15th Triennial World Congress, Barcelona, Spain

QUALITY OF SERVICE OF COMMUNICATION NETWORKS AND DISTRIBUTED AUTOMATION: MODELS AND PERFORMANCES

Guy Juanole *

* *LAAS-CNRS. 7, avenue du colonel Roche.31077 Toulouse FRANCE*

Abstract: Today networks are more and more used to implement distributed applications and, in particular, industrial applications like closed loop process control which require a timeliness service to guarantee their performances. The timeliness property depends on the message and task schedulings and the communication protocol. The goal of this paper is to present a pluridisciplinary work which must operate at the distributed control system level (computer science and telecommunication view), at the automation level (automatic control view) and to link these two views.

Keywords: Communication networks. Quality of Service. Automatic control. Formal models. Performance evaluation

1. INTRODUCTION

In the technological context of today, distributed systems (computers connected through a communication network) are more and more used and, particularly, for implementing industrial applications like, for example, closed loop process control systems (Astrom and Wittenmark, 1997 ; Coughanour and Koppel, 1965). Such systems must be "real time" and still more "hard real time" i-e the timeliness of the service which is provided is essential to garantee the performances of the applications. This timeliness depends on the timeliness of the computations (beginning, duration), of the message productions and consumptions, and of the message exchanges (sending on the network, reception from the network). A lot of works have been done in the last decade, particularly, on scheduling and ressource management (Stankovic and Ramaritham, 1988 ; Tilborg and Koob, 1991 ; Rajkuman, 1991 ; Montuschi et al., 1992 ; Malcolm and Zhao, 1995 ; Sevcik and Johnson, 1994 ; Tindell and Hansson, 1994 ; Carmo et al., 1994) both in terms of algorithms for tasks and messages, and of performances. All these works are mainly concentrated on a particular point in a distributed system and then give partial informations on the distributed system.

In our opinion, it lacks works which: 1. try to integrate several partial results in order to evaluate, in the context of a given (industrial) application, the Quality of the Service which is provided; 2. to link this Quality of Service to the performances of the (industrial) application. These performances are expressed in terms specific to the application (often not understandable by a "computer science man"!).

Such works concern the global system and are then a pluridisciplinary work (at least two disciplines) which must operate at the distributed system level ("Computer Science view") and at the (industrial) application level (if we have a closed loop control system, it is an "Automatic Control view"). These works require, at least, two kinds of formal modelling: related to the point 1 (computer and communication domain), we need models which allow, on the one hand, to express the mechanisms of distributed systems (parallelism, synchronisation, choice, mutual exclusion, priority...) as well as timed aspects (constraints, durations...), and, on the other hand, to make qualitative analysis (deadlocks for example) and quantitative analysis (response time, mean time to the first failure...); related to the point 2 (if we consider closed loop control systems (Automatic Control domain)), we need models which allow to represent the transfert of signals (by considering the frequency of the signals) and to deal with the feedback structure. Petri nets based models like Stochastic Timed Petri Nets (STPN) (Juanole and Gallon, 1995 ; Atamna and Juanole, 1995) are well suited for the point 1 and will be used here. The concept of transfer function based on Laplace transform (Astrom et Wittenmark, 1997 ; Coughanou and Ray, 1988) is well suited for the point 2 and will be used here.

The objective of this paper is precisely to make such a work by considering a closed loop control system based on a distributed system around a fieldbus.

This paper includes three parts. The first part presents the global system in terms of architecture and of the model of the closed loop control system. In this model, the parameters expressing the Quality of the Service provided by the distributed system are emphasized. Finally, the performance which is considered for the closed loop control system is presented and expressed as a function of these parameters.

The second part presents the general modelisation, using the STPN model (this model consider immediate transitions (null duration) represented with thin bars and timed transitions (non null duration) represented with rectangles), of the part of the distributed system which is concerned by the considered application. We can then evaluate the influence of the main distributed system parameters (for the network part: access delay, sample loss, sample transfer protocol; for the local computers: task scheduling) on the Quality of the Service which is provided by the distributed system. We finally present the impact on the performance of the closed loop control system.

The third part concerns the translation of the general study made in the second part to an actual case i-e by considering a type of medium access control (FIP network (NFa, 1990)) and task scheduling.

2. GLOBAL SYSTEM

2.1 Architecture

The global system architecture is represented on the figure 1. The distributed system is composed of several sites connected through a bus. The sites i and j are only concerned by the considered application (industrial process control where the controlled variable s(t) is captured in the site i and the returned variable r(t) is compared to the set point in the site j). The role of the sites i and j (with the associated computers) is very simple: they implement (by using the bus) the feedback loop.

In the site i, we have, at first, the sampling (period T_o) of the analog signal s(t) (analog samples s_k(t)) and the analog to digital conversion (A-D). The digital samples are considered in the computer of the site i by the task called the Producer task (this task adds informations to the digital samples like, at least, an identifier). Then the Producer task requires the service of the Data link entity for transmitting data (sample concatenated with the identifier). Finally, the Data link entity sends frames (each data is encapsulated in a frame) on the bus (note that, in a frame, the Data link entity puts an error detector code in order to control errors induced on the bits by the transmission on the bus).

In the site j, after the reception of a frame and the positive test of the error detector code (if the test is negative, the Data link entity does not consider the data which is included in the frame, which means that we have a loss of this data), the Data link entity transmits the data included in the frame to the task, called the Consummer task, which reads the sample value and sends it to the digital analog computer (D-A). After the Hold, we have the returned variable r(t).

In this context, the distributed system includes three layers (user, data link, physical).

The distributed system provides then the service of sample transfer between a Producer task and a Consumer task. During the service relative to each sample, the sample can be either delayed (due to the scheduling time resulting from the internal and/or external environnement) or lost (due to errors on the bits during the transmission of the frame on the bus).

Then, about the sample r_k(t) wave, we can say: with respect to the sample s_k(t) wave, we have a delay τ_d which can be characterized by a jitter (the superior bound is the worst case); between the samples r_k(t), the time relations (due to the losses, in particular) can be different of the period T_0 (strictly speaking, these time relations can no more be strictly periodic but we can approximate them by a period T (T> T_0) which represents a mean value).

In the following, the Quality of the Service (QoS), which is provided by the distributed system, is then represented be the two parameters: delay τ_d and period T.

2.2 Closed loop control system

2.2.1. Model We consider the model, based on the transfer function concept, given on the figure 2. The

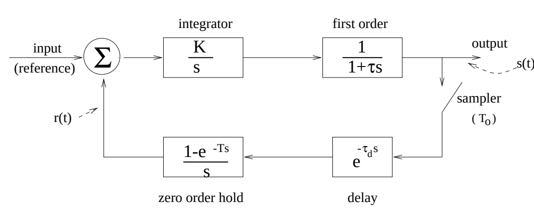

Figure 2. Model of the closed loop control system

feedback loop integrates the parameters τ_d and T which represent the distributed system QoS. Note that the transfer function of the zero order hold (period T which is either equal to T_o if we have no losses in the distributed system (and the delay τ_d is smaller

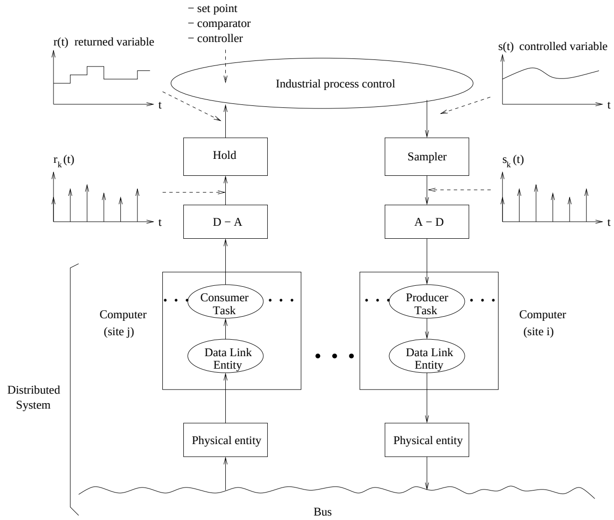

Figure 1. Global system

than the period T_0, what we suppose in this study) or different of T_o in the opposite case) can still be approximated by a continuous transfer function which is a pure delay $e^{-\frac{T}{2}s}$ (the conditions are (Astrom et Wittenmark, 1997): sampling frequency (1/T) $\gg$ higher frequency of signal s(t)). In the following, we consider this approximation and the feedback loop has then the continuous transfer function $e^{-(\tau_d+\frac{T}{2})s}$. In this way, we can analyse the closed loop control system as a continuous linear system.

2.2.2. Performance A fundamental performance is the stability which is characterized by the phase margin φ_m:
$\varphi_m = 180^0 - arg\left(\frac{K}{j\omega} * \frac{e^{-j(\tau_d+\frac{T}{2})\omega}}{1+j\tau\omega}\right)$ (for ω such that $\mid \frac{K}{j\omega}\frac{1}{1+j\tau\omega} \mid= 1$). φ_m then depends on the two parameters τ_d and T. The numerical values ($K = 50\sqrt{2}$ rd/s; $\tau = 2\ 10^{-2}$ s; $T_0 = 10^{-2}$ s) are considered for our study. Note that, in this study, we are not concerned by the very adequate value of φ_m but we want to show why and how the distributed system influences this performance.

3. DISTRIBUTED SYSTEM MODELISATION AND EVALUATION: GENERAL CASE

We model the distributed system concerned with the application presented on the figure 1. We consider two data transfer protocols : a simple protocol without loss control, and a protocol with loss control (it is the alternating bit protocol (Bartlett et al, 1969) which by means, of a Timer in the data sender and a modulo 2 numbering scheme controls the losses of the frames carrying data and of the acknowledges).
The STPN models of the protocols without loss control and with loss control are presented respectively on figures 3 and 4. The explaination of these models (not given here for space reasons) can be found in (Juanole and Blum, 1998).

The QoS parameters (delay τ_d and period T), which show the influence of the protocol, are given on the tables 1 and 2.
The table 1 shows the following important points about the QoS parameters: the delay τ_d and the period T increase with the loss probability; the period T only depends on the loss probability; the delay τ_d depends on the hold simultaneously effects of the initial delay (delay when no loss) and the loss probability (this can be well seen when we have a delay with jitter and we use different firing rules for the STPN model). The table 1 still shows the advantage of the STPN model (thanks to the different firing rules) to analyse the jitter and then to evaluate the worst case (with the MAX rule). The table 2 shows two points: the great advantage of a protocol with loss control (the delay τ_d and the period T have now values very close to respectively the initial delay and the period T_0; only when the loss probability has high values we have changes); the influence of the value of the timer on the delay (what is normal).
The performances of the closed loop system are represented on fogures 5 and 6 which show the evolution of phase marging as a function of the mechanisnms internal to the distributed system.

4. APPLICATION TO AN ACTUAL CASE

The real system that we consider (for implementing the closed loop control system) is based on a network FIP (NFa, 1990) which has a centralized medium access control. Futhermore, we consider that the user tasks, in the computers, are scheduled with the Rate Monotonic (RM) (Liu and Layland, 1973) algorithm and that the priority inversion is controlled either by a priority inheritance protocol or a ceiling priority protocol (Rajkuman, 1991).

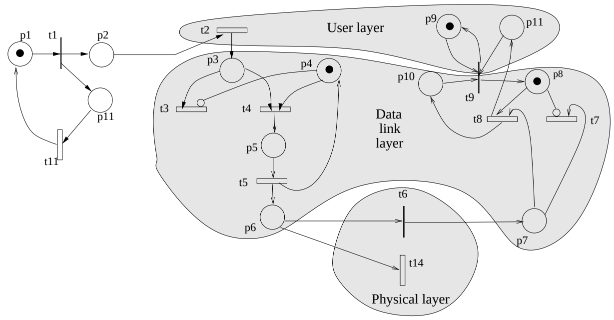

Figure 3. STPN model

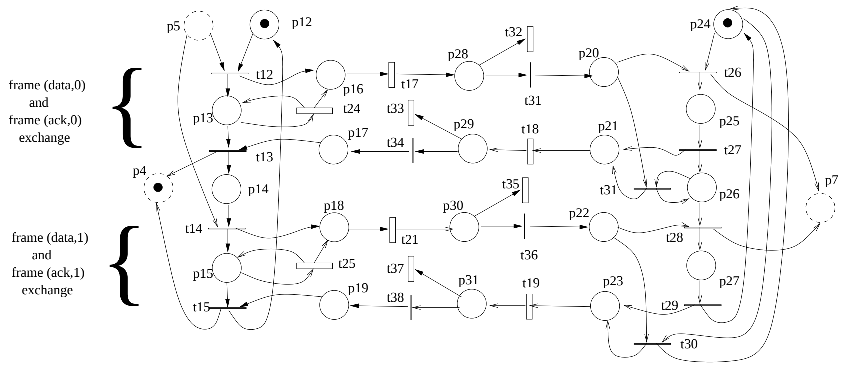

Figure 4. STPN model of the loss control protocol

	Transition t_5 (transmission delay)			
	deterministic	delay with jitter $2.5\ 10^{-4} \leq delay \leq 7.5\ 10^{-4}$		
Loss probability	delay: $5\ 10^{-4}$	MIN firing rule	mean firing rule	MAX firing rule
0	T= 1.0 10^{-2} τ_d= 5.10^{-4}	T= 1.0 10^{-2} τ_d= 2.5 10^{-4}	T= 1.0 10^{-2} τ_d= 5.10^{-4}	T= 1.0 10^{-2} τ_d= 7.5 10^{-4}
10^{-3}	T= 1.001 10^{-2} τ_d= 5.3 10^{-4}	T= 1.001 10^{-2} τ_d= 2.8 10^{-4}	T= 1.001 10^{-2} τ_d= 5.3 10^{-4}	T= 1.001 10^{-2} τ_d= 7.8 10^{-4}
10^{-2}	T= 1.010 10^{-2} τ_d= 6.21 10^{-4}	T= 1.010 10^{-2} τ_d= 3.71 10^{-4}	T= 1.010 10^{-2} τ_d= 6.21 10^{-4}	T= 1.010 10^{-2} τ_d= 8.71 10^{-4}
10^{-1}	T= 1.111 10^{-2} τ_d=16.3 10^{-4}	T= 1.111 10^{-2} τ_d=13.38 10^{-4}	T= 1.111 10^{-2} τ_d=16.3 10^{-4}	T= 1.111 10^{-2} τ_d=18.8 10^{-4}
$2\ 10^{-1}$	T= 1.249 10^{-2} τ_d=30.02 10^{-4}	T= 1.249 10^{-2} τ_d=27.68 10^{-4}	T= 1.249 10^{-2} τ_d=30.02 10^{-4}	T= 1.249 10^{-2} τ_d=32.68 10^{-4}

Table 1. Qos with respect to a Protocol without loss control

Loss probability	time-out=1.2 10^{-3}	time-out=2.4 10^{-3}
0	T=1.0 10^{-2} τ_d=5.10^{-4}	T= 1.0 10^{-2} τ_d= 5.10^{-4}
10^{-3}	T=1.0002 10^{-2} τ_d=5.21 10^{-4}	T= 1.0005 10^{-2} τ_d= 5.22 10^{-4}
10^{-2}	T=1.0024 10^{-2} τ_d=5.32 10^{-4}	T= 1.0048 10^{-2} τ_d= 5.44 10^{-4}
10^{-1}	T=1.0274 10^{-2} τ_d=6.33 10^{-4}	T= 1.0556 10^{-2} τ_d= 7.87 10^{-4}
$2\ 10^{-1}$	T=1.0641 10^{-2} τ_d=8.2 10^{-4}	T= 1.1132 10^{-2} τ_d=11.2 10^{-4}

Table 2. QoS with respect to a Protocol with loss control (t_5 : 5.10^{-4} s)

4.1 Hypothesis for the QoS evaluation

Call, with respect to the application of the closed loop control system, task i and variable i respectively the producer task and the variable generated by the producer task (from the samples $s_k(t)$) with the period $T_0 = 10^{-2}$ s.

The QoS evaluation requires to specify the external environnement (point 1) and the internal environnement (point 2).

For the point 1, we suppose that we have four other periodic variables with a period equal to T_0. By considering an RM algorithm, all the variables have the same priority and the scheduler must then defines arbitrarily a priority. We consider two cases: the variable i has the highest priority (it is scheduled the first); the variable i has the lowest priority (it is scheduled the last).

For the point 2, we suppose that, in the task i site, there are four other user tasks but the task i has the highest priority (and furthermore each task has the same excution time noted c). We consider three scheduling

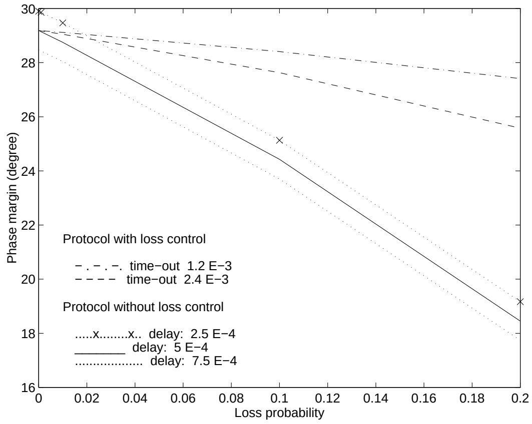

Figure 5. Phase margin with respect to the protocols

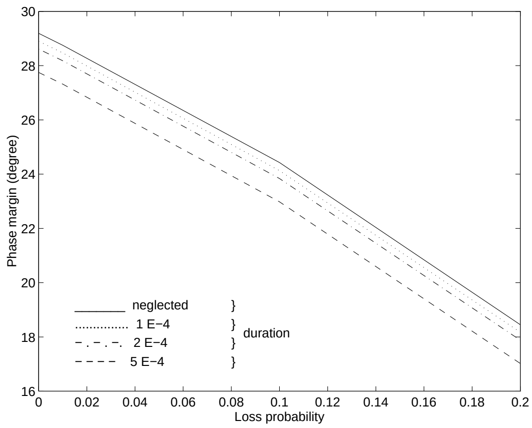

Figure 6. Phase margin with respect to the producer task duration

algorithms: RM algorithm (we suppose that there is no ressource sharing problem); RM with priority inheritance and RM with ceiling priority for dealing with the sharing of one ressource (we call c_{cs} the duration of the critical section for each task).
The points 1 and 2 allow to calculate respectively the variable i transmission delay and the task i busy time, and then to specify the probability densities associated to the transitions t5 and t2 of the STPN model of the figure 3 (which is used to evaluate QoS).
Concerning the variable i transmission delay, its expression is (we suppose that the arbiter starts the exchange relative to the variables of the same period at the instant where these variables are written by the producer in the data link buffer; we neglect the propagation time):

- case when the variable i has the highest priority: id-dat+rp-dat+tr
- case where the variable i has the lowest priority: 5(id-dat+rp-dat)+9tr

Concerning the task i busy time, its expression is:

- RM algorithm: c
- RM with priority inheritance: c+4c_{cs}
- RM with ceiling priority: c+c_{cs}

We consider the following numerical values:

- FIP network (we consider a bit rate of 1 Mb/s and in this case we have: 10 μs $\leq$ tr $\leq$ 70 μs).
  - id-dat has 61 bit length; rp-dat has a length of 45 bits plus n bytes of data (we take here n=8)
  - then (id-dat+rp-dat) has a length of 170 bits and a duration of 170 μs
- user tasks: c=100 μs; c_{cs}=100 μs.

4.2 *Results*

We present the results obtained by considering the numerical values of the works made in (Juanole and Blum, 1998).

4.2.1. Only considering the external environnement
The result of the analysis on the QoS of the distributed system and on the phase margin of the closed loop control system are summarized on the table 3. The main comments are: the delay τ_d and the period T increase with the loss probability and decrease with the priority of the variable (it is why the phase margin decreases with the loss probability and increases with the variable priority); the period T is independant of the turn over time (it only depends on the losses); on the other hand, the delay increases with the turn over time.

4.2.2. Considering both the external and internal environnements This study has been done by considering no losses (then the period T is always equal to T_0). The main results are summarized on the table 4. We can see (in more of the result already got in the previous subsection, about the influence of the priority of the variable i): the RM (only) algorithm gives the more positive phase margin (that is normal because there is no ressource sharing which consumes time); the algorithm RM with priority inheritance induces a phase margin smaller than the algorithm RM with ceiling priority (that is normal as the blocking time is higher with the priority inheritance).

5. CONCLUSION

We have presented a global study of a system (i-e considering both the underlying distributed system and the application (closed loop control system) running on it) which then requires to handle, at least, two types of formal models (one for the distributed system, another one for the application). We are convinced that this kind of work (pluridisciplinary) is more and more necessary today, taking into account the very sophisticated applications which are implemented on the top of the distributed systems. If some years ago, as the communication networks were mainly used for very simple applications of data transfer, the evaluation of delays, losses ...was enough, it is now necessary to make the linkage with application performance. Furthermore, in this context, concerning the distributed system, we must integrate all the different

Loss probability	Variable i with the highest priority			Variable i with the lowest priority		
	tr=10	tr=50	tr=70	tr=10	tr=50	tr=70
0	T= 1.0 10^{-2} τ_d= 2.0 10^{-4} φ_m=30.10	T= 1.0 10^{-2} τ_d= 2.4 10^{-4} φ_m=29.99	T= 1.0 10^{-2} τ_d= 2.6 10^{-4} φ_m=29.93	T= 1.0 10^{-2} τ_d= 9.6 10^{-4} φ_m=27.93	T= 1.0 10^{-2} τ_d=13.2 10^{-4} φ_m=26.89	T= 1.0 10^{-2} τ_d=15.0 10^{-4} φ_m=26.38
10^{-3}	T= 1.001 10^{-2} τ_d= 2.1 10^{-4} φ_m=30.08	T= 1.001 10^{-2} τ_d= 2.5 10^{-4} φ_m=29.94	T= 1.001 10^{-2} τ_d= 2.7 10^{-4} φ_m=29.89	T= 1.001 10^{-2} τ_d= 9.7 10^{-4} φ_m=27.88	T= 1.001 10^{-2} τ_d=13.3 10^{-4} φ_m=26.85	T= 1.001 10^{-2} τ_d=15.1 10^{-4} φ_m=26.34
10^{-2}	T= 1.010 10^{-2} τ_d= 3.01 10^{-4} φ_m=29.67	T= 1.010 10^{-2} τ_d= 3.41 10^{-4} φ_m=29.55	T= 1.010 10^{-2} τ_d= 3.61 10^{-4} φ_m=29.49	T= 1.010 10^{-2} τ_d=10.61 10^{-4} φ_m=27.49	T= 1.010 10^{-2} τ_d=14.21 10^{-4} φ_m=26.46	T= 1.010 10^{-2} τ_d=16.01 10^{-4} φ_m=25.94

Table 3. QoS and phase margin

	RM only	RM with priority inheritance	RM with ceiling priority
Variable i with the highest priority	τ_d= 3.4 10^{-4} φ_m=29.60	τ_d= 7.4 10^{-4} φ_m=28.56	τ_d= 4.4 10^{-4} φ_m=29.42
Variable i with the lowest priority	τ_d=14.2 10^{-4} φ_m=26.61	τ_d=18.2 10^{-4} φ_m=25.46	τ_d=15.2 10^{-4} φ_m=26.32

Table 4. QoS and phase margin

main mechanisms (task scheduling, message scheduling, data transfer protocol);

REFERENCES

Karl J. Aström and Björn Wittenmark. Computer controlled Systems: Theory and Design. Prentice Hall. Information and system Science Series. 1997.

D. R. Coughanour, L. B. Koppel. Process control systems. Mc Graw-Hill. 1965.

Yoram Halevi, Asok Ray, Integrated communication and control systems. Journal of Dynamic systems, Measurement and Control. December 1988.

J. A. Stankovic and K. Ramaritham. Hard Real-Time Systems. IEEE Computer Society Press. 1988.

A. M. V. Tilborg, G. M. Koob, Foundations of real time computing. Scheduling and Ressource management. Kluwer academic publishers. 1991.

R. Rajkuman, Synchronization in real-time systems. A priority inheritance approach. Kluwer academic publishers. 1991.

P. Montuschi, L. Ciminiera and A. Valenzano, Time Characteristics of IEEE 802.4 Token Bus Protocol, IEE Proceedings-E, vol. no.1. January 1992.

N. Malcolm,W. Zhao, Hard Real-Time Communication in Multiple-Access Networks. The Journal of Real-Time Systems, 8(1). 1995.

K. Sevcik and M. Johnson, Cycle Time Properties of the FDDI Token Ring Protocol. IEEE Trans. on Software Eng., vol SE-13, no.3.1994.

K. Tindell, H. Hansson, A. Wellings, Analysing Real-Time Communications: Controller Area Network (CAN). IEEE RTSS'94, S. Juan, December 1994.

R. L. R. Carmo, F. Vasques, G. Juanole, Real-Time Communication services in a DQDB network. IEEE RTSS'94, S. Juan, December 1994.

G. Juanole and L. Gallon, Critical Time Distributed Systems: Qualitative and Quantitative Analysis based on Stochastic Timed Petri Nets. In FORTÉ95, the 8th International IFIP Conference on Formal Description Techniques for Distributed Systems and Communication Protocols. Montreal. Canada. October 1995.

Y. Atamna and G. Juanole, Methodology for Obtaining Abstract Views of State Graphs Labeled with Probabilities and Times: an Example of Application to a Communication Protocol, In MASCOTS'95, the Third International Workshop on Modeling, Analysis and Simulation of Computer and Telecommunication Systems. Durham. North Carolina. January 1995.

K.A. Bartlett, R.A. Scantlebury, P.T. Wilkinson, A note on reliable full-duplex transmission over half-duplex links. Commun. Ass. Comput. Mach., Vol. 12. May 1969.

NFa: FIP bus for exchange of information between transmitters, actuators and programmable controllers, data link layer, NF C46-603. June 1990.

C. Liu and J. Layland, Scheduling Algorithms for Multiprogramming in a Hard Real-Time Environnement. Journal of ACM, vol. 29, no. 1. 1973.

G. Juanole and I. Blum, Evaluating the Quality of Service of a real-time distributed system and its impact on the performance of an industrial application. Rapport LAAS 98202, Mai 1998.