS parameters such as jitter and cell transfer delay.

It has been used extensively as a guidance to dimensioning ATM switches, and in call admission control algorithms.

Jitter

An important QoS parameter for voice and video. It refers to the variability of the inter-arrival times at the destination

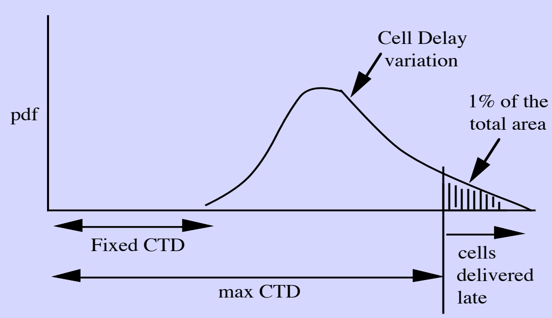


Cell transfer delay (CTD)

- The time it takes to transfer a cell end-to-end, that is, from the transmitting end-device to the receiving end-device. It comprises of
  - Fixed cell transfer delay
    - Propagation delay, fixed delays induced by transmission systems, and fixed switch processing times
  - Variable cell transfer delay, known as the peak-to-peak cell delay variation
    - Queueing delays in switches

Maximum cell transfer delay (max CTD)

This is a statistical upper bound on the end-to-end cell transfer delay.



Cell error ratio (CER) and Cell misinsertion rate (CMR)

- The CER of a connection is the ratio of the number of *errored* cells to the total number of cells transmitted by the source. An *errored* cell is a cell delivered with erroneous payload.
- CMR is the rate of cells delivered to a wrong destination, calculated over a fixed period of time.

ATM service categories

- *Constant bit rate* (CBR),
- *Real time variable bit rate* (RT-VBR),
- *Non-real time variable bit rate* (NRT-VBR),
- *Available bit rate* (ABR),
- *Unspecified bit rate* (UBR), and
- *Guaranteed frame rate* (GFR)

The constant bit rate (CBR) service

- Intended for real-time applications which require tightly coupled constrained delay and delay variations, such as
  - circuit emulation services,
  - constant-bit rate video, and
  - high-quality audio.
- Sources are expected to transmit at a constant rate

The real-time VBR service

- It is intended for real-time applications, i.e. applications that require constrained delay and delay variations, such as
  - video and
  - voice.
- Sources are expected to transmit at a variable rate and be bursty.

- The non-real time VBR service:
  - It is for variable bit rate and bursty sources which do not require real-time constraints.
- The unspecified bit rate (UBR) service:
  - It is intended for delay tolerant applications. It has no quality of service guarantees.
- The available bit rate (ABR) service:
  - Feedback-based service for sources that can adjust their transmission rate

Attributes for: CBR, RT-VBR, NRT-VBR, UBR

- CBR
  - Class attributes: PCR, CDVT
  - QoS attributes: peak-to-peak CDV, MaxCTD, CLR
- rt-VBR
  - Class attributes: PCR, CDVT, SCR, MBS, CDVT
  - QoS attributes: peak-to-peak CDV, MaxCTD, CLR
- nrt-VBR
  - Class attributes: PCR, CDVT, SCR, MBS, CDVT
  - QoS attributes: CLR
- UBR
  - PCR is specified, but it may not be subject to CAC and policing
  - No QoS parameters are signaled

Attributes for ABR and GFR

- ABR
  - Class attributes: PCR, CDVT, MCR
  - QoS attributes: CLR (possible, depends on network)
  - Other attributes: feedback messages
- GFR
  - Class attributes: PCR, CDVT, MCR, MBS, MFS, CDVT
  - QoS attributes: CLR (possible, depends on network)

Congestion control

- Preventive
  - It prevents the occurrence of congestion using
    - *Call admission control (CAC)*
    - *Policing (GCRA)*
- Reactive
  - It is based on feedback from the network to control transmission rates
    - *Available bit rate (ABR) service*

Preventive congestion control

- When a new connection is requested, each ATM switch on the path has to decide whether to accept it or not.
- Two questions need to be answered:
  - *Will the new connection affect the quality-of-service of the existing connections already carried by the switch?*
  - *Can the switch provide the quality-of-service requested by the new connection?*

Call admission control (CAC)

- The CAC algorithm is used by an ATM switch to decide whether to accept or reject a new connection.
- CAC algorithms may be classified into
  - *non-statistical allocation (or peak bit rate allocation)* , and
  - *statistical allocation.*

Non-statistical allocation

- It is used for connections requesting a CBR service.
- CAC algorithm is very simple.
  - The decision to *accept* or *reject* a new connection is based purely on whether its peak bit rate is less than the available bandwidth on the link.

Example

- Let us consider a non-blocking switch with output buffering, and suppose that a new connection with a peak bit rate of 1 Mbps has to be established through output link 1.
- Then, the new connection is accepted if the link's available capacity is more or equal to 1 Mbps.

Statistical allocation

- In this case, the allocated bandwidth is less than the peak bit rate of the source.
- In the case where statistical allocation is used for all the connections on the link, the sum of the peak bit rates of all the connections may exceed the link's capacity.

Statistical allocation makes economic sense when dealing with bursty sources.

However, it is difficult to implement effectively because:

- *It is difficult to characterize the traffic of a source and how it is shaped deep in the network.*
- *The CAC algorithm has to run real-time. There is no time for CPU-intensive calculations.*

A variety of CAC algorithms have been proposed

- Most CAC algorithms are based on the CLR
  - A new connection is accepted if the switch can provide the requested cell loss rate without affecting the cell loss rate of the existing connections. Jitter, or CTD are not taken into account
- New algorithms are emerging based on the cell transfer delay

The equivalent bandwidth of a source

- Let us consider a finite capacity queue served by a server at the rate of μ.
- We assume that this queue is fed by a single source, whose equivalent bandwidth we wish to calculate.
- Now, if we set μ equal to the source's peak bit rate, then we will observe no accumulation of cells in the buffer. This is due to the fact that the cells do not arrive faster than they are served.

- Now, if we slightly reduce the service rate μ, then we will see that cells are beginning to accumulate in the buffer.
- If we reduce the service rate further, then the buffer occupancy will increase.
- If we keep repeating this experiment and each time we lower slightly the service rate, then we will see that the cell loss rate begins to increase.

- *The equivalent bandwidth of the source is defined as the service rate e of the queue that corresponds to a cell loss rate of ε.*

- It falls somewhere between its average bit rate and the peak bit rate.
  - If the source is very bursty, it is closer to its peak bit rate, otherwise, it is closer to its average bit rate.
- We note that the equivalent bandwidth of a source is not related to the source's SCR.

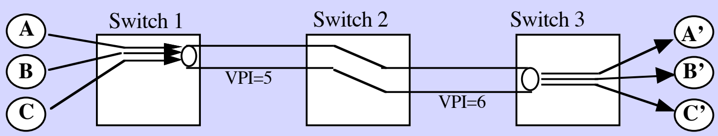
There are various approximations that can be used to compute quickly the equivalent bandwidth of a source.

The equivalent bandwidth of a source is used in statistical bandwidth allocation in the same way that the peak bit rate is used in non-statistical bandwidth allocation

Virtual path connections

- Used by operators to control congestion within their network.
- A virtual path connection is used to create a dedicated connection between two switches. Within this connection, individual virtual circuit connections can be set-up.

An example of virtual paths



1	3	1	5	47	3
1	4	3	5	39	3
2	5	5	5	41	3

5	47	2	6	47	8
5	39	2	6	39	8
5	41	2	6	41	8

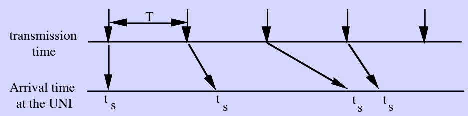
6	47	3	1	54	2
6	39	3	7	41	3
6	41	3	2	52	7

Bandwidth enforcement

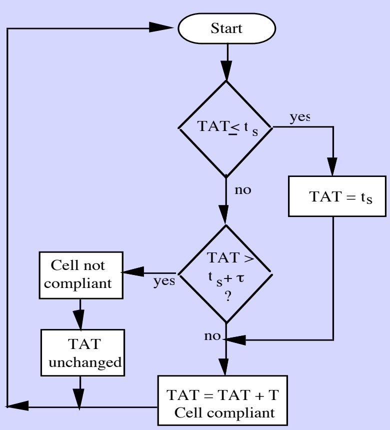
- Used to ensure that the traffic generated by a source conforms with the *traffic contract*, agreed between the user and the network at call set-up.
- The traffic contract consists of
  - *A connection traffic descriptor,*
  - *A requested quality of service class, and*
  - *A definition of conformance.*

The cell delay variation tolerance

- It's due to interleaving at the UNI level.
- A source may appear to transmit higher than its PCR.
- How much of that the network should tolerate?



The generic cell rate algorithm (GCRA) for policing PCR



- TAT, t_s: Theoretical and actual arrival time of a cell
- τ: Cell delay variation tolerance (CDVT). It is set by the administrator, and it permits the network to tolerate a number of cells that arrive faster than the agreed upon peak cell rate.

- Example:
  - $T=10$, $\tau=15$, actual arrival times: 0, 12, 18, 20, 25.
- Cell 1: $t_s=TAT=0$. The cell is accepted, $TAT=TAT+10=10$.
- Cell 2, $t_s=12$. and $TAT \le t_s$. The cell is accepted and $TAT= t_s+T=22$.
- Cell 3: $t_s = 18$ and $TAT>t_s$. Since $TAT \le t_s+\tau$ the cell is accepted and $TAT=TAT+T=32$.
- Cell 4: $t_s=20$, and $TAT>t_s$. Since $TAT \le t_s+\tau$ the cell is accepted and $TAT=TAT+T=42$.
- Cell 5: $t_s=25$ which makes $TAT>t_s$. Since $TAT> t_s+\tau$ the cell is considered as non-compliant.

Violation tagging

- GCRA will either accept a cell or classify it as non-compliant.
- Non-compliant cells are either *dropped* or *tagged* (CLP=1)
- If congestion arises inside the network the tagged cells are dropped

Push-out and threshold schemes

- Violation tagging introduces two types of cells: the untagged cell(CLP=0) and the tagged cell (CLP=1).
- A simple way to handle tagged cells is through a priority mechanism, such as the *push-out* scheme and the *threshold* scheme

Reactive congestion control

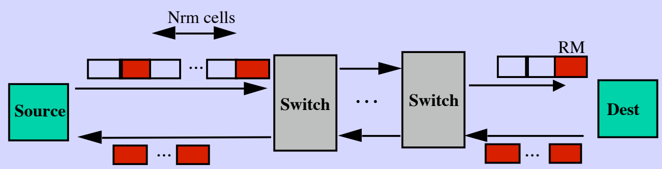
- In reactive congestion control, we let sources transmit without bandwidth reservation and policing, and we take action only when congestion occurs.
- High-speed links become a temporary storage
- In view if this, large buffers are required

Available bit rate (ABR) service

- This is a feedback-based mechanism whereby the sending end-device is allowed to transmit more during the time that there is a slack in the network.
- It utilizes RM messages to control the transmitting end-device.

- At connection set-up time, the sending end-device requests a *minimum cell rate* (MCR). It also specifies a maximum cell rate, which is its PCR.
- The network accepts the new connection if it can satisfy its requested MCR.
- The transmission rate of the source may exceed its requested MCR, if the network has slack capacity.
- When congestion begins to build up in the network, the sending end-device is requested to decrease its transmission rate.

RM messages are used to implement ABR



- The source sends an RM cell every Rnm-1 data cells. The defaulted value for Rnm is 32.
- The RM cells and data cells may traverse a number of switches before they reach their destination end-device.
- The destination end-device turns around the RM cells, and transmits them back to the sending end-device.
- Each switch writes information about its congestion status onto the RM cells, which is used by the sending end-device to adjust its transmission rate.

RM cell structure - fields

- *Message type field:* This is a one-byte field and it contains the following 1-bit sub-fields.
  - *DIR:* This bit indicates the direction of the RM cell.
  - *BN:* It indicates whether the RM cell is a *backward explicit congestion notification* (BFCN) cell.
  - *CI:* Congestion indication bit used by an ATM switch or the destination end-device, to indicate congestion in the network.
  - *NI:* No increase indicator, used to prevent the sending end-device from increasing its *allowed cell rate* (ACR), which is its current transmission rate.

- *Explicit rate (ER):*
  - A 2-byte field used to carry the explicit rate calculated by an ATM switch. The ER is used to limit the transmission rate of the sending end-device. This field may be subsequently reduced by another ATM switch, if it calculates a lower ER than the one indicated in the ER field of the RM cell.
- *Current cell rate (CCR)*:
  - A 2-byte field used by the sending end-device to indicate its ACR, i.e. its current transmission rate.
- *Minimum cell rate (MCR):*
  - This the minimum cell rate that the connection has requested and the network has agreed to guarantee.

The ABR mechanism: Binary mode

- In the binary mode, the switch marks the EFCN bit in the header of the data cells to indicate pending congestion.
- The destination translates the EFCN information into bits such as the CI or NI, which are marked in the corresponding backward RM cell.
- Upon receipt of this, the source takes appropriate action.
  - *a) increase the transmission rate,*
  - *b) decrease the transmission rate, or*
  - *c) no change to the transmission rate.*

The ABR mechanism: ER mode

- In the explicit rate mode, a switch computes a local fair share for the connection and marks the rate at which source is allowed to transmit in the ER field of the backward RM cell.
- The source, upon receipt of the backward RM cell, sets its transmission rate to the ER value.
- When detecting congestion, a switch can generate a backwards RM cell in order to convey the congestion status, without having to wait for a backwards RM cell to arrive.

Source behavior

- The source adjusts its ACR according to the information received in the RM cell. ACR should be always:

$$MCR \leq ACR \leq PCR$$

- After ACR has been adjusted, it is set to the ER field if it is less than the adjusted ACR. It cannot be lower than MCR.

The ACR is adjusted as follows:

- If CI=1, ACR is reduced by at least ACRxRDF, where RDF is a pre-specified *rate decrease factor*. If the reduction results to a value below MCR, then ACR is set equal to MCR.
- If both CI=0 and NI=0, then the ACR may be increased by no more than RIFxPCR, where RIF (*rate increase factor*) is a pre-specified quantity. The resulting ACR should not exceed PCR.
- If NI=1, then the ACR is not increased.

Destination behaviour

- When a data cell is received, its EFCN is saved in the EFCN status of the connection.
- On receiving a forward RM cell, the destination transmits it back to the source. The DIR bit is changed from forward to backward, BN=0, and CCR, MCR, ER, CI, and NI fields in the RM cell are unchanged, unless
  - If the saved EFCN status of the connection is set, then the destination set CI=1 in the RM cell, and re-sets the EFCN state.
  - If the destination is experiencing internal congestion, it may reduce the ER to whatever rate it can support and set either CI=1 or NI=1.
- The destination may also generate a new backward RM cell, where CI=1 or NI=1, DIR=1, and BN=1.

Switch behaviour

A switch shall implement at least one of the following methods:

- *EFCN marking:* The switch may set the EFCN bit in the header of the data cells.
- *Relative rate marking:* The switch may set CI=1 or NI=1 in forward and/or backward RM cells.
- *Explicit rate marking:* The switch may reduce the ER field of forward and/or backward RM-cells.