Layer 4 Switching Experiments for Burst Traffic and Video Sources in ATM

Virgil DOBROTA, Daniel ZINCA, Cristian Mihai VANCEA, Aurel VLAICU
Technical University of Cluj-Napoca, Department of Communications
26 Baritiu Street, 3400 Cluj-Napoca, Romania
Tel: +40-64-191689, 195699/208 Fax: +40-64-191689
E-mail: {Virgil.Dobrota, Daniel Zinca}@com.utcluj.ro

EXTENDED ABSTRACT

This paper is focused on the results of a Layer 4 switching experiment, aiming to evaluate the performances at the interface between the applications and the nonblocking stream-oriented sockets in TCP/IP. One major objective is to demonstrate that some of the traffic models, mainly designed for ATM sources, could be also applied at transport layer. Prior to our studies on Fast Ethernet and ATM traffic parameters, presented at LANMAN'96 and LANMAN'98, we are trying to obtain better results for burst traffic and video sources by involving departure schedules for cells or frames. This means that the applications should not send the information directly to the sockets without taking into account the behaviour of TCP/IP entities within a broadband network. We have selected the real-time experiments, carried out on both Classical IP over ATM and IP over Fast Ethernet, in order to get the answers to the following questions:

1. *Is it possible to apply ATM traffic models to the TCP/IP environment?*
2. *Which are the advantages of Layer 4 switches implemented by software for point-to-point, point-to-multipoint and broadcast services?*
3. *What is the influence of the lower layers technologies against the transport layer exchange of information?*

The first paragraph is devoted to burst traffic generated by ON/OFF sources of constant throughput. A Matlab-based scheduler is able to determine the number of ON cells to the number of OFF cells ratio, for every burst, until the transmission process is completed [Dob99].

Due to different types of correlation between successive frames, the video services are mainly different than voice and data, involving discrete-state continuous-time Markov models. *M1_X* is the unidimensional model, whilst *M2_X* is the bidimensional one. *X* represents the type of experiment: *(A)* Probability of being in a given state versus average throughput (state *i*, where *i=0,1,...N* for unidimensional model, or state *(i,j)*, where *i=0,1,...N-low* and *j=0,1,..N-high*, for bidimensional model); *(B)* Average throughput D versus activation/deactivation rates (a , b for unidimensional model, respectively a , b , g , d for bidimensional model); *(C)* Average throughput D versus probability of being in a given state. A detailed description of these video models is given in [Dob98b], [Dob98c].

The testing configuration included four workstations connected either to ATM 25.6 Mbps ports of VIRATAswitch 1000, either to Fast Ethernet 100 Mbps ports of HP ProCurve hub. The most powerful station within the tested network was based on Intel's Pentium II/400 MHz, running the server and acting as a Layer 4 switch. The client software was installed on three different workstations (with Celeron 366 MHz, Pentium 233 MHz MMX and Pentium 120 MHz). Note that these machines were not connected simultaneously to ATM and Fast Ethernet, in order to avoid the uncontrolled influences.

Note that the application's buffer for sends and the application's buffer for receives are different than those of Windows Sockets related to TCP/IP. The last ones could be modified through *setsockopt* function (integer values SO_SNDBUF and SO_RCVBUF). We tried also the influence of disabling the Nagle algorithm (by enabling TCP_NODELAY option), but the general suggestion is to leave it enabled (by default).

The evaluation accuracy of the proposed software tool (client and server) is given by the clock period of the CPU (2.5 nanoseconds at Pentium II/400 MHz). The measurement of the sends and receives on the sockets is also dependent on RDTSC *(Read Time Stamp Counter)* and other instructions included in the loop. Obviously the processes are guided by the TCP/IP entity, as we rely on the Windows Sockets *select* function to determine the status of the sockets and to perform synchronous I/O.

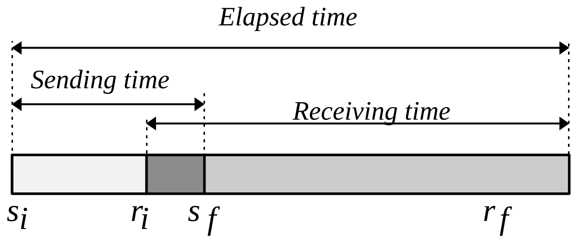


Figure 1. Four time stamps for measuring the sending, receiving and elapsed times

Sometimes it is more efficient to send the information using a model. However, the general throughput for point-to-point service could be higher (about 15 % for *Model1*) or lower (about 20 % for *Model2*) compared to the case of classical one-block sending. This observation is not valid for point-to-multipoint and broadcast services because in this case it seems that any model generates better performances.

The model-based transmission could also reduce the network congestion. For instance, the experiment of broadcasting video frames to three workstations, including the transmitter, shows that the serving rate of about 4.94…5.11 Mbps (Pentium II/400) is comparable to the incoming rate of any station. We come to the conclusion that the CPU's frequency of the sender does not have a great influence at the level we are discussing in this paper. The elapsed time is less than 5% higher for Pentium 120 MHz, compared to Celeron 366 MHz, in a 3-station broadcast trial.

CONCLUSIONS

1. Some of the ON/OFF and video models, usually describing the departure schedules for ATM sources, could be used also for nonblocking stream-oriented sockets in TCP/IP.
2. The Layer 4 switching has advantages due to its status information about the sockets traffic. By exploiting the specific non-linear behaviour of TCP/IP-based networks, it can reduce the traffic congestion. The resulting switching and arrival schedules are significantly different than the departure ones.
3. The highest throughput, calculated at the application/Windows Sockets interface, is less than 10 Mbps for 25.6 Mbps ATM, and less than 20 Mbps for 100 Mbps Fast Ethernet.

FUTURE WORK

Although the voice models have been also studied, by the time the paper was submitted the experiments were under progress. The overall performance of the Layer 4 switching is expected to be improved by adding Layer 2-3 switching on the same machine. It is for further work to determine the optimum model by anticipating the consequences of the self-similarity behaviour of the network.

REFERENCES

[Dob99] V. Dobrota, D. Zinca, "Experimental Results of Traffic Models for Burst Data and Voice Sources in ATM Networks", *ACTA Tehnica Napocensis*, ISSN 1221-6542, Vol.39, No.1, 1999, pp. 5-12.

[Dob98a] V. Dobrota, D. Zinca, "Traffic Models for Data, Voice and Video Sources in ATM", *Proceedings of the Symposium on Electronics and Telecommunications ETc'98*, Timisoara, Sept.17-18, 1998, Vol. 1, pp. 166-170.

[Dob98b] V. Dobrota, *Digital Networks in Telecommunications. Volume 2: B-ISDN with ATM, SS7*, Mediamira Science Publishers, Cluj-Napoca 1998

[Dob98c] V. Dobrota, D. Zinca, A. Vlaicu, K.Pusztai, "Evaluation of ATM Traffic Parameters in Heterogeneous Networks", *Proceedings of the 9th IEEE Workshop on Local and Metropolitan Area Networks LANMAN'98*, Banff, Alberta, Canada, May 17-20, 1998, pp. 304-309

[Zin96] D. Zinca, V. Dobrota, M. Cosma, A. Vlaicu, "Software Traffic Analyzer and Frame Generator for IEEE 802.3u", *Proceedings of the 8th IEEE Workshop on Local and Metropolitan Area Networks LANMAN'96*, Potsdam, Germany, August 25-28, 1996, pp. 243-248

[***] *http://193.226.6.190/people/dobrota/book9.htm*

Layer 4 Switching Experiments for Burst Traffic and Video Sources in ATM

V. Dobrota, D. Zinca, C. M. Vancea, A. Vlaicu
Technical University of Cluj-Napoca
26 Baritiu Street, 3400 Cluj-Napoca, Romania
E-mail: Virgil.Dobrota@com.utcluj.ro
http://www.utcluj.ro/utcn/eltc

SCHEME OF TALK

- Models for burst traffic and video sources
- Testing configuration and files
- Experimental results
  - Classical IP over ATM
  - Fast Ethernet
- Conclusions
- Future work

Models for burst and video sources

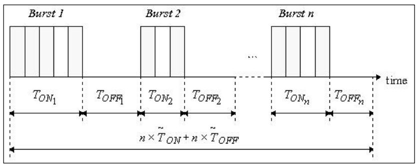


Figure 1. Geometrical distribution of burst traffic

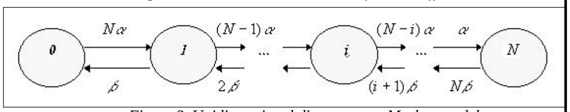


Figure 2. Unidimensional discrete-state Markov model

Test files

Figure 3. Testfile1 (7,990 bytes)

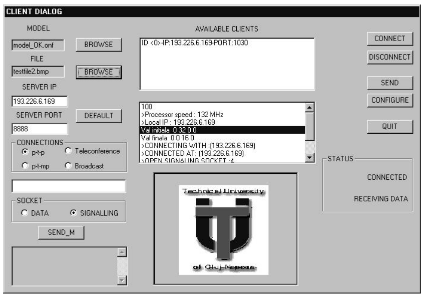


Figure 4. Screen capture of the client's GUI used as Testfile2 (240,118 bytes)

Departure schedules

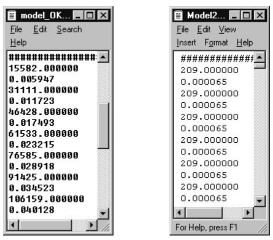


Figure 5. Model 1 and Model 2 for burst traffic: the 1st article is the number of bytes during the 1st burst (ex:15582 B) , the 2nd article is the duration of the 1st period ON+OFF (ex: 0.005947 seconds) etc.

Results for Model 1

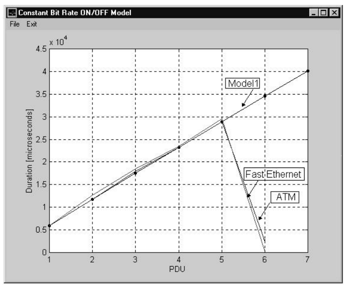


Figure 6. The sending TCP entity could follow the Model 1 for both Classical IP over ATM and IP over Fast Ethernet

Results for Model 2

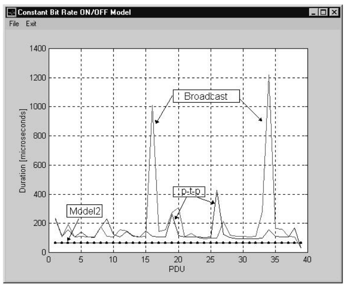


Figure 7. The sending TCP entity could not follow the Model 2 for all the transport services

LANMAN'99, 21-24 November 1999, Sydney SLIDE 7

Results for M2_B (video)

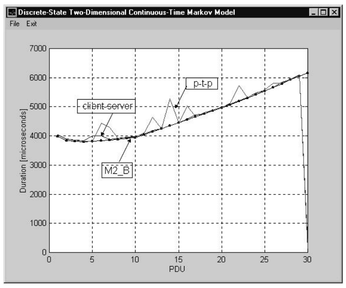


Figure 12. The sending TCP entity could generally follow M2_B but there are some differences at the actual departure schedule of p-t-p. Each PDU has 7990 bytes.

LANMAN'99, 21-24 November 1999, Sydney SLIDE 8

Numerical results - burst traffic

point-to-point

Interval	Model	Measured time [μ s]	Throughput [Mbps]
- ELAPSED	-	252512…260383	7.37…7.60
	1	223194…236111	8.13…8.60
	2	309603…326569	5.88…6.20
$s_f - s_i$ SEND	-	35987…39626	48.47..53.37
	1	90362…91211	21.06..21.25
	2	265194…277370	6.92..7.24
$r_f - r_i$ RECV.	-	195191…208155	9.22..9.84
	1	200324…220329	8.71..9.58
	2	305731…322555	5.95..6.28
$r_i - s_i$	-	52227…57321	N.A.
	1	14465…22870	N.A.
	2	3652…4014	N.A.
$r_f - s_f$	-	216524…220756	N.A.
	1	132100…144899	N.A.
	2	44384…49198	N.A.

broadcast

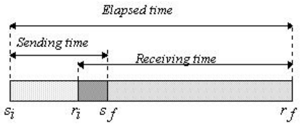
Interval	Model	Measured time [μ s]	Throughput [Mbps]
$r_f - s_i$ ELAPSED	-	434926…436824	4.39… 4.41
	1	409365…418930	4.58… 4.69
	2	419113…431807	4.44 …4.58
$s_f - s_i$ SEND	-	20408…20534	93.54..94.12
	1	87871…88381	21.73..21.86
	2	191061…257949	7.44…10.05
$r_f - r_i$ RECV.	-	405426…408241	4.70…4.73
	1	398744…402074	4.77…4.81
	2	416132…428589	4.48…4.61
$r_i - s_i$	-	28583…29499	N.A.
	1	10621…16856	N.A.
	2	2980…3218	N.A.
$r_f - s_f$	-	414391…416416	N.A.
	1	321494…330549	N.A.
	2	161163…240746	N.A.

LANMAN'99, 21-24 November 1999, Sydney SLIDE 9

Numerical results - video sources

broadcast

Interval	Station	Measured time [μ s]	Throughput [Mbps]
$r_f - s_i$ ELAPSED	Cel366	401870…405505	4.72…4.77
	P233	N.A.	N.A.
	P120	N.A.	N.A.
$s_f - s_i$ SEND	Cel366	138185…145189	13.20..13.87
	P233	-	-
	P120	-	-
$r_f - r_i$ RECV.	Cel366	389869…393386	4.87..4.91
	P233	391094…394638	4.85..4.90
	P120	390398…392943	4.88..4.91



LANMAN'99, 21-24 November 1999, Sydney SLIDE 10

Serving time for Layer 4 switch

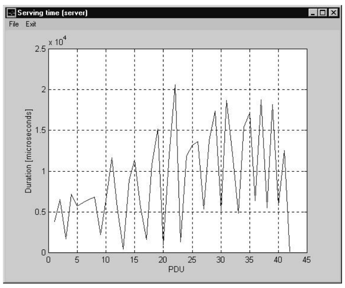


Figure 15. 3-station broadcast. The incoming traffic is the result of the model M2_B for video sources.

LANMAN'99, 21-24 November 1999, Sydney SLIDE 11

Arrival schedule

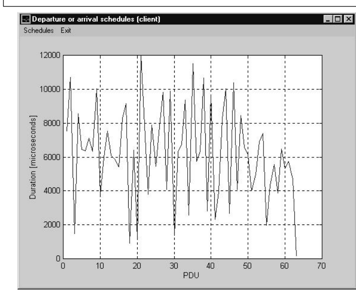


Figure 16. Classical IP over ATM, M2_B video model, broadcast. The arrival schedule is different than the departure schedule

LANMAN'99, 21-24 November 1999, Sydney SLIDE 12