CONTRIBUTION TO T1 STANDARDS PROJECT

STANDARDS PROJECT: Specification and Allocation of ISDN Performance (T1Q1-10)

TITLE: Frame Delay Through ATM Switches and networks: MIMO Latency and its aggregation.

AUTHOR: Arjan Durresi, Raj Jain, and Gojko Babic

SOURCE: Ohio State University

CONTACT: Raj Jain

Raj Jain is now at
Washington University in Saint Louis
Jain@cse.wustl.edu
http://www.cse.wustl.edu/~jain/

n Science
)H 43210-1277
.1, Email: jain@cis.ohio-state.edu

DATE: May 3, 1999

DISTRIBUTION: Working Group T1A1.3

ABSTRACT: This contribution addresses the problem of measuring frame latency in ATM switches. The frames consisting of several ATM cells may arrive with numerous gaps between cells. It is important that the gaps present in the input stream be not counted towards the switch's contribution to the frame delay. The proposed solution called "MIMO" (Message-In Message-Out) latency improves upon FILO (First-In Last-Out) and LILO (Last-In Last-Out) latencies commonly used for continuous frame technologies such as frame relay.

Briefly, the MIMO latency is defined as the difference between the LILO latency through the switch and that through an ideal switch. The definition and the discussion also apply to any network of switches as well.

This contribution shows also how to compute the MIMO latency of a network path consisting of several components from the MIMO latencies of individual components.

NOTICE

This document has been prepared to assist the Standards Committee T1 - Telecommunications. It is offered to the Committee as a basis for discussion and is not a binding proposal on Ohio State University. The requirements presented in this document are subject to change in form and numerical value after more study. Ohio State University specifically reserves the right to add to, or amend, the statements contained herein.

1. Problem Statement

The performance of ATM equipment and the quality of services have been defined in terms of cell-level metrics such as cell transfer delay (CTD), cell delay variation (CDV) and cell loss ratio (CLR). However, cell-level metrics do not very often reflect the performance as experienced (or desired) by end users. For example, a video user sending 30 frames/sec would like frames to be completely delivered every 33 ms and it does not matter whether the cells belonging to a frame arrive back-to-back or regularly spaced. Thus, it is the frame delay and its variation that matters, not CTD and CDV. A frame is defined here as the ATM Adaptation Layer (AAL) protocol data unit (PDU).

One problem in measuring the frame delay in ATM networks is that when seen inside the network, the frames may be discontinuous with numerous gaps between the cells as well as cells of other frames. Note that the monitoring equipment, if placed inside the host, will be affected by the performance of the host and may not accurately reflect the performance of the switch. Thus, the test probes of the monitoring equipment should be placed at the measurement points (MP) at the entrance and the exit of the system to be measured, as in Figure 1 at physical connection layer.

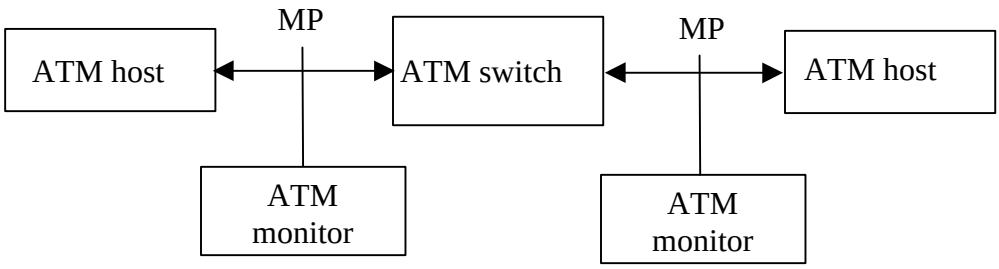


Figure 1. Measurement Point

To describe frame-level performance parameters it is important to define frame-level referents events (FRE) observable at MP at physical layer. Using the definitions of cell entry and exit events in ATM Forum Traffic Management Specification Version 4.0 [2] and in ATM Forum/BTD-TEST-TM-PERF.00.11 [1], FREs can be defined as follows:

Frame-level Reference Exit Event (FRE1): the occurrence of the cell exit event for the first user data cell of the frame;

Frame-level Reference entry Event (FRE2): the occurrence of the cell entry event for the last user data cell of the frame.

Using these definitions of FRE1 and FRE2, we define the following latency terms:

FILO frame latency is the difference between the time t2 of a FRE2 at MP2 at the output switch port and the time t1 of the corresponding FRE1 at MP1 at the input switch port.

LILO frame latency is the difference between the time t2 of a FRE2 at MP2 at the output switch port and the time t1 of the corresponding FRE2 at MP1 at the input switch port.

LIFO frame latency is the difference between the time t2 of a FRE1 at MP2 at the output switch port and the time t1 of the corresponding FRE2 at MP1 at the input switch port.

The latency of switch at the cell level is generally measured by FILO (first-bit in to the last-bit out) latency as indicated in Figure 2. Other alternatives such as FIFO (first-bit in to the first-bit out), LILO (last-bit in to the last-bit out), and LIFO (last-bit in to the first-bit out) latencies can be easily obtained from the figure. Most ITU documents measure cell level latency using FILO metric.

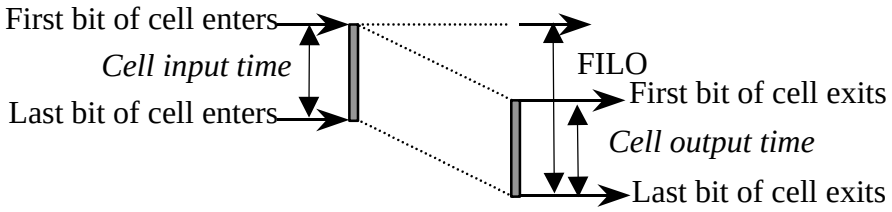


Figure 2. FILO latency at cell level

Although we use the term "switch" throughout this contribution, the discussion applies equally well to any network element (including switches, routers, multiplexers, inverse-multiplexers, wires, etc). Similarly, although this contribution relates to frames segmented in to ATM cells, it can be easily applied to IP frames segmented into fragments by IP routers.

One way to measure switch latency at the frame level is to measure the delay between the first bit in and the last bit out events for the frame. This is so called FILO latency introduced by the switch at the frame level. For example, consider the configuration shown in Figure 3 consisting of a switch with input and output ports of 155 Mbps (OC-3). The input frame is composed of two cells with a gap of one cell time. Let us assume that the switch A introduces a delay of 5c to each cell, where c is the cell time at OC-3. As a result, the FILO latency (interval between the first bit in and the last bit out events of the frame) is 8c and LILO latency is 5c.

Generally, the measured performance of a system depends upon the system as well as the workload. Some metrics are highly workload dependent while others are less dependent. A metric, which depends more on the system and less on the workload, is generally preferred particularly if the users are interested in comparing the systems and not the workloads. It turns out that the FILO frame latency as defined above has the undesirable property that it depends heavily on the workload. For example, see Figure 4. Here the two cells of the frame arrive with a gap of 5c, the switch B delays each cell by c. The FILO frame latency is 8c, which is the same as in Figure 3 for switch A. Clearly, switch B is better, but the FILO latency does not reflect that fact. The LILO frame latency is c and indicates the fact that switch B is better than switch A.

To show the problem in its extreme case, consider the situation in Figure 5, where the two cells of the frame arrive two days apart. Switch B delays each cell by c. But the FILO frame latency is two days plus 3c. It mostly reflects the arrival gap and is nowhere

close to the actual delay introduced by the switch. Again LILO frame latency is c and indicates properly the switch latency.

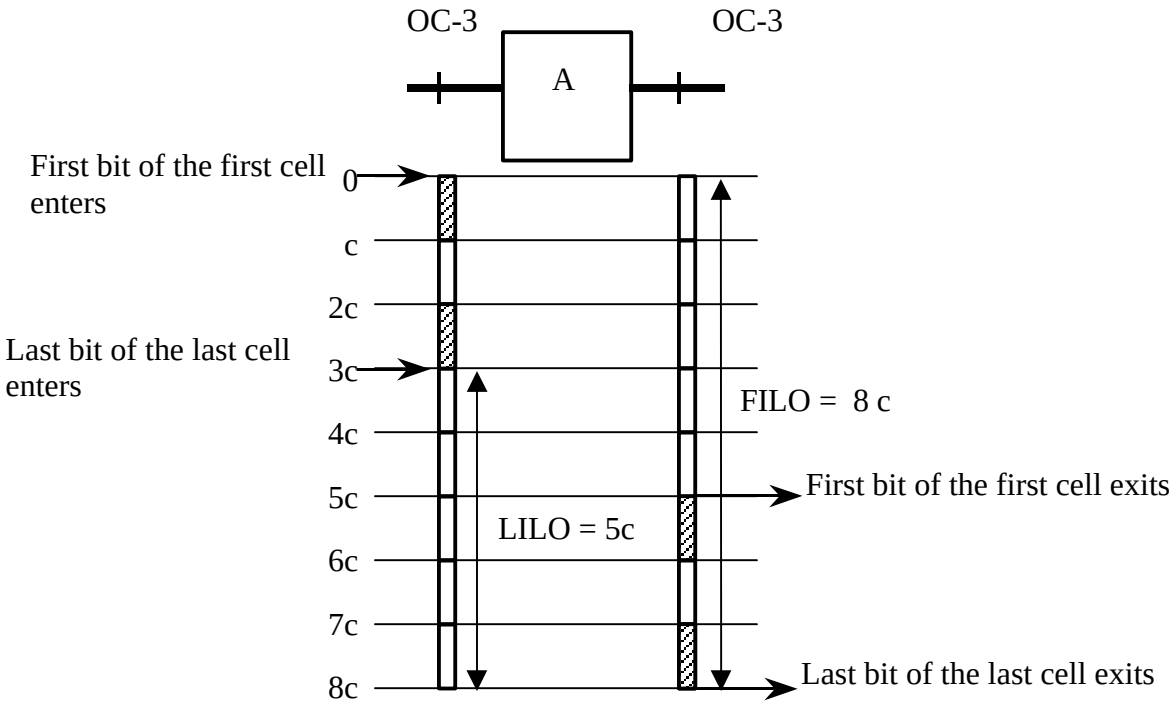


Figure 3. FILO and LILO latencies for the switch A that delays each cell by 5c

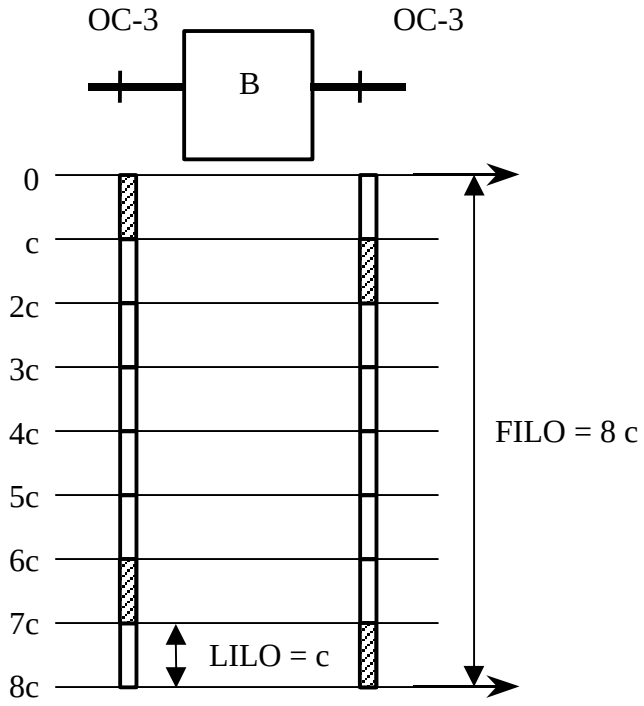


Figure 4. FILO and LILO latencies for the switch B that delays each cell by c

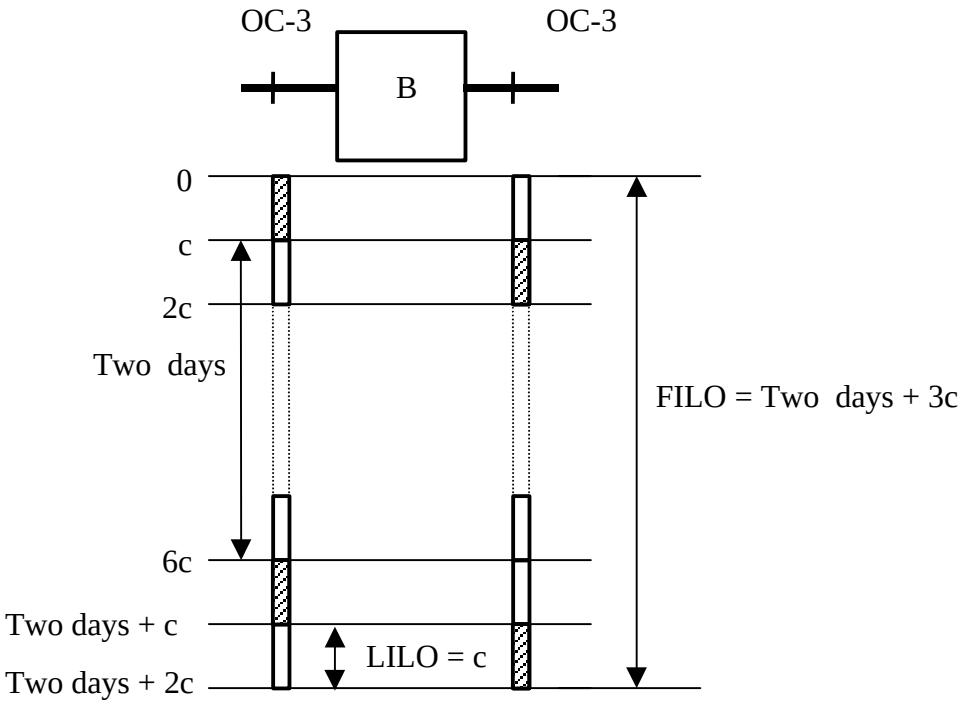


Figure 5. FILO and LILO latencies for the switch B that delays each cell by c

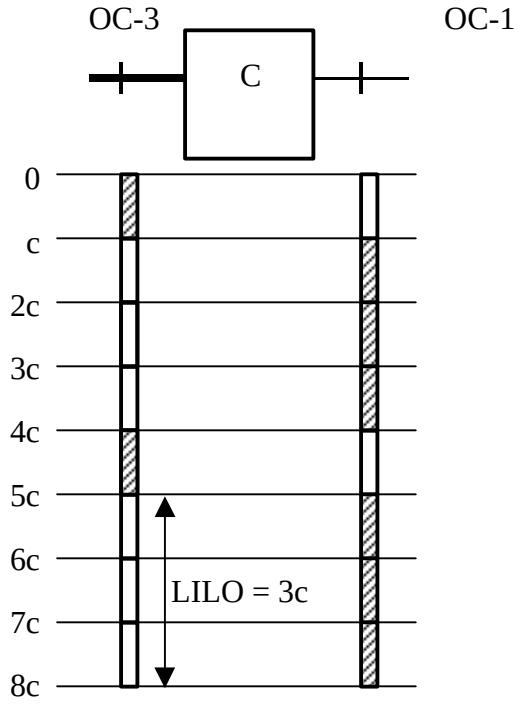


Figure 6. LILO latency for the switch C that delays each cell by c

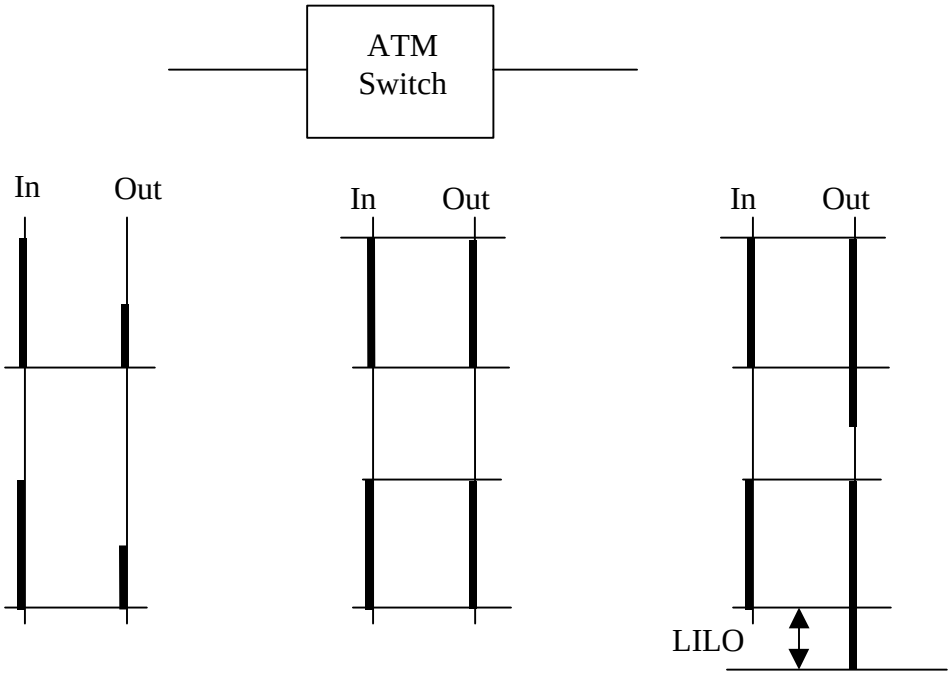
The LILO frame latency also depends on the workload. For example, see Figure 6. The input and output ports of Switch C are OC-3 and OC-1 respectively. Here the two cells of the frame arrive with a gap of 3c, the switch B delays each cell by c. The LILO frame latency is 3c and is not a good indication of delay attributed to the switch.

In this contribution, we propose a new metric called MIMO (Message In Message Out) latency that measures the true contribution of the switch to the frame latency and is not affected by the arrival patterns (gaps) of the cells constituting the frame. We introduce the concept of an ideal switch that does the best possible processing of its frames. MIMO latency is calculated for any given arrival pattern as the LILO frame latency for the pattern through the ideal switch ($LILO_0$) subtracted from the measured LILO frame latency of the switch under test gives, i.e.:

$$\text{MIMO latency} = \text{LILO latency} - LILO_0 \qquad (1)$$

Here, LILO is the measured LILO latency and $LILO_0$ is the LILO latency of an ideal switch for the same input pattern. Given input and output speeds, $LILO_0$ can be easily computed. Figure 7 shows three possible cases. The figure shows that $LILO_0$ is zero unless input speed is faster than the output speed.

$$LILO_0 = 0 \quad \text{if input speed} \leq \text{output speed}$$



a. Input Speed < Output Speed **b**. Input Speed = Output Speed **c**. Input Speed > Output Speed

Figure 7. An ideal switch introduces a nonzero LILO latency only when input link speed is greater than the output link speed.

In the example shown in Figure 3, the LILO frame latency through the ideal switch, $LILO_0$, is 0 and so MIMO latency for the switch A is 5c-0c or 5c. Similarly, for the example shown in Figure 4, $LILO_0$ is 0 and so MIMO latency from the switch B is c. For the example shown in Figure 5, $LILO_0$ is 0 and so MIMO latency for the switch B is again c. Finally for the example shown in Figure 6, $LILO_0$ is 2c and so MIMO latency for the switch C is c. Notice that in each case, MIMO latency reflects the switch behavior and is not affected by the arrival pattern.

In Section 2 of this contribution we present a more rigorous definition of the MIMO latency. The ideal switch is defined in Section 3. Section 4 presents some of our measurement tests of MIMO latency.

2. MIMO Latency Definition

As discussed above, MIMO latency is defined as:

$$\text{MIMO latency} = \text{LILO latency} - LILO_0$$

$LILO_0$ has a nonzero value when input link speed is greater than the output link speed. $LILO_0$ for a given frame is equal to the LILO latency of that frame passing through *an ideal switch*. An ideal switch is defined as a switch that handles incoming frames in such way that they are transmitted on the output link without any unnecessary time consumption, i.e. the best any switch can do. By definition, MIMO latency for an ideal switch is zero. Hence, an ideal switch can also be called a zero-delay switch.

$LILO_0$ is calculated as:

$$LILO_0 = FOT_0 - FIT \qquad (2)$$

Where FIT is the Frame Input Time and FOT_0 is the corresponding Frame Output Time produced by an ideal switch. FOT_0 is calculated on the fly as follows:

a. Initially $FOT_0 = 0$ and time t is measured from the arrival of the first bit of the first cell.
b. For each cell with its first bit arriving at time t, update FOT_0 as follows:

$$FOT_0 = \max\{t, FOT_0\} + CT$$

where:

CT = cell time = Max{CIT, COT}
CIT = Cell Input Time = 424 bits / Input Link Rate in bps
COT = Cell Output Time = 424 bits / Output Link Rate in bps

Note that MIMO latency, as a switch delay metric, accounts only for delays caused by node processing, such as switching, routing and queuing delays, and not by transmission delays introduced by communication links.

3. Cell and Frame Latency through an Ideal Switch

The concept of an ideal switch is explored in depth in this section. Figure 8 illustrates how an ideal switch would handle a cell. The switch behavior depends upon the relationship between the input and output link rates. In the case when the input link rate is equal to the output link rate, as presented in Figure 8a, an ideal switch transmits each bit as soon as it arrives. Thus, each bit of the cell experiences zero latency in an ideal switch.

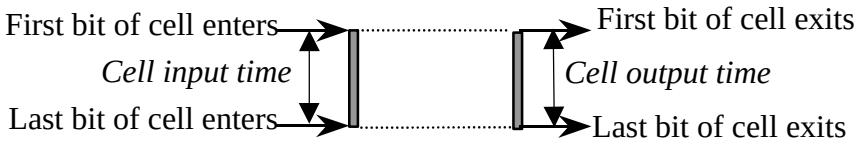


Figure 8a. Cell Processing of an Ideal Switch for Input Rate = Output Rate

Figure 8b illustrates the case when the input link rate is higher than the output link rate. In this case, outputting (transmitting) a bit takes longer than inputting it. The ideal switch can transmit only the first bit as soon as it is received. The other bits of the cell can not be transmitted immediately as they arrive, because the transmission of all previously received bits has not yet finished. Bits at the end of the cell wait longer then bits at the beginning. Thus, an ideal switch in this situation should be intelligent to do appropriate buffering of incoming bits.

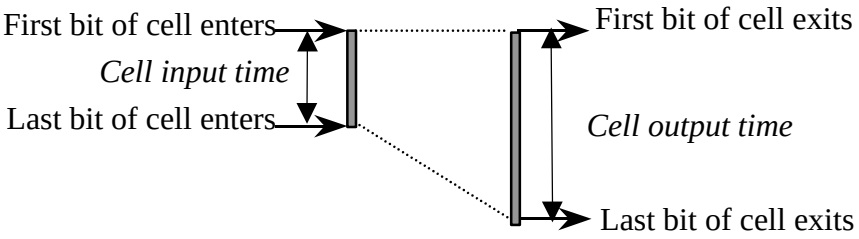


Figure 8b. Cell Processing of an Ideal Switch for Input Rate >Output Rate

Figure 8c illustrates the case when the input link rate is lower than the output link rate. An ideal switch does not start transmission of the first bit immediately after it is received, but after an appropriate delay. Bits at the beginning of the cell are delayed more than bits at the end, with larger delays for slower output link rates. Only the last bit of a cell has no delay and it is transmitted immediately upon its arrival. Thus, an ideal switch should be intelligent to avoid under-runs by appropriately delaying the transmission of incoming bits.

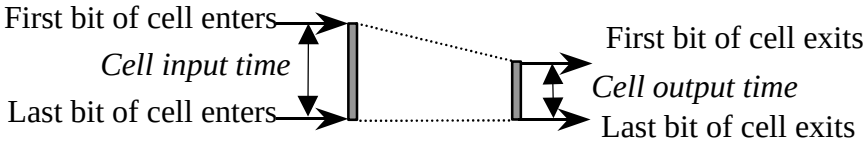


Figure 8c. Cell Processing of an Ideal Switch for Input Rate < Output Rate

It should be easily realized that the illustrations in Figures 8 apply not only to cells, but also to contiguous frames. Note that none of the usual latencies (FILO, LILO, FIFO, or LIFO) has a zero value in all three cases, as it should be for delays of a cell (frame) passing through an ideal switch.

The rest of this section considers how an ideal switch handles discontinuous frames in an ATM environment.

Figures 9 presents two possible cases of a frame passing through an ideal switch with the input link rate higher than the output link rate. Figure 9a illustrates the case when cells of a frame do not have to wait. The given frame includes two cells and the input link rate is four times the output link rate. The two cells start arriving at time t = 0c and t = 5c, respectively, where c is the input cell time. An ideal switch will start transmitting the first cell at time t = 0c and finish at time t = 4c. The second cell can be transmitted without waiting and the transmission is finished at t = 9c. This is how long an ideal switch will take to transmit this frame. Hence, FOT_0 of an ideal switch for this frame is 9c and $LILO_0$ is equal to 3c.

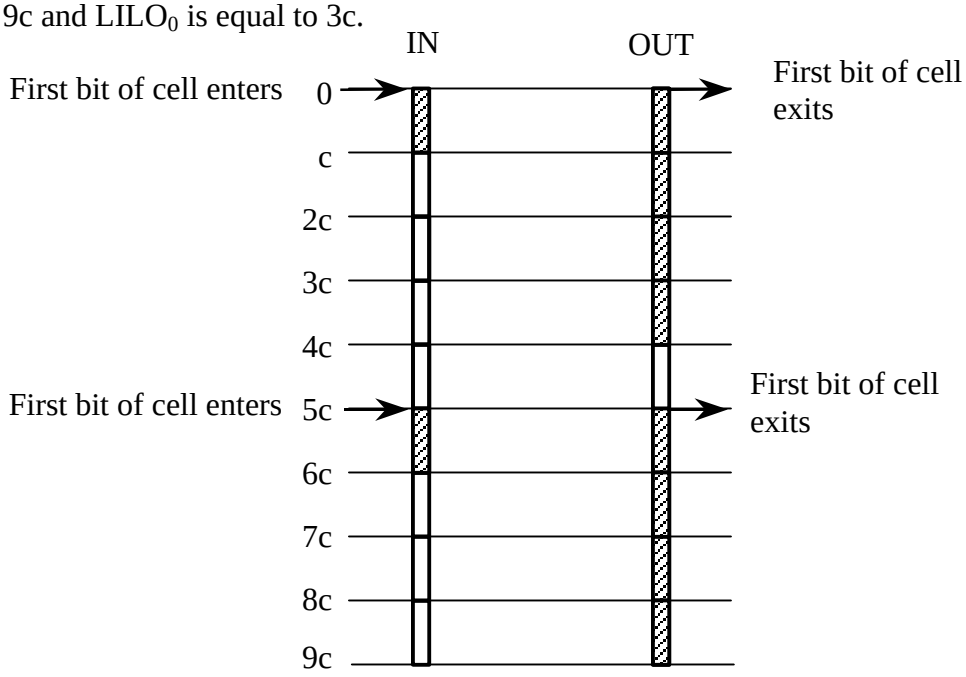


Figure 9a. No-Cell-Waiting Operation of an Ideal Switch for Input Rate > Output Rate

Figure 9b shows the another possible case of a frame passing through an ideal switch with an input link rate higher than the output link rate when cells of a frame have to wait. As in Figure 9a, the given frame has two cells and the input link rate is four times the output link rate. However, the frame has a different gap pattern. The second cell arrives at time t = 2c and thus has to wait. An ideal switch will start transmitting the first cell at time t = 0 and finish at time t = 4c. The second cell transmission starts at t = 4c and it is finished at t = 8c. Hence, FOT_0 of an ideal switch for this frame is 8c, i.e. $LILO_0$ = 5c.

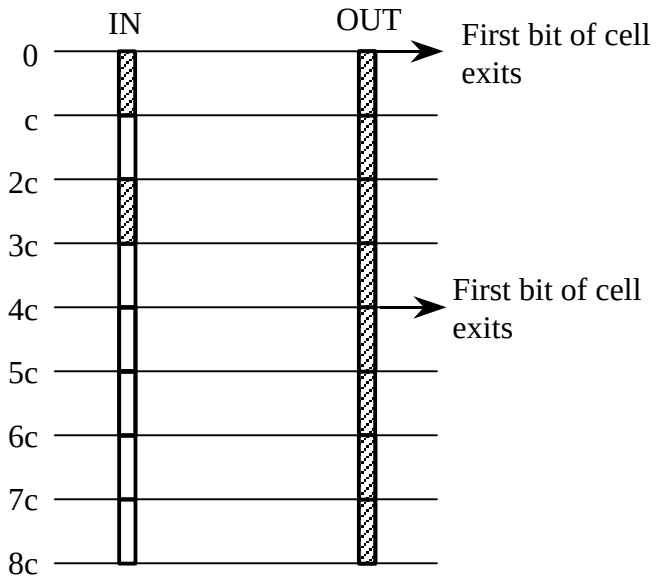


Figure 9b. Cell-Waiting Operation of an Ideal Switch for Input Rate > Output Rate

Thus, Figures 9 illustrate possibilities that an incoming cell can be transmitted immediately without waiting and that an incoming cell has to wait for previously received cells of the same frame to be transmitted.

In general, for a given discontinuous frame when the input link rate is higher than the output link rate, it is possible that some cells have to wait on previously received cells of the same frame, while some cells can be transmitted without waiting. Also, notice that ideal switch on output decreases the size of each gap from input, with some gaps being completely removed.

Figure 10 illustrates the only possible case of a frame passing through an ideal switch with an input rate lower than the output rate. Again, the frame includes two cells but the output link rate is now four times the input link rate. The two cells arrive at time t = 0c and t = 5c, respectively. An ideal switch will start transmitting the first cell at time t = 3c (not at t = 0, in order to avoid an underrun), and finish at time t = 4c. The second cell transmission starts at t = 8c and finishes at t = 9c. This is how long an ideal switch will take to transmit this frame. In this case $LILO_0$= 0.

Figure 11 illustrates the only possible case of a frame passing through an ideal switch with an input rate equal to the output rate. Again, the frame includes two cells. The two cells arrive at time t = 0c and t = 5c, respectively. An ideal switch will start transmitting the first cell at time t = 0c and finish at time t = 1c. The second cell transmission starts at t = 5c and finishes at t = 6c. This is how long an ideal switch will take to transmit this frame. Hence, $LILO_0$ = 0.

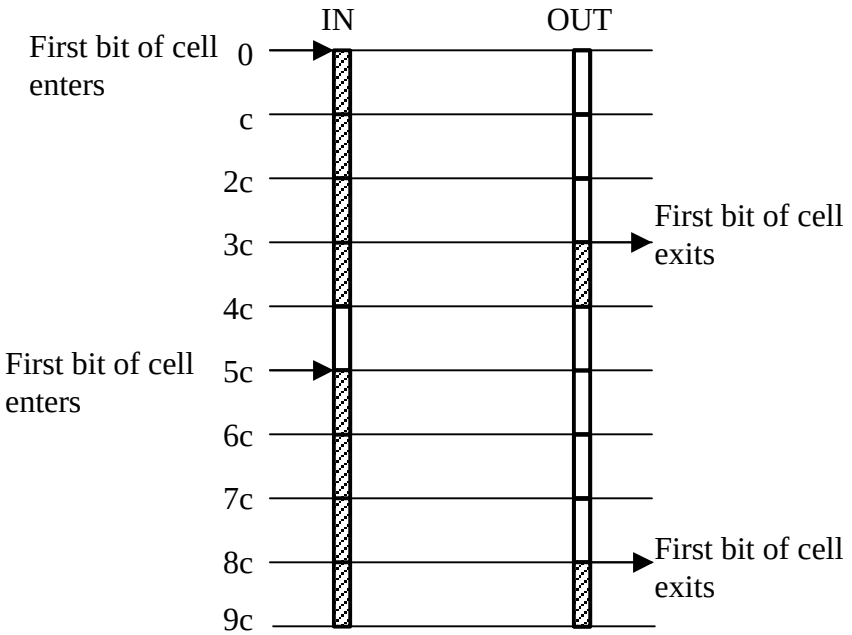


Figure 10. Operations of an Ideal Switch for Input Rate < Output Rate

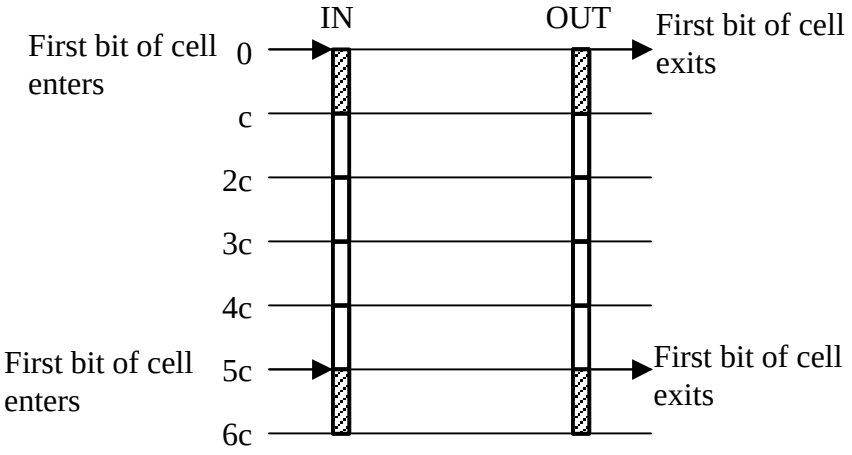


Figure 11. Operations of an Ideal Switch for Input Rate =Output Rate

4. Measurement Experiences

MIMO latency can be measured by using data from existing ATM monitors. In this section we describe several measurements performed in our performance laboratory using a commercial available ATM monitor as a traffic generator as well as a traffic analyzer. This monitor and, as far as we are aware all other similar systems, can provide measurement data on delays and inter-arrival times at the cell level.

The following relation, which can be easily derived, are used later in this section for MIMO latency calculation:

FILO latency = First cell to last cell inter-arrival time at the output
+ First cell transfer delay (3)

LILO latency = Last cell transfer delay – CIT (4)

Where CIT is Cell input time. Note that the inter-arrival time between two cells is defined by current ATM monitors as the time between the arrivals of the last bits of the two cells. The cell transfer delay is defined as the FILO latency for the cell.

4.1. Tests with Input Rate equal to Output Rate

The test configuration for MIMO latency measurements is shown in Figure 10. The configuration includes one ATM monitor and one ATM switch with 155 Mbps UTP-5 link between the monitor port 1 and the switch port A1 and 155 Mbps OC-3c link between the monitor port 2 and the switch port B1. The switch has two network modules A and B with four ports on each module. A permanent virtual channel connection (VCC) is established between the monitor ports 1 and 2 through the switch ports A1 and B1. That VCC is used for transmission of frames whose latency is measured. Figure 10 also indicates the traffic flow direction.

For each test run, a sequence of equally spaced 192-cell frames (cells of each frame are generated back-to-back) was sent over the VCC at a rate of 4.63 frames/s. After the flow had been established, we recorded the transfer delays of the last cells in the next 1,000 consecutive frames. The average last cell transfer delay was found to be 20.78 μs.

Since in this configuration:

- CIT = 424[bits] / Input Link Rate = 424[bits] / 149.76 [Mbps] = 2.83 μs

the average MIMO latency calculation using expressions (1) and (4) is given as:

MIMO latency = Last cell transfer delay - CIT = 20.78 – 2.83μs = 17.95 μs

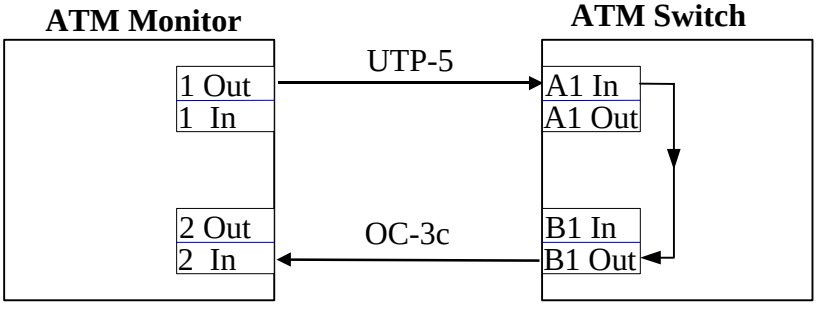


Figure 10. Test configuration for measurements of MIMO latency

Table 1 presents measurement data for two randomly chosen frames and calculated MIMO latency.

Table 1. (all times in μs)

Last cell CTD	1st cell CTD	1st cell to last cell inter-arrival time	MIMO latency	FILO latency
21.5	21.5	541.0	18.67	562.5
21.0	18.5	543.5	18.17	562.0

4.2. Tests with Input Rate Higher Than Output Rate

The test configuration for the MIMO latency measurements for the case with the input link rate higher than the output link rate**,** shown in Figure 11. It uses a 155 Mbps UTP-5 link between the monitor port 1 and the switch port A1 and a 25 Mbps link between the monitor port 2 and the switch port D1. Figure 11 also indicates the traffic flow direction.

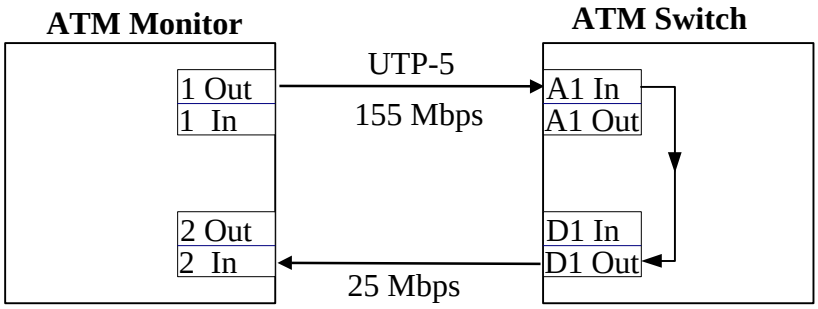


Figure 11. Test configuration for measurements of MIMO latency

In this configurations:

- CIT = 2.83 μs
- COT = 424[bits] / Output Link Rate = 424[bits] / 25.6 [Mbps] = 16.56 μs

We performed all our tests with 32-cell frames. One of the measurements used contiguous frames, i.e. cells of the test frame were transmitted back-to-back. In the rest of the tests, we introduce identical gaps (unassigned cells or cells of other frames) between cells of the test frame.

Table 2 presents measurement results for eight test runs, from which MIMO latency is calculated. The first test uses a contiguous test frame on input. All other tests use discontinuous frames on input, with gaps between cells of the test frame, as indicated in the second column. Our tests do not show any significant difference if gaps include unassigned cells or cells of other frames, which leave the switch through output links other than the one used by the test frames. The third, fourth and fifth columns present measurement results for the last cell delay, first cell delay and inter-arrival time between the first and the last cells, respectively. The sixth column includes calculated values for

$LILO_0$, as explained in Section 2, given a frame pattern on input. Here is how we calculate those values. For the first five tests, each cell entering an ideal switch has to

Table 2: (All times are in μs)

Test No.	Frame Pattern	Last cell CTD	1st cell CTD	1st cell to last cell inter-arrival time	$LILO_0$	LILO	FILO	MIMO
1	No gap	387.84	36.8	526.5	351.71	385.01	563.3	33.3
2	1-cell gaps	298.61	35.8	526.0	263.98	295.78	561.8	31.8
3	2-cell gaps	211.88	36.8	526.0	176.25	209.05	562.8	32.8
4	3-cell gaps	122.65	34.8	526.5	88.52	119.82	561.3	31.3
5	4-cell gaps	46.83	40.8	519.5	13.67	44	560.3	30.3
6	5-cell gaps	36.4	36.8	526.5	13.67	33.57	562.8	19.9
7	6-cell gaps	38.7	36.8	616.0	13.67	35.87	652.8	22.2
8	7-cell gaps	38.4	35.3	705.0	13.67	35.57	740.3	21.9

wait for transmission of the previously received cell to finish. Thus, on output we have back-to-back cells, i.e. a contiguous frame. Therefore, we can calculate FOT_0 for 32-cell frames in all those cases as:

$FOT_0 = 32 \times COT = 32 \times 16.56 = 530$ μs

In the first case:

FIT = 32 x CIT = 32 x 2.83 = 90.56 μs

Hence:

$LILO_0 = FOT_0 - FIT = 439.44$

In the second case:

FIT = (CIT + 1-cell gap) × 31 + CIT = 63 x 2.83 = 178.29 μs

Hence:

$LILO_0 = FOT_0 - FIT = 351.71$

Similarly in the cases with 2-cell gaps, 3-cell gaps and 4-cell gaps $LILO_0$ is calculated as 263.98 μs, 176.25 μs and 88.52 μs, respectively.

In the last three tests, the gaps on input are large enough that no cells have to wait on a previously received cell. In the case with 5-cell gaps, the first bit of the 32^{nd} (last) cell arrives at an ideal switch at time t, where

$t = (CIT + 5\text{-cell gap}) \times 31 = 6 \times CIT \times 31 = 526.38\ \mu s$

and then

$FOT_0 = t + COT = 526.38 + 16.5 = 542.88\ \mu s$

$FIT = (CIT + 5\text{-cell gap}) \times 31 + CIT = 529{,}21\ \mu s$

Hence:

$LILO_0 = FOT_0 - FIT = 13.67\ \mu s$

In the cases when the last cell doesn't wait on the previous received cell $LILO_0$ can be calculated also as:

$LILO_0 = COT - CIT = 13.67\ \mu s$

The eighth column shows FILO latency calculated, according to the expression (3) as the sum of terms in the third and the fourth column. In the last column MIMO latency values are obtained according to the expression (1) and (4).

Note that the switch latency is higher in the first five tests due to cell queueing. In the last three tests, the gap between the cells is large and there is no queueing. MIMO latency clearly reflects this effect.

4.4. Tests with Input Link Rate Lower Than Output Link Rate

We also performed tests using the configuration in Figure 11, but with the traffic flow in the opposite direction as indicated in the figure. Thus, this is the configuration with the input link rate lower than the output link rate. In this case, we have:

- CIT = Cell input time = 16.56 μs
- COT = Cell output time = 2.83 μs

We performed tests with 32-cell frames, with random idle periods between cells. Table 3 includes measurement data from two tests for which MIMO latency is also calculated. Since the input link rate is lower than the output link rate, $LILO_0$ is equal to 0.

Table 3. (All times are in μs)

Last cell CTD	1st cell CTD	1st cell to last cell inter-arrival time	FILO Latency	MIMO latency
32.0	31.0	535.0	566.0	15.44
32.5	33.0	1067.5	1100.5	15.94

The results in Table 3 show clearly that MIMO latency reflects the switch behavior and is not affected by the arrival pattern. On the other hand, it is shown that FILO latency is strongly affected by the arrival pattern. It can be observed that good agreement of MIMO latency values can be obtained using any of the two expressions for its calculation.

5. MIMO Latency of a Path

Consider a network path consisting of n components in a series. Subscript i is used for the latency of the i^{th} component and subscript Σ for the combination. Thus,

$$MIMO_i = LILO_i - LILO_{0i}$$

Similarly for a network path:

$$MIMO_\Sigma = LILO_\Sigma - LILO_{0\Sigma}$$

Since LILO is additive:

$$LILO_\Sigma = \Sigma\, LILO_i$$

The above three relationships lead us to the following identity:

$$MIMO_\Sigma + LILO_{0\Sigma} = \Sigma\,(MIMO_i + LILO_{0i})$$

Or

$$MIMO_\Sigma = \Sigma\, MIMO_i + \Sigma\, LILO_{0i} - LILO_{0\Sigma}$$

This relationship allows us to compute MIMO of a series of components from the measured MIMO values of individual components. Note that $LILO_{0i}$ and $LILO_{0\Sigma}$ can be computed given the input pattern and the input/output speeds. We illustrate this with a few examples.

Example 1:

Consider the configuration shown in Figure 12 consisting of two switches interconnected via a wire. All links and ports are 155 Mbps (OC-3). The input frame is composed of two cells with a gap of three cell times. Let us suppose that each switch introduces a MIMO equal to c, where c is the cell time at OC-3. Also, for simplicity assume that the wire between the switches also introduces a MIMO of c. (Other wire lengths can be handled similarly).

Since all input/output speeds are same, an ideal switch will produce zero LILO latency. Hence,

$$LILO_{0i} = 0$$
$$LILO_{0\Sigma} = 0$$
$$MIMO_{\Sigma} = \Sigma\ MIMO_i = c + c + c = 3c$$

That is, the MIMO latency is simply the sum of individual MIMO latencies.

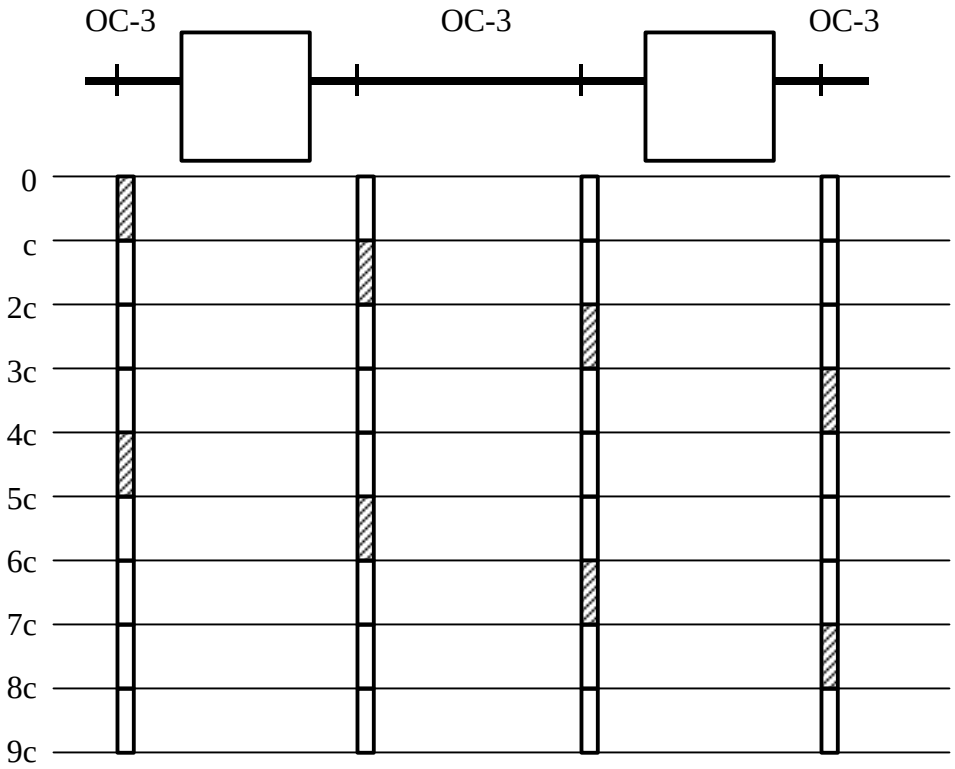


Figure 12. MIMO aggregation for Example 1.

Example 2:

Consider the configuration shown in Figure 13. This is similar to the configuration of Example 1, except that the intermediate link is 51.4 Mbps and therefore, introduces a delay of 3c.

In this case, the first switch has an input speed of OC-3, while the output speed is OC-1. An ideal switch with these I/O speeds will produce a LILO latency of 2c, where c is the cell time at OC-3. That is,

$$LILO_{01} = 2c$$

For the wire as well as the second switch, the input speed is equal to or less than the output speed and so the $LILO_0$ is zero:

$$LILO_{02} = 0$$
$$LILO_{03} = 0$$

If the whole network is replaced by a single ideal switch, that switch will have an input speed of OC-3 and an output speed of OC-3 and therefore, will have a zero LILO latency.

That is,

$$LILO_{0\Sigma} = 0$$

Using the above values, we get:

$$MIMO_{\Sigma} = \Sigma\, MIMO_i + \Sigma\, LILO_{0i} - LILO_{0\Sigma} = (c+3c+c) + (2c+0+0) - 0 = 7c$$

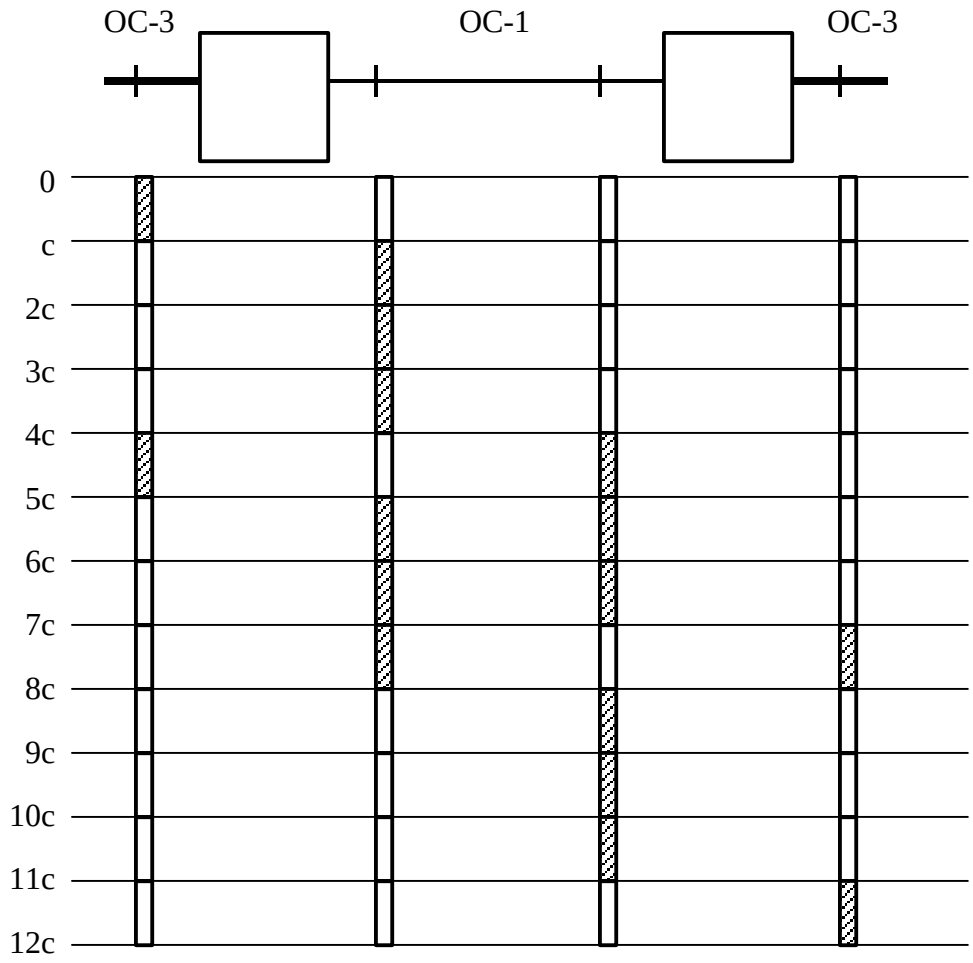


Figure 13. MIMO aggregation for Example 2.

Conclusion:

Frames on ATM networks are generally discontinuous and so the usual measures of latency such as FILO, LILO, LIFO, or FIFO give results that do not reflect the delay introduced by the network components. The MIMO latency, which is defined as the difference between the LILO latency of a component and that through an ideal component, always gives the correct component latency. MIMO latency can be

measured using data produced by current ATM monitors. Also, it is possible to compute MIMO latency of a path from the latencies of the components in the path.

Acknowledgements:

We would like to thank members of the ANSI T1A1.3 committee who, during our December 98 presentation to the committee, encouraged us to work on aggregation of MIMO latency. Also, VFILO and VLIFO discussions by Ken Glossbrenner (AT&T), Geoffrey Garner (Lucent) and others during the January 99 ANSI T1A1.3 meeting were helpful in getting MIMO aggregation results.

Reference:

[1] ATM Forum/BTD-TEST-TM-PERF.00.11 (Draft) February 1999
[2] ATM Forum Traffic Management Specification Version 4.0
[3] CCITT Recommendation X.135, "Speed of Service (Delay and Throughput) Performance Values for Public Data Networks when Providing International Packet Switched Service," 1992
[4] ITU-T Recommendation I.356,"B-ISDN ATM Layer Specification," ITU-Study Group 13, Geneva, 1995