

MODELLING TRAFFIC IN IMS NETWORK NODES

BA Alassane¹, OUYA Samuel² and FARSSI Sidi Mohamed³

^{1,2}Laboratory of Computer, Network and Telecommunications, Cheikh Anta Diop
University of Dakar, Dakar, Senegal
alassane.mail@gmail.com, samuel.ouya@gmail.com

³Laboratory of Medical, Imaging and Bioinformatics Cheikh Anta Diop
University of Dakar, Senegal
farsism@yahoo.com

ABSTRACT

IMS is well integrated with existing voice and data networks, while adopting many of their key characteristics.

The Call Session Control Functions (CSCFs) servers are the key part of the IMS structure. They are the main components responsible for processing and routing signalling messages.

When CSCFs servers (P-CSCF, I-CSCF, S-CSCF) are running on the same host, the SIP message can be internally passed between SIP servers using a single operating system mechanism like a queue. It increases the reliability of the network [5], [6]. We have proposed in a last work for each type of service (between I-CSCF and S-CSCF (call, data, multimedia.))[23], to use less than two servers well dimensioned and running on the same operating system.

Instead dimensioning servers, in order to increase performance, we try to model traffic on IMS nodes, particularly on entries nodes; it will provide results on separation of incoming flows, and then offer more satisfactory service.

KEYWORDS

IMS, NODE, CSCF, SIP, QoS

1. INTRODUCTION

The interconnection between different telecommunication networks with different access technologies has evolved to a heterogeneous environment with many challenges for real time applications.

In converging towards an All-IP network and new technologies, the introduction of multimedia with services strict requirements implies QoS guarantees.

One of main aspect of the study of the quality of service (QoS) is revisiting the traffic behavior at the nodes of the network. There will be a need for modelling IMS nodes network and multimedia traffic.

Relation between queue load and output rate in the core network nodes will be studied. Also, we study the probability distribution of the number of messages in each queue and the distribution of the delay of a message through nodes and discipline services associated.

Modelling traffic on IMS nodes, particularly on CSCFs and core services nodes will provide results on separation of incoming flows, and then reduce SIP overload and latency.

This paper is organized as follows: section 2 gives the node traffic equations. In section 3, three cases messages models distributions are studied. The paper concludes in Section 4.

2. THE NODE TRAFFIC EQUATIONS

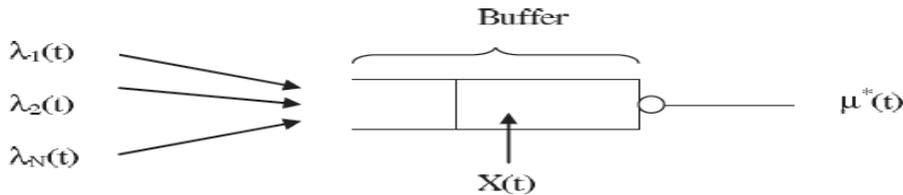


Figure 1

Nodes are modelled as D(t)/D/1/N queues [1],[18],[14],[22].

D(t) arrivals are considered when we model the exact evolution of TCP/IP not its stochastic behaviour. Figure 1 depicts the queue and the different parameters associated with it: D(t)/D/1/N D(t) means transient deterministic arrival

$\lambda(t) = \sum_i \lambda_i$ is the total rate entering the node at date t

$X(t)$ is the queue load at date t

$\mu^*(t)$ the output rate of the queue at date t

dt the integration step (infinitesimal variation time).

The following fundamental relation must always be verified (at any time t)[18].

$$X(t + dt) = (\lambda(t) - \mu^*(t)) \cdot dt + X(t) \quad (1)$$

If $X(t+dt)$ is bigger than buffer size, we start the loss processing mode. An overflow has occurred while the source doesn't know about it yet.

Suppose that the network is composed of N nodes, each containing an independent exponential server and an infinite buffer. The external arrival process to each node is Poisson with average rate λ_i $i = 1, 2, \dots, N$. Messages are routed probabilistically.

The total flow into each node of the network $A_i \ i = 1,2, \dots, N$, is composed of external arrivals plus flows of each of the other nodes in the network. This relationship is represented as:

$$A_i = \lambda_i + \sum_{j=1}^N q_{ji} A_j \quad i = 1,2, \dots, N \quad [20],[22] \quad (2)$$

where q_{ji} is the probability that a message is routed from node j to node i the associated matrix version is :

$$A = \lambda + AQ, \text{ where} \quad (3)$$

$$Q = \{q_{ji}\} \text{ is an } N \times N \text{ routing matrix } i,j = 1,2, \dots, N$$

$$A = [A_1, A_2, \dots, A_N]$$

$$\lambda = [\lambda_1, \lambda_2, \dots, \lambda_N]$$

When the inputs and the routing matrix are known, the total input traffic into each node is

$$A = \lambda[I - Q]^{-1}, \quad (4)$$

where I is the $N \times N$ identity matrix
From equation (1)

$$\lambda(t) = \frac{X(t + dt) - X(t)}{dt} + \mu^*(t)$$

From equation (3) $\lambda = A - AQ$

Hence $\frac{X(t+dt)-X(t)}{dt} + \mu^*(t) = A - AQ \quad (5)$

$$\Delta X + \mu^* = A - AQ$$

$$\mu^* = A(I - Q) - \Delta X \quad (6)$$

For example:

Define A_i as the total flow into node $i, i = 1,2,3,4$. In terms of the external arrivals, suppose the following set of equations for the total flow into each node:

$$A_1 = \lambda_1$$

$$A_2 = \lambda_2 + 0.25A_1$$

$$A_3 = \lambda_3 + 0.75A_1$$

$$A_4 = \lambda_4 + 0.2A_2 + 0.6A_3$$

The routing matrix is:

$$\begin{bmatrix} 0 & 0,25 & 0,75 & 0 \\ 0 & 0 & 0 & 0,2 \\ 0 & 0 & 0 & 0,6 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

If the external flow is represented by the vector $\lambda = [\lambda_1, \lambda_2, \dots, \lambda_N] = [2,1,0.5,3]$

We solve (4) and find $A = [A_1, A_2, \dots, A_N] = [2,1.5,2,4.5]$

Relation between queue load and output rate is:

$$\begin{cases} \mu_1^* = 2 - X_1 \\ \mu_2^* = 1 - X_2 \\ \mu_3^* = 0.5 - X_3 \\ \mu_4^* = 3 - X_4 \end{cases} \quad (7)$$

The CSCFs servers (P-CSCF, I-CSCF, S-CSCF), the HSS (database containing all subscribers data, the services allowed, the granted network and location information) and the Presence Server (PS) (inform about reachability, availability, and willingness of users) are the main components responsible for processing and routing messages on IMS network [5],[6],[7],[8].

Consider, Proxy-CSCF(entry node of IMS network) and its external flows a_0, a_1, a_2, a_3 respectively to I-CSCF, S-CSCF, HSS, PS;

Relation between queue load and output rate in this core entity is:

$$\begin{cases} \mu_1^* = a_0 - X_1 \\ \mu_2^* = a_1 - X_2 \\ \mu_3^* = a_2 - X_3 \\ \mu_4^* = a_3 - X_4 \end{cases} \quad (8)$$

3. MESSAGE DISTRIBUTION

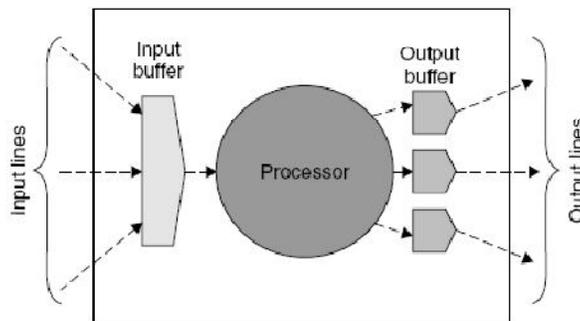


Figure2

Messages enter the node from the external lines and are fed into an input buffer, where they are held for processing. The processor performs such operations as correcting and detecting errors, sending positive and negative acknowledgments and routing. After processing, the message is placed into an output queue for transmission over an output line. In this model the input buffer and processor form a single queue as does each output buffer and line.

Let us be interested in the probability distribution of the number of messages in each queue and the distribution of the delay of a message through the node.

We study distribution associated to three discipline service: the processor sharing, the infinite server model and the preemptive arbitrary service.

3.1 THE PROCESSOR SHARING MODEL

In a processor sharing (PS) queuing system, the server capacity is shared equally among all the customers that are present in the system. This model is applicable to a time-shared computer system where a central processor serves all the jobs present in the system simultaneously at an equal service rate, suppose the Proxy-CSCF server. It is the entry node and can host traffic engineering intelligence of the system.

Another important application of PS is for a multiplicity of TCP connections that share a common bottleneck. The Internet router at the bottleneck simultaneously switches the flows generated by the users, while TCP congestion control mechanism guarantees that the service rate obtained by the different flows are equal [13], [14], [22].

Service regime is assumed to be processor sharing. In particular, we assume that the process of the number of customers in the system is a continuous time Markov-chain, where customers arrive according to a Poisson process with parameter λ and that the service time required by an arriving customer is exponentially distributed with parameter μ . We also assume the stability condition of $\lambda < \mu$.

The average arrival rate $\rho = \frac{\lambda}{\mu}$

Model D(t)/D/1/N applied to sharing processor is statistically the same as the continuous-time Markov chain that describes the M/M/1 (FCFS) queue [22].

The mean queue-size $E[Q]$ is given by:

$$E[Q] = \frac{\rho}{1-\rho} \quad (9)$$

The delay distribution is:

$$E[D] = \frac{1}{(1-\rho)\mu} = \frac{1}{\mu-\lambda} \quad (10)$$

The mean delay of a customer that requires amount of service x is $E[D(x)]$

The variable x here represents the time that the customer spends in the system to complete its service assuming that there are no other customers being served and all the server capacity can be dedicated to it.

$$E[D(x)] = \frac{x}{1-\rho} \quad (11)$$

$E[Q]$, $E[D]$ and the mean delay of a customer with service requirement x are insensitive to the shape of the distribution of the service-time requirements of the customers.

In other words, these results apply also to model characterized by Poisson arrivals, generally distributed service-time requirements, and a processor sharing service policy.

For example:

Consider packets arriving at a multiplexer where the service discipline is based on processor sharing. Assume that the service rate of the multiplexer is 2.5 Gb/s. The mean packet size is 1250 bytes. The packet arrival process is assumed to follow a Poisson process with rate of 200,000 [packet/sec] and the packet size is exponentially distributed.

1. Find the mean number of packets in the multiplexer.
2. Find the mean delay of a packet.
3. Find the mean delay of a packet of size 5 kbytes.

Solution

Find the mean number of packets in the multiplexer.

mean packet size = 1250 x 8 = 10 000 bits

$$\mu = \frac{2500\ 000\ 000}{10000} = 250\ 000 \text{ paquets/s}$$

$$\rho = \frac{200\ 000}{250\ 000} = 0,8$$

$$E[Q] = \frac{0,8}{1-0,8} = 4$$

Find the mean delay of a packet.

By Little's formula $E[D] = \frac{E[Q]}{\lambda} = \frac{4}{200\ 000} = 0,00002s = 20\mu s$

Find the mean delay of a packet of size 5 kbytes.

Let x be the time that the 5 kbytes packet is delayed if it is the only one in the system

$$X = \frac{5000 \times 8}{2,5 \times 1000\ 000\ 000} = 16\ \mu s$$

Now we will use the time units to be microseconds.

$$E[D(x)] = \frac{16}{1-0,8} = 80\ \mu s$$

A packet four times larger than the average sized packet will be delayed four times longer

3.2 THE INFINITE SERVER MODEL

The infinite-server model is a model in which a message is immediately assigned a server as it enters the system.

Notice that, for both processor sharing and the infinite-server case, all messages are simultaneously in service.

There is no dependence between the probability distribution and the service time.

These results can be extended in an analogous fashion to nodes serving several different classes of messages, each with its own arrival rate and service time distribution.

n_i : number messages of class i

ρ_i : probability distribution of message i

$\rho = \frac{\lambda}{\mu}$ μ and λ are the arrival and service rates

JEREMIAH F. HAYES and THIMMA V. J. GANESH BABU shown that the probability distribution is [14]:

$$P(n_1, n_2, \dots, n_c) = (1 - \rho)m! \prod_{i=1}^c \frac{\rho_i^{n_i}}{n_i!} \quad [14], \quad (12)$$

Where $\rho = \sum_{i=1}^c \rho_i$, $m = \sum_{i=1}^c n_i$

3.3 THE PREEMPTIVE ARBITRARY SERVICE

For the last case, we turn to the preemptive resume last come first served (LCFS) discipline.

For the LCFS queue discipline, a message at the head of the line having reached a certain stage in the service process may be displaced by a new arrival. This message may be displaced in turn by a new message, and so on it goes.

In our treatment of individual nodes we consider multiple classes of messages in a LCFS node and the probability distribution is given by [14]:

$$P(n_1, n_2, \dots, n_c) = (1 - \rho)m! \prod_{i=1}^c \frac{\rho_i^{n_i}}{n_i!}, \quad (13)$$

where parameters are the same defined in previously.

In general, a network can be viewed as consisting of a number of closed chains, one for each source-destination pair or node. The flow can be viewed as number of messages circulating between nodes. Suppose in an N-node network there is a single class of traffic (more in general) between each pair. The number of such chains or nodes is $J=N(N-1)$.

With the three CSCFs, HSS and PS points $J=5 \times 4 = 20$

$$P(n_1, n_2, \dots, n_{20}) = (1 - \rho)20! \prod_{i=1}^{20} \frac{\rho_i^{n_i}}{n_i!}, \quad (14)$$

4. CONCLUSION

The mean number of packet and the mean delay of a customer with service requirement x are insensitive to the shape of the distribution of the service-time requirements of the customers.

These results apply also to model characterized by Poisson arrivals with a processor sharing service policy.

Notice that, for both processor sharing and the infinite-server case, all messages are simultaneously in service. The final service type for which the product form holds is preemptive resume last-come first-served (LCFS), in which newly arrived messages are served immediately. Displaced messages are queued and resume where they were left off when the server is available again.

When the service times of messages vary widely, these alternatives to the FCFS discipline may offer more satisfactory service. For example, when most messages are short, with only the occasional long message, processor sharing has the effect of giving priority to short messages. This model can be deployed in Proxy-CSCF in this case, and priority given to voice and short messages.

Future work: Future research needs to work on limited storage case, blocking case when buffer overflow and at last managing priorities.

REFERENCES

- [1] Hassan HASSAN, Jean-Marie GARCIA and Olivier BRUN LAAS-CNRS, Toulouse, France, Bandwidth Allocation and Session Scheduling using SIP, JOURNAL OF COMMUNICATIONS, VOL. 1, NO. 5, AUGUST 2006,
- [2] Ward Whitt, Tail probabilities with statistical multiplexing and effective bandwidths in multiclass queues, AT&T Bell Laboratories, Murray Hills, NJ 07974-0636, USA, August 1993
- [3] Daizo Ikeda, Toshihiro Suzuki and Akira Miura, Performance Analysis of a Mobile Network for API-Oriented Traffic, Research Laboratories, NTT DOCOMO, Inc. 3-5 Hikarino-oka, Yokosuka-shi, Kanagawa 239-8536, Japan Prefectural University of Kumamoto 3-1-100, Tsukide, Kumamoto, 862-8502, Japan , International Journal of Control and Automation Vol. 4, No. 4, December, 2011
- [4] Tomáš Mácha, Luboš Nagy, Zdeněk Martinásek , Vít Novotný, IMS Mapping of QoS Requirements on the Network Level, Fakulta elektrotechniky a komunikačních technologií VUT v Brně , JUNE 2010
- [5] C. Chi Bell Laboratories, Alcatel-Lucent chic, Hao Bell Laboratories, Alcatel-Lucent D.Wang Bell Laboratories, Alcatel-Lucent Z. Cao Institute of Information Science Beijing Jiaotong University, Modelling IMS Presence Server: Traffic Analysis & Performance, 2008 IEEE
- [6] Mesud Hadžialić, Mirko Škrbić, Nerma Šečić, Mirza Varatanović, Elvedina Zulić, Nedim Bijedić University of Sarajevo , Problem of IMS modelling – Solving Approaches, CTRQ 2012
- [7] Mlindi Mashologu, Performance optimization of IP Multimedia Subsystem, Dissertation.com, Boca Raton, Florida 2010
- [8] V.S. Abhayawardhana, R. Babbage, A Traffic Model for the IP Multimedia Subsystem (IMS), 2007 IEEE
- [9] Michael T. Hunter Russell J. Clark Frank S. Park Security Issues with the IP Multimedia Subsystem (IMS): A White Paper College of Computing Georgia Institute of Technology {frank,mhunter,russ.clark}@gatech.edu Version 1.0 September 1, 2007
- [10] Pierre COATANE, IMS: IP Multimedia Subsystem, EFORT
- [11] Nils Berglund, Chaînes de Markov, Université d'Orléans Décembre 2007
- [12] Alexander Klemm, Christoph Lindemann, and Marco Lohmann Modeling IP Traffic Using the Batch Markovian Arrival Process, Modeling IP Traffic Using the Batch Markovian Arrival Process

University of Dortmund Department of Computer Science August-Schmidt-Str. 12 44227 Dortmund, Germany

- [13] Modelling Markovian Queues and Similar Processes Winfried Grassmann Department of Computer Science University of Saskatchewan, CORS - SCRO 2000 ANNUAL CONFERENCE Energy, Natural Resources, and the Environment MAY 29-31, 2000 EDMONTON, ALBERTA
- [14] JEREMIAH F. HAYES THIMMA V. J. GANESH BABU Modeling and Analysis of telecommunications network 2004 by John Wiley & Sons, Inc. All rights reserved
- [15] Jeffrey P. Buzen Harvard University and Honeywell Information Systems, Computational Algorithms for Closed Queueing Networks with Exponential Servers Communications September 1973 of Volume 16 the ACM Number 9
- [16] Muhammad T. Alam, Graduate Student Member IEEE, Zheng Da Wu, Member IEEE Admission Control Approaches in the IMS, Presence Service, international journal of computer science volume 1 number 1 2006
- [17] Henning Schulzrinne, Personal Mobility for Multimedia Services in the Internet, IDMS'96 (European Workshop on Interactive Distributed Multimedia Systems and Services), Berlin, Germany, March 4-6, 1996
- [18] Hassan HASSAN Modélisation et analyse de performances du trafic multimedia dans les réseaux hétérogènes, mémoire Doctorat de l'Université Paul Sabatier-Toulouse III 2007
- [19] PETTER LINDGREN, Diameter service creation investigation and HSS, Evaluation, Master's Thesis Supervisor: Stefan Östergaard, Kajsa Goffrich Examiner: Prof. Gerald Q. Maguire Jr, 2010-02-23
- [20] Balaji Prabhakar Stanford University, Network Processor Algorithms: Design and Analysis, Stochastic Networks Conference Montreal July 22, 2004
- [21] Vern Paxson and Sally Floyd, Wide-Area Traffic: The Failure of Poisson Modelling, Lawrence Berkeley Laboratory and, EECS Division, University of California, Berkeley 1 Cyclotron Road, Berkeley CA 94720, July 18, 1995
- [22] Moshe Zukerman, Introduction to Queuing Theory and Stochastic Teletraffic Models, 2000-1012
- [23] Improving traffic performance on IMS network, BA Alassane, OUYA Samuel, FARSSI Sidi Mohamed, may 2013

Authors

BA Alassane is a PhD student in telecommunications holder of a DEA in Mathematics



Dr OUYA Samuel is the head of the Laboratory of Computer, Network and Telecommunications, Cheikh Anta Diop University of Dakar, Senegal. Doctor OUYA is a mathematician IT specialist having a long experience in teaching and search. He participates in several IT meets.



Dr FARSSI Sidi Mohamed is the head of the Laboratory of Medical, Imaging and Bioinformatics, Cheikh Anta Diop University of Dakar, Senegal. Doctor FARSSI is a IT specialist who participate in several IT and medical meetings

