
ATM Forum Document Number: ATM_Forum/96-0180

Title: Scope For ATM Forum's Performance Benchmarking Work Item

Abstract:

This contribution discusses the scope of the ATM Forum's test working group. It also presents update on he performance metrics proposed in our December 1995 contribution.

Source:

Raj Jain, Bhavana Nagendra, and Gojko Babic
The Ohio State University

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

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

Date: February, 1996, Los Angeles

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

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.

SUMMARY OF DECEMBER 1995 DISCUSSION:

In October 1995 meeting of AF-TEST, it was agreed that performance benchmarking is essential and that instead of forming a separate "birds of a feather (BOF)", AF-TEST will schedule presentations on performance issues. As a result, three presentations were made in December 1995 meeting.

During the presentation, some issues were raised about what exactly should be the scope of ATM Forum's work in this area. Five different views were expressed by five different people. They were all asked write up and present their view of the scope for February meeting.

This contribution fulfills that commitment from our side. Many of the ideas presented here are enhancements of ideas presented in

the two earlier meetings.

SCOPE OF ATM FORUM'S WORK ON PERFORMANCE BENCHMARKING:

Performance benchmarking is related to user perceived performance of ATM technology. For the success of ATM technology, it is important that the performance of existing and new applications be better than that on other competing networking technologies. In other words, goodness of ATM will not be measured by cell level performance but by frame-level performance and performance perceived at higher layers.

Most of the Quality of Service (QoS) metrics, such as cell transfer delay (CTD), cell delay variation (CDV), cell loss ratio (CLR), and so on, may or may not be reflected directly in the performance perceived by the user. For example, while comparing two switches if one gives a CLR of 0.1% and a frame loss ratio of 0.1% while the other gives a CLR 1% but a frame loss ratio of only 0.05%, the second switch will be considered superior by many users.

ATM Forum and ITU have standardized the definitions of QoS metrics. We need to do the same for higher level performance metrics. Without a standard definition, each vendor will use their own definition of common metrics such as throughput and latency resulting in a confusion in the market place. Avoiding such a confusion will help buyers eventually leading to better sales resulting in the success of the ATM technology.

GOALS OF THE ATM FORUM WORK:

-
- a. ATM Forum should define higher level performance metrics that will help a user to compare various ATM equipment (and possibly non-ATM equipment) in terms of performance.
 - b. The metrics should be such that they are independent of switch or NIC architecture. The same metrics should apply to all architectures.
 - c. The metrics should help the user predict the performance of their application or design their network configurations to meet their required performance.
 - d. ATM Forum should develop precise methodology for measuring these metrics. The methodology includes a set of configurations and traffic patterns. This will allow vendors as well as users to conduct their own measurements and come up with comparable results.
 - e. The key goal of this effort is to enhance the marketability of ATM technology and equipment. Any other extension of the above that helps in achieving that goal can be added later to this list.
 - f. The benchmarking should eventually cover all classes of service. Many past performance measurements concentrated on CBR service. We need to extend those to real time VBR, non-realtime VBR, ABR, and UBR. This may be phased such that most important service classes are covered first and less important ones are added later.
 - g. The metrics and methodology for different service classes can be different.

h. The benchmarking should cover as many protocol stacks as possible. For example, data traffic may use UBR or ABR service class. Some ATM networks (switches) may offer one or both classes. The user may care more for the application throughput rather than the underlying mechanism used. The performance is, therefore, measured on several alternative protocol stacks.

i. The benchmarking work should include performance of network management, connection setup, along with normal data transfer.

NON-GOALS OF THE ATM FORUM WORK:

a. ATM Forum is not responsible for conducting any measurements. This is similar to other tests such as conformance testing. Tests are defined but not conducted by the standards bodies.

b. ATM Forum is not responsible for certifying any measurements. Again this is not different from that for conformance testing. Certification has legal issues. Only defining metrics and methodologies has no legal consequences over and above what ATM forum is already doing.

c. ATM Forum is not responsible for setting particular performance thresholds such that equipment below those thresholds are called "unsatisfactory." For example, whether a switch which loses 50% of packets is good or bad may depend upon the applications and cost. The users and designers should be free to make their own cost-performance tradeoffs. Setting such thresholds inhibits such trade-offs. For example, if we set the packet loss threshold at, say 99%. This will prevent manufacturers from making low-cost switches that may be good enough for many applications. Generally, users have flexibility to design their applications that they get satisfactory performance in spite of lower grade equipment (for example, by forward error correction or retransmissions in case of packet errors and loss). In other words, ATM Forum should not set any requirements that prevents reducing cost while reducing performance as well.

As another example of the above argument consider the problem of setting a delay value for ATM switches. Lets say a delay of 30 ms is set as the standard. Switch manufacturers are compelled to manufacture switches with that delay value. But what prevents them to manufacture switches with various delays, depending on applications and requirements? For some applications a switch with a delay of 40 ms might suffice. Such switches are cheaper and needn't be precluded from the market. At the same time, manufacturers should continue to invest on better switches (with lower delays). So applications should not be bound by numbers as the switch manufacturers will be compelled to manufacture switches with the prescribed parametric values. This will hurt competition and will bring in legalities. It would also hurt progress into better technology. This holds good for all the other metrics of the switches like throughput, latency and the like.

AN EXAMPLE PROPOSAL:

The metrics, methodologies, and traffic patterns discussed below are presented as a starting point for discussion. These are enhancements of those in our December 95 contribution. During

December 95 meeting, a number of good suggestions were made. We have tried to update the presentations with those suggestions.

Since this is a new endeavor, it is limited in several respects. This particular proposal concentrates on the data traffic (ABR and UBR service classes) since that is expected to be the bulk of traffic on ATM networks initially. Other service classes will be added later.

TRAFFIC PATTERNS:

We define two types of traffic based on application's response to network congestion.

- a. Open loop traffic
- b. Closed loop traffic

Case a) With open loop traffic, the application does not reduce its load when the network performance degrades in terms of throughput or delay. Periodically occurring events generally lead to such traffic patterns.

Case b) With closed loop traffic, the application does slow down when the network response is slower. In many client-server applications, clients will not generate new requests if the previous requests have not been served. TCP/IP, which is expected to be a big part of the ATM market at least initially, is an example of a closed loop application. If the network performance degrades and TCP packets are delayed excessively or lost, TCP will reduce its window and resulting load on the network.

UDP is an example of an open loop traffic. The following figure shows some of the application layer protocols that run on TCP and UDP, respectively.

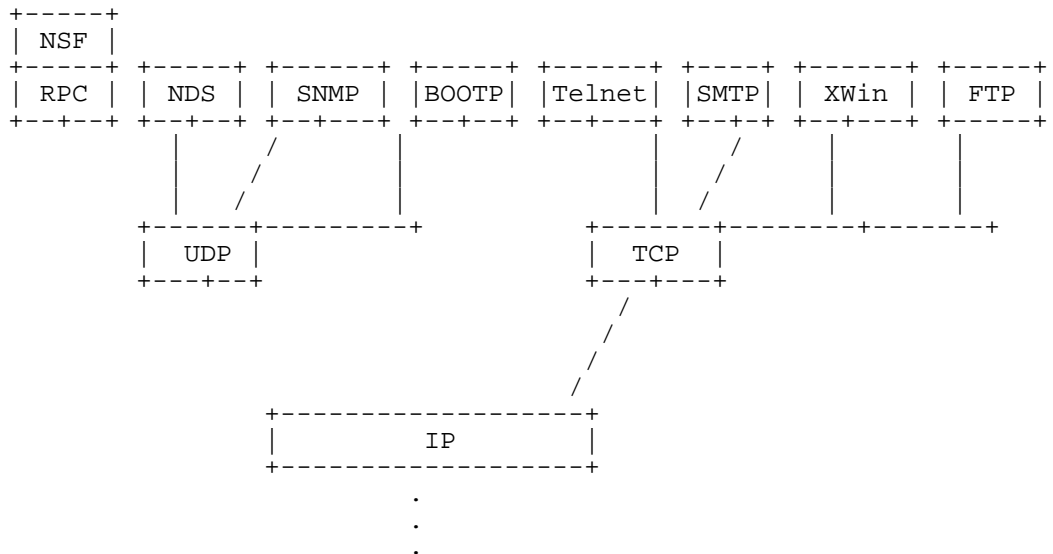


Figure 1 - The protocol stack above UDP and TCP

One reason for differentiating between open-loop and closed-loop traffic patterns is that ATM layer has to provide proper resource control for open-loop traffic. The closed-loop traffic can live with looser controls. For example, TCP can work over UBR or ABR. It can work even under high loss conditions.

WHICH LAYER TO MEASURE THE PERFORMANCE?

The performance can be measured at several layers (above ATM layer), for example, network (e.g., IP), transport (e.g., TCP), application (e.g., FTP). At each layer, several alternative stacks are possible. For example, IP can use "Classical IP over ATM" (RFC 1577) or "LAN Emulation (LANE)."

As shown in Figure 2, performance could be measured at any of the three layers: AAL5, RFC 1577/LANE and IP.

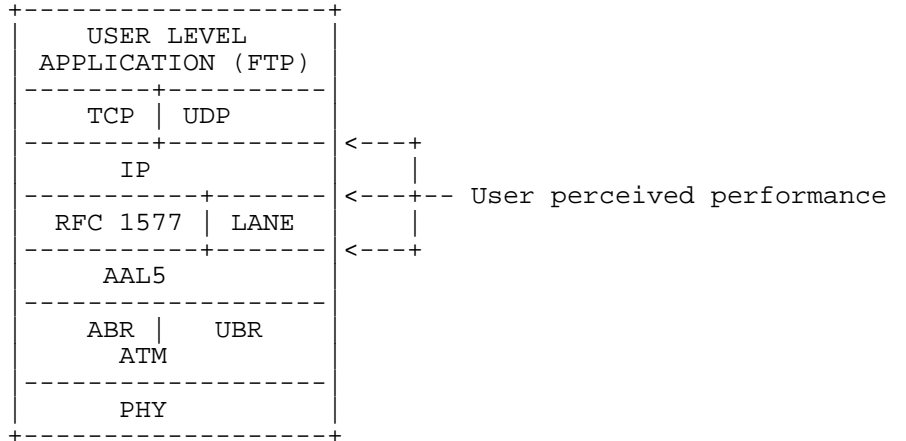


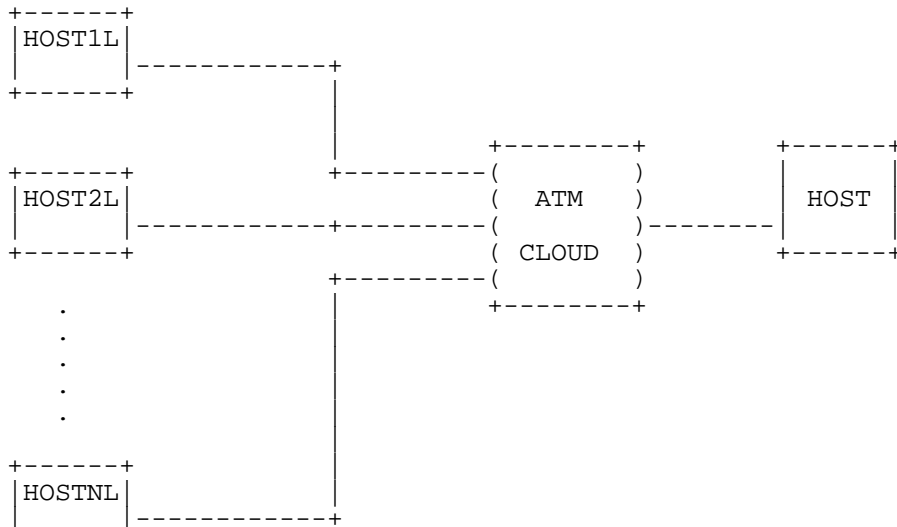
Figure 2 - Examples of measurement alternatives

At the AAL5 layer, one can measure ATM performance, but cannot compare technologies. At the LANE/RFC 1577 layer or at the IP layer, different technologies can be compared.

TEST CONFIGURATIONS:

We propose considering the following two configurations. These will be used in defining the metrics. The hosts are connected by a ATM cloud and can be a single switch or a collection of switches.

Configuration A: N inputs and 1 output



+-----+

Figure 3 - A configuration with N inputs and a single output

Configuration B: N inputs and N outputs

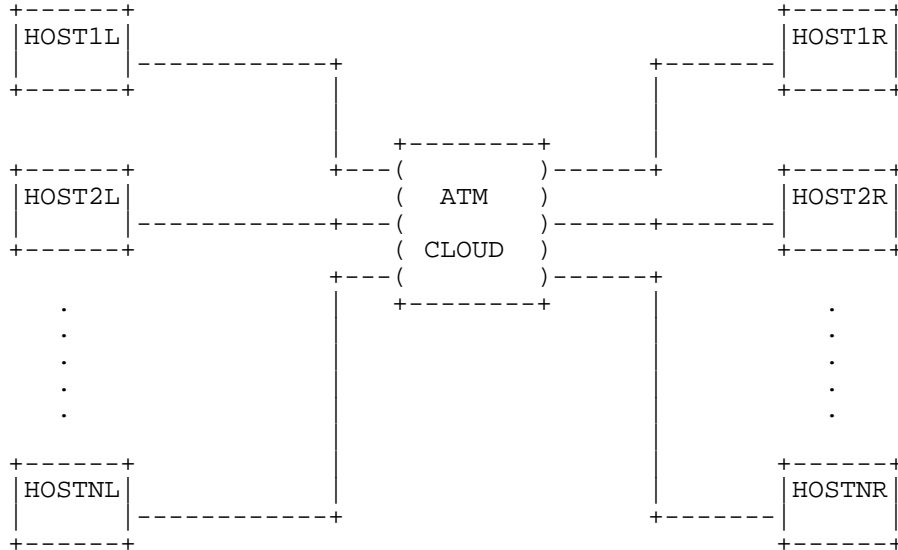


Figure 4 - A configuration with N inputs and N outputs

Here are illustrations of tests that can be performed using the above configurations.

For Configuration A, increase load symmetrically on N ports and measure the output. This configuration represents an overloaded condition as N inputs are flowing into one switch and there is a single output. Such a condition would result in lower throughput, increased frame loss, lower back to back burst size, higher latency etc. Fairness can also be measured, i.e. if the switch discriminates the sources.

For Configuration B, the traffic can be sent in the following 3 ways.

- i) Host_iL sends traffic to Host_iR all of its traffic, $i = 1, 2, \dots, N$.
- ii) Host_iL sends traffic to Host_jR, $j = 1, 2, \dots, N$, $1/N$ of traffic, $i = 1, 2, \dots, N$.
- iii) Same as i), but with bidirectional traffic.
- iv) Same as ii), but with bidirectional traffic.

N needs to be determined, overloading depends on the number of sources.

Increase load symmetrically on all ports and measure on the corresponding outputs and this configuration can be used to measure the fairness of the switch.

PERFORMANCE METRICS:

We propose that the metrics be grouped as follows: - General metrics - Protocol-Stack specific metrics - Traffic Management metrics - Network Management metrics

General Performance Metrics : These metrics apply to most ATM networks and are not protocol specific. The tests for these metrics effectively characterize the basic features of the switch.

Protocol-Stack Specific Metrics : These metrics apply to particular protocol stacks and need only be measured and tested if particular protocols are being used. Examples, of such protocols are RFC1577 and LANE, as discussed earlier.

Traffic Management Metrics: These measure ability of the switches to avoid overload and to efficiently and fairly resolve contention among various VCs when there is overload.

Network Management Metrics : These metrics are defined to aid characterization of the switch in responding to network management requests.

Some of the discussion below is from RFC 1242 and its current version (an internet draft) [Bradner] and is a modification of [Jain, Nagendra]. We are of course, open to comments, suggestions, and discussion, for tailoring these metrics and configurations.

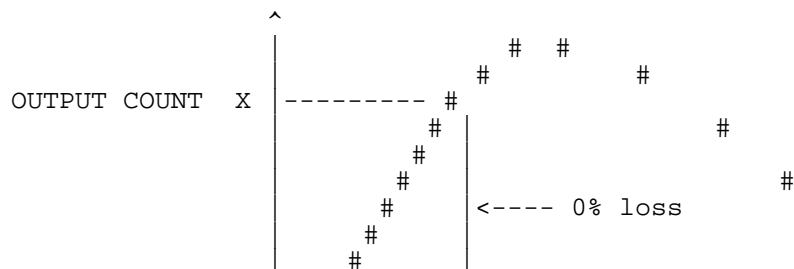
GENERAL PERFORMANCE METRICS

1. Throughput

The throughput can be measured for UBR case (open loop) and ABR case (closed loop).

For UBR, throughput is defined as the maximum rate at which none of the frames are dropped by the ATM switch. Essentially we are looking at the behavior of a perfect switch which works with an efficiency of 100%. Data traffic is passed through the switch and then the frames that are transmitted by the switch are counted. If the input and the output count are the same then the load is increased and the test is conducted again. The throughput is the highest load at which the count of the output frames equals the count of the input frames. A graph of input count vs output count can be shown. Instead, the load can be kept constant and the frame size can be varied and its effect on the throughput can be studied.

A model graph of input count vs output count would be: Point X defines the throughput without loss.



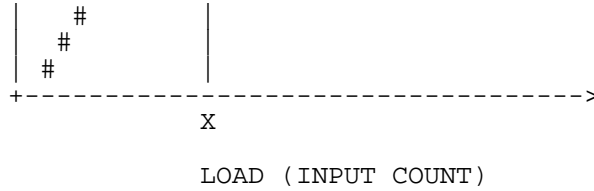


Figure 5 - Graph of output count vs load (input count) for UBR

Throughput can be expressed in bits/sec, frame/sec or cells/sec. Cells/ sec is not a good unit as cells in the ATM layer have significant overhead and relatively low overhead at AAL5. It is preferred to express the same in bits/sec, because expressing it in frame/sec would involve the frame size which is a variable. However bits/sec and frame/sec are related by the following equation.

$$\text{Throughput (bits/sec)} = \text{Throughput (frame/sec)} * \text{Average frame size (bits)}$$

For ABR, we propose two throughputs definitions:

- o without loss
- o after congestion mechanism is triggered.

The definitions and tests will now be explained.

Data traffic is passed through the switch from the sources and then the frames that are transmitted by the switch are counted. All frames are of the same size (Frame size is open to discussion). If the input and the output count are the same then the load is increased and the test is conducted again. The throughput without loss is the highest load at which the count of the output frames equals the count of the input frames. When the load is increased beyond a certain point, the congestion mechanism is activated and sends a warning to the sources to decrease load rate. The system will stabilize at some point (meaning that the input count and output count are identical) and that load defines the throughput after congestion mechanism is triggered. Instead, the load can be kept constant and the frame size can be varied and its effect on the throughput can be studied.

Throughput in configuration A equals or is close to the capacity of the sink. It is noted that a well behaved switch would allow equal load from all sources without giving preference to any source.

2. Latency

We use the following table to define the beginning (at input) and ending (at output), of the time interval, for which latency has to be defined. The definition of latency can be defined in 4 ways depending on the time interval considered. We wish to deviate from the usual definition of latency which is stated in [Jain and Nagendra]

ON INPUT			ON OUTPUT		
SL#	FIRST BIT	LAST BIT	FIRST BIT	LAST BIT	
1	X		X		
2		X	X		

3	X			X
4		X		X

Table 1 - Four ways of defining the time interval for latency

Definition 1 and 2 in the above table is not appropriate since the whole frame is the user's concern and the complete frame has not been received as yet. Definitions 3 and 4 are appropriate for the user but in case 3, the latency is dependent on the message length. So it appears that definition 4 is a good measure for switch latency measurement. Hence we choose to define latency as follows :

The time interval starting from when the last bit of the input frame is transmitted and ending when the last bit of the output frame is received by the host.

This is valid for all types of devices both cut-through devices and store and forward devices and the measure cannot be negative for cut-through devices as stated earlier. This helps in treating the devices uniformly and not be bothered by the internal architecture.

Latency depends on the load. Hence it has to be measured in two extreme loads.

- o zero load
- o throughput load

Other loads may also be considered.

The time at which the frame is fully transmitted is recorded (timestamp A). The receiver logic in the test equipment should be able to the tag information in the frame stream and record the time at which the entire tagged frame was received (timestamp B).

$$\text{Latency} = \text{Timestamp B} - \text{Timestamp A}$$

The reporting format would be load and resultant latency for each frame size.

3. Frame loss rate

Percentage of frames that should have been forwarded by the switch under steady state traffic that were not forwarded due to lack of resources.

Frame loss rate is an interesting metric only under open loop (UBR case), conditions as under closed loop conditions, the network will warn the source of potential losses (congestion mechanism). When congestion mechanism is activated, frame loss is possible, but it is not constant.

This measurement reports the performance of the switch at an overloaded state. The device might lose frames that contain routing information and this may further reduce the performance as more frames need to be retransmitted. The frame errors could be CRC errors and/ or cell termination errors.

$$\text{Frame loss rate} = (\text{input_count} - \text{output_count}) / \text{input_count}$$

Configuration A) The first trial should be run at the load that

corresponds to 100% of the maximum rate for the frame size from N sources. The load is progressively decreased until there are two successive trials with no frame loss.

Configuration B) The switch receives traffic from N sources simultaneously at maximum rate for the frame size. The output is measured at the N outputs. The load is progressively decreased until there are two successive trials with no frame loss.

The results of the frame loss test should be reported as a graph of % loss vs load.

4. Back-to-Back Burst Size

Fixed length frames presented at a rate such that there is the minimum legal separation between frames over a short to medium period of time, starting from an idle state. This determines buffering capabilities of the ATM switch in hand. NFS, remote disk backup systems like rdump, and remote tape access systems, can be configured such that a single request can result in a block of data being returned, as much as 64k octets. The length of the frame is to be decided.

Case a) Bursts of frames with minimum inter-frame gaps are sent to the switch from the sources and the number of frames that have been forwarded by the switch to the single host is counted. If there are no losses or congestion mechanism is not triggered (ABR) then the length of the burst is increased and the test is rerun. The back-to-back burst size is the longest burst that the device will handle without the loss of any frames. It measures the extent of data buffering in the switch.

5. Call establishment time

This is the time taken to setup a connection with the destination by the calling party.

For short duration VCs, call establishment time is an important part of the user perceived performance. The time between the submission of a "call request" and the reception of the corresponding "ready indication" is defined as the call establishment time.

The call establishment time is measured at zero load and load corresponding to the throughput. Other loads may also be considered.

TRAFFIC MANAGEMENT METRICS:

1. Load Control Latency: A set of VCs are established. After the system reaches the steady state, the load on one VC is suddenly increased, the time for the system to reach the steady state again is measured. Similarly, when the load is decreased, the time to reach steady state is measured.

2. Burst Throughput: Frames are sent at differing burst (frame burst) sizes and the steady state throughput is measured. Depending upon the underlying service class (UBR, ABR), the bursty performance may be different than steady state performance. This is particularly important for request-response (client-server) applications.

3. Throughput in the Presence of Higher Priority Traffic: The throughput of ABR traffic is measured when a VBR VC shares the path with data traffic. The characteristics of the VBR traffic need to be clearly specified.

4. Fairness: The fairness can be discussed for both the configurations in figures 3 and 4.

In the configuration A, N sources are connected to a single host through a switch. Increase load symmetrically on N ports and measure the output. The switch might cut out a host and only allows traffic from the remaining hosts and so the fairness of the situation can be studied.

In the configuration B, each of the N hosts is connected to either 1 host or all the N hosts on the output through a switch. Increase load symmetrically on N ports and measure the output on the corresponding hosts. If the traffic in all the lines is not equal, then the switch is partial and the fairness criteria has been violated.

NETWORK MANAGEMENT METRICS: [To be discussed]

APPLICATION SPECIFIC PERFORMANCE METRICS: [To be discussed]

BIBLIOGRAPHY:

[Bradner] Scott Bradner, "Benchmarking Methodology for Network Interconnect Devices", Internet Draft.

[Mandeville] Robert Mandeville, European Network Laboratories, Data Comm Magazine, March 1995, p 69.

[Wakid] Wakid et al, "Architectures for BISDN Networks : A Performance Study", Advanced Systems Division, National Institute of Standards and Technology, (301)-975-4855, <http://www.hpcc.gov/blue94/section.4.7.html>

[LANQuest] ATM Cell Congestion Loss Across Switch (CCLAS) Throughput Analysis, LANQuest Labs, (408) 894-1000.

[RFC1577] M. Laubach, Classical IP and ARP over ATM, RFC 1577, Jan 1994

[SNCI] Scott Bradner, "The 1995 Ethernet to ATM Evaluation", SNCI

[Mier] Mier and Smithers, "ATM to the Desktop", Product Testing, Communications Week, September 25, 1995

[RFC1242] Scott Bradner, "Benchmarking Terminology for Network Interconnection Devices", RFC 1242, July 1991.

[Rowe] Martin Rowe, "Wealth of ATM Testers - Answers Most Needs", Test and Measurement World, Sept 1995, p 55

[Krivda] Cheryl D. Krivda, "Analyzing ATM Adapter Performance: The Real-World Meaning of Benchmarks," <http://www.efficient.com/dox/EM.html>

[Jain] Raj Jain, "Performance Benchmarking BOF," AF-ALL/95-1347, October 1995.

[Jain, Nagendra] Raj Jain, Bhavana Nagendra, "Performance

Benchmarking of ATM Switches", AF-TEST/95-1662, December 1995

Note: All our past ATM forum contributions and presentations are available on-line:

<http://www.cse.wustl.edu/~jain/>