

.....
ATM Forum Document Number: ATM_Forum/97-0612.
.....

Title: Revised MIMO Definition
.....

Abstract: A minor change to MIMO definition to avoid negative values is presented.
.....

Source:

Gojko Babic, Arjan Durresi, Raj Jain, Justin Dolske
The Ohio State University
Department of CIS Columbus, OH 43210-1277
Phone: 614-292-3989, Fax: 614-292-2911, Email: Jain@ACM.Org

The presentation of this contribution at the ATM Forum is sponsored by NASA.
.....

Date: July 1997
.....

Distribution: ATM Forum Technical Working Group Members (AF-TEST, AF-TM)
.....

Notice:

This contribution has been prepared to assist the ATM Forum. It is offered to the Forum as a basis for discussion and is not a binding proposal on the part of any of the contributing organizations. The statements are subject to change in form and content after further study. Specifically, the contributors reserve the right to add to, amend or modify the statements contained herein.
.....

In the Appendix A of the current performance baseline test, the frame latency through a network element is defined as follows:

$$\text{MIMO latency} = \begin{cases} \text{LILO latency} & \text{if Input Link Rate} \leq \text{Output Link Rate} \\ \text{FILO latency} - \text{NFOT} & \text{if Input Link Rate} \geq \text{Output Link Rate} \end{cases}$$

Here, NFOT is the nominal frame output time. It is computed as follows:

$$\text{NFOT} = \text{Frame input time} \times \text{Input rate} / \text{output rate}$$

The NFOT was intended to reflect the expected delay through the zero-delay switch. It turns out the formula has a minor inaccuracy if the input rate is more than the output rate. In this case, the input gaps also get multiplied by speed ratio (which is larger than 1). This is not necessary. The result is that NFOT is overestimated and MIMO latency can be negative. There is no problem in other cases and so the rest of the discussion is limited only to the case of input rate more than the output rate.

A simple fix to the problem is to consider the gaps more carefully. The zero-delay switch will not multiply gaps by the speed ratio. Instead, it will start transmitting each cell as soon as possible. This can be done as soon as the first bit of the cell arrives and the output link is idle. Since the cell output time (COT) is larger than cell input time, it is possible that an arriving cell has to wait for transmission of previous cells to be finished. This wait is due to the speed difference and cannot be attributed to the switch. The following simple modification in NFOT definition is sufficient to fix the problem:

Initially NFOT = 0 and time t is measured from the arrival of the first bit of the first cell. For each cell with its first bit arriving at time t \Rightarrow NFOT = max{t, NFOT} + COT.

The explanation is as follows. At intermediate points during frame transmission, NFOT represents the expected finish time for all cells that have arrived so far. A cell arriving at an idle link at time t should start transmitting immediately and finish at time t + COT. If a cell arrives earlier than NFOT, the previous cells have not been (actually cannot be) transmitted and so the link is not idle. In this case, the transmission will start at NFOT and finish at NFOT + COT. The new expected finish time is the new NFOT. Thus, the new NFOT will be either t + COT or NFOT + COT, whichever is larger.

Example 1: Figure 1 shows one case of two characteristic cases of a frame consisting of two cells passing through the zero delay switch. The input speed is 4 times the output speed. The two cells start arriving at time t = 0 and t = 5, respectively. The zero-delay switch will start transmitting the first cell at time t = 0 and finish at time t = 4. The second cell can be transmitted without waiting and finished at t = 9. This is how long the zero-delay switch will take to transmit this frame. Hence NFOT for this frame is 9.

The same result can be obtained from the algorithm presented above. Initially NFOT = 0. After the first bit of the first cell arrives at t = 0, NFOT = max{0, 0} + 4 = 4. After the first bit of the second cell arrives at t = 5, NFOT = max{5, 4} + 4 = 9

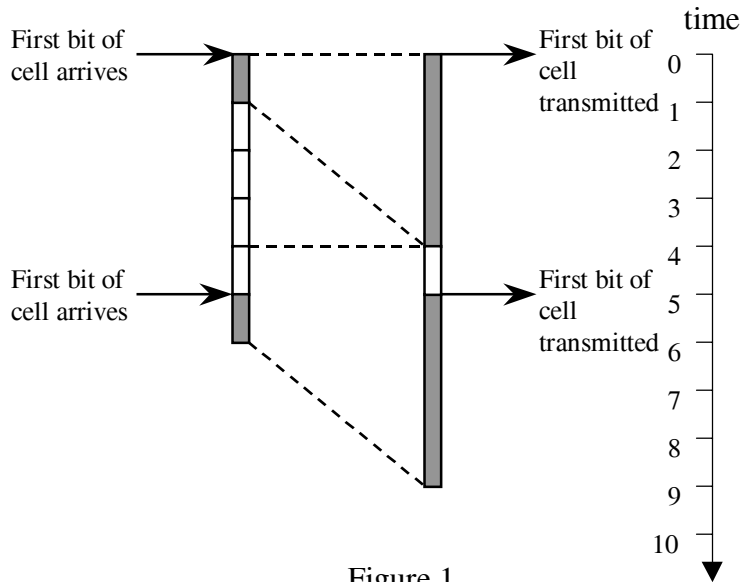


Figure 1

□

Example 2: Figure 2 shows another characteristic case of a frame consisting of two cells passing through the zero delay switch. As in Example 1, the input speed is 4 times the output speed. However, the second cell arrives at time $t = 2$ and thus has to wait. The zero-delay switch will start transmitting the first cell at time $t = 0$ and finish at time $t = 4$. The second cell can be transmitted at $t = 4$ and finished at $t = 8$. Hence NFOT for this frame is 8.

The same result can be obtained from the algorithm presented above. Initially $NFOT = 0$. After the first bit of the first cell arrives at $t = 0$, $NFOT = \max\{0, 0\} + 4 = 4$. After the first bit of the second cell arrives at $t = 2$, $NFOT = \max\{2, 4\} + 4 = 8$

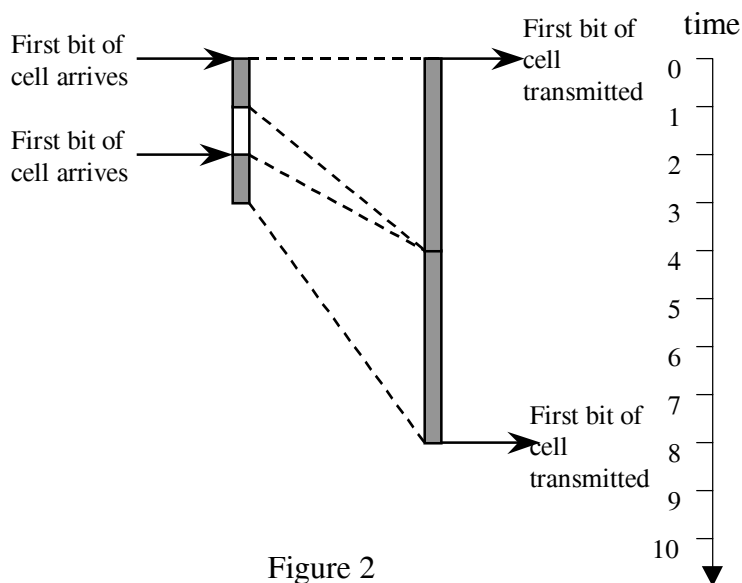


Figure 2

□

The above expression for NFOT is valid only if the input rate is larger than the output rate. A more general expression, which applies to all rates is:

Initially $NFOT = 0$ and time t is measured from the arrival of the first bit of the first cell. For each cell with its first bit arriving at time $t \Rightarrow NFOT = \max\{t, NFOT\} + CT$.

Here, the cell time CT is defined as the cell input time (CIT) or cell output time (COT), whichever is larger. For input rates less than the output rate, $CT = CIT$. For input rates larger than output rate, $CT = COT$.

Example 3: Figure 3 illustrates the case of two cell frame passing through the zero delay switch with input rate less than output rate. The output speed is 4 times the input speed. The two cells arrive at time $t = 0$ and $t = 5$, respectively. The zero-delay switch will start transmitting the first cell at time $t = 3$, not at $t = 0$ to avoid an underrun, and finish at time $t = 4$. The second cell starts at $t = 8$ and finishes at $t = 9$. This is how long the zero-delay switch will take to transmit this frame. Hence NFOT for this frame is 9.

The same result can be obtained from the new algorithm presented. Initially $NFOT = 0$. After the first bit of the first cell arrives at $t = 0$, $NFOT = \max\{0, 0\} + 4 = 4$. After the first bit of the second cell arrives at $t = 5$, $NFOT = \max\{5, 4\} + 4 = 9$.

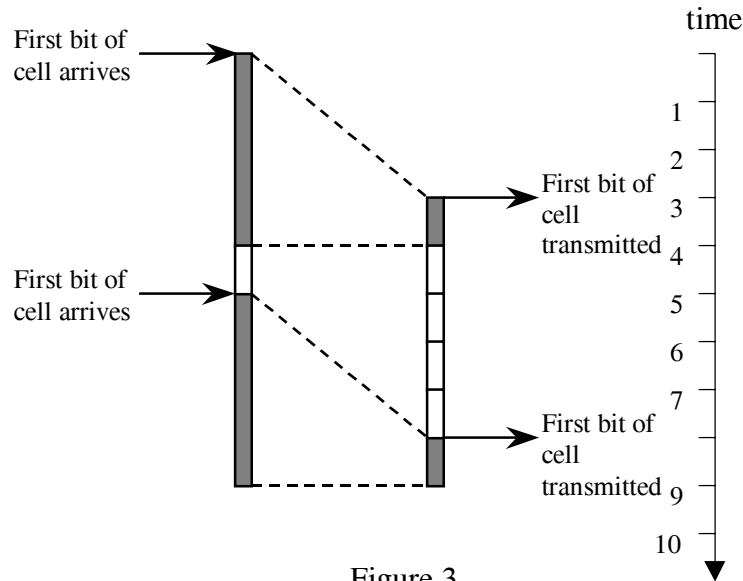


Figure 3

□

Note that in the case of input rate less than output rate, a cell never has to wait for completion of transmissions of previously received cells.

An advantage of the NFOT generalization is that the MIMO latency can be expressed simply as:

$$\text{MIMO Latency} = \text{FILO Latency} - \text{NFOT}$$

That is, the MIMO latency is the difference between the measured FILO latency of the given switch and the same for the zero-delay switch. This formula applies for all input/output rates as long as the NFOT is computed appropriately.

Proposed Revised Text for Section 3.2.1

3.2.1. Definition

The frame latency for a system under test is measured using a "Message-in Message-out (MIMO)" definition. Succinctly, MIMO latency is defined as follows:

$$\text{MIMO latency} = \text{FILO latency} - \text{NFOT}$$

where

- FILO latency = Time between the first-bit entry and the last-bit exit
- NFOT = Nominal Frame Output Time, defined as the time a frame needs to pass through *the zero-delay switch*, that can be calculated using the following procedure:

Initially NFOT = 0 and time t is measured from the arrival of the first bit of the first cell. For each cell with its first bit arriving at time t \Rightarrow NFOT = max{t, NFOT} + CT.

Here CT is the larger of the cell input time or cell output time. Cell times are computed as the cell size of 424 bits divided by the respective link rates in bits per sec.

An equivalent MIMO latency definition is:

$$\text{MIMO latency} = \begin{cases} \text{LILO latency} & \text{if Input Link Rate} \leq \text{Output Link Rate} \\ \text{FILO latency} - \text{NFOT} & \text{otherwise} \end{cases}$$

where

- LILO latency = Time between the last-bit entry and the last-bit exit

Frame Latency Measurements and Calculation

To obtain MIMO latency for a given frame, the time of occurrence for the following two events need to be recorded:

- First-bit of the frame enters into the SUT,
- Last-bit of the frame exits from the SUT.

The time between the second and the first events is FILO latency. If measurement data are available at cell level, what is usually the case with contemporary ATM monitors, it can be shown that:

$$\text{FILO latency} = \text{First cell's transfer delay} + \text{First cell to last cell inter-arrival time}$$

where

- cell transfer delay (CTD) is the time between the first bit of the cell entering the switch and the last bit of the cell leaving the switch,
- cell inter-arrival time is the time between arrival from the switch of the last bit of the first cell and arrival from the switch of the last bit of the second cell.

Given the cell pattern of a frame on input, NFOT can be obtained using the procedure from its definition. Then, substituting FILO latency and NFOT in the MIMO latency formula would give the SUT delay for the given frame.

In the cases when $\text{Input Link Rate} \leq \text{Output Link Rate}$, MIMO latency can be obtained easier. In those cases, the time of occurrence for the following two events need to be recorded:

- Last-bit of the frame enters into the SUT,
- Last-bit of the frame exits from the SUT.

The time between the second and the first events is LILO latency. When measurement data are available at cell level, it can be shown that:

$$\text{LILO latency} = \text{Last cell's transfer delay} - \text{Cell input time}$$

and in these cases, LILO latency would give the SUT delay for the given frame.

An explanation of MIMO latency and its justification is presented in Appendix A.