INSTITUTO DE COMPUTAÇÃO
UNIVERSIDADE ESTADUAL DE CAMPINAS

Experiments about Quality of Service of Failure Detectors under Message Loss Bursts

I. Sotoma *E. R. M. Madeira*

Technical Report - IC-08-001 - Relatório Técnico

January - 2008 - Janeiro

Experiments about Quality of Service of Failure Detectors under Message Loss Bursts

Irineu Sotoma[a,b,1], Edmundo Roberto Mauro Madeira[a]

[a] *Institute of Computing, State University of Campinas, Campinas, SP, Brazil*

[b] *Department of Computing and Statistics, Federal University of Mato Grosso do Sul, Campo Grande, MS, Brazil*

Abstract

This paper presents a Markov Model used to configure the Chen,Toueg and Aguilera's NFD-U/NFD-E failure detector algorithm for unsynchronized clocks according to QoS requirements. The Markov Model uses the loss burst probabilities to capture the message loss burst information. The experiments with data gathered from a link between two long distance networks show that the proposed configurator performs better than the Chen, Toueg and Aguilera configurator when long loss bursts occur and performs similarly to their configurator when no loss burst occurs.

Keywords: Failure detectors; Quality of service; Fault tolerance; Distributed system.

1 Introduction

A failure detector uses monitoring messages to catch information about the crash of monitored processes. In distributed systems, failure detectors have been used to solve several problems such as distributed consensus [11, 9, 6] and group membership [8, 3, 23, 26]. In computer networks, the timers are a kind of failure detector which have as the main goal the detection and retransmission of lost packets [28, 22].

This paper is about the quality of service (QoS) of failure detectors, which is defined in the Chen et al paper [7]. They define metrics to measure and evaluate failure detectors about a fast crash detection and an accurate suspicion of process crashes. They developed a new failure detector (NFD-S) algorithm for synchronized clocks, and unsynchronized ones (NFD-U and NFD-E). However, the model defined for these failure detectors assumes only *message mean loss probability* as the system parameter for message losses.

Long distance [1, 27, 29] and wireless networks [2, 22] have a high rate of message loss bursts. Markov chain models have been used to model loss bursts in these networks [27, 29, 14, 16]. Markov chains usually use 2^m states, where m represents the last consecutive losses which are considered. So, Sanneck [21] proposes limited and unlimited state space Markov chain models which have only $m + 1$ states. In these models, the state and state transition probabilities are approximated by using the probability distribution of loss burst lengths.

[1]Corresponding author. Tel: +55 67 33457459; fax: +55 67 33457455. This research was supported by FAPESP, process number 00/05369-2, and CNPq, process number 62.0123/04-4. The email addresses are: isotoma@dct.ufms.br (I. Sotoma) and edmundo@ic.unicamp.br (E. R. M. Madeira).

Sotoma and Madeira [24] propose a Markov model, based on the limited state space model of Sanneck, to model the QoS of failure detectors in the presence of message loss bursts. That work was reformulated by the authors with the unlimited state space model of Sanneck, until NFD-S configurator when message delay distribution is known [25]. In this current paper, we further improve this recent work with a new configurator to NFD-S when message delay distribution is unknown, and a new configurator to NFD-U/NFD-E. This work shows that the new configurator to NFD-U/NFD-E performs better than Chen et al ones, mainly when there are long loss burst lengths, and performs similarly to Chen et al ones when there are no loss bursts. The results also corroborate previous works which indicate that the choice of a message delay estimator has much impact on failure detector QoS and that it is difficult to meet all main QoS requirements at the same time.

This paper is organized as follows. Section 2 shortly describes the basic Sanneck model, the Chen et al work, and related works which use the QoS metrics. Section 3 reviews the Sotoma and Madeira Markov chain model for the configurator to NFD-S. Section 4 proposes a new configurator to NFD-S when message delay distribution is unknown. Section 5 proposes a new configurator for NFD-U/NFD-E for unsynchronized clocks. Section 6 presents the simulation setting and the results with the configurator for NFD-E, and Section 7 offers some conclusions.

2 Background

The following subsections provide a short description of the loss burst model of Sanneck, the quality of service of failure detectors, and some related work on failure detectors and QoS.

2.1 Loss Run-Length Model

The Sanneck [21] model for loss run-length with a Markov chain with unlimited state space (m+1 states) is shown in Figure 1). The random variable X provides information about loss bursts lengths and is defined as follows: $X = 0$ means *no* lost packet, $X = z$ $(0 < z < m)$ means *exactly* z consecutive lost packets, $X \geq z$ means *at least* z consecutive lost packets. The Markov chain state transition occurs according to transition probabilities p_{ij}, with $i < j$ (for loss burst lengths lower than or equal to m) or $i \geq j = 0$ (for a packet arrival). The state probability of the system for $0 < z \leq m$ is $Pr(X \geq z)$, and for $z = 0$ is $Pr(X = 0)$.

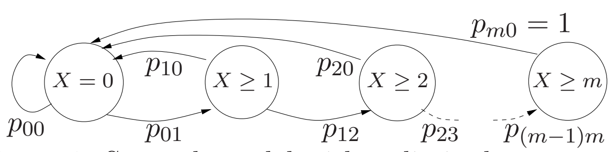

Figure 1: Sanneck model with unlimited state space.

2.2 QoS of Failure Detectors

This Section shortly describes the Chen et al assumptions to their model to QoS of failure detectors. A monitored process p sends heartbeat messages periodically to a failure detector process q which periodically verifies p has crashed. They assume the following probabilistic network model:

1) The link between p and q does not create or duplicate messages, but may delay or drop messages;

2) *message loss probability* p_L is the probability of a message be dropped by the link, and *message delay* D is the delay from the time a message is sent to the time it is received, since the message is not dropped by the link;

3) the delay expected value $E(D)$ and the delay variance $V(D)$ of D are finite;

4) p and q have access to their own local clocks, which have no drift;

5) the probabilistic behavior of the network does not change over time;

6) the crashes can not be predicted;

7) the delay and loss behaviors of the messages that a process sends are independent of whether (and when) the process crashes;

8) the link from p to q has the *message independence* property: the behaviors of any two heartbeat messages sent by p are independent.

At a time t, q can inform S whether it suspects of p crash, or T if it believes that p is alive. A *transition* is an alternated changing between the outputs T and S. An *S-transition* occurs when q informs S just after it has informed T; and a *T-transition* occurs when the output of q changes from S to T. The following QoS metrics are defined as random variables.

Detection time (T_D) is the QoS metric for failure detector speed, and describes how fast q detects the p crash. T_D represents the time that elapses from the time that p crashes to the time when the final *S-transition* (of q) occurs and there are no transitions afterward. There are QoS metrics for *accuracy* which describe how well q avoids mistakes. A mistake occurs when q outputs S, but p is still alive.

The two primary accuracy metrics are: 1) *Mistake recurrence time* (T_{MR}): the time that elapses from an *S-transition* to the next one; and 2) *Mistake duration* (T_M): the time that elapses from an *S-transition* to the next *T-transition.*

Besides these two primary accuracy metrics, there are four accuracy metrics which are derived from T_{MR} and T_M: 1) *Average mistake rate* (λ_M): the rate at which a failure detector mistakes; 2) *Query accuracy probability* (P_A): the probability that the failure detector's output is correct at a random time; 3) *Good period duration* (T_G): the time that elapses from a *T-transition* to the next *S-transition*; and 4) *Forward good period duration* (T_{FG}): the time that elapses from a random time at which q trusts p to the time of the next *S-transition.*

Theorem 1 of Chen et al, at next, explains how the six accuracy metrics are related. $Pr(A)$ denotes the probability of event A; $E(X)$, $E(X^k)$, and $V(X)$ denote the expected value (or mean), the kth moment, and the variance of random variable X, respectively. A failure detector history is a sequence of outputs (S or T) which the failure detector provides. In failure-free runs, an ergodic failure detector is that which outputs histories which follow an ergodic probabilistic distribution. This means that, in failure-free runs, the failure detector slowly "*forgets*" its past history: from any given time on, its future behavior may depend only on its recent behavior.

Theorem 1. *For any ergodic failure detector, the following results hold: 1)* $T_G = T_{MR} - T_M$*. 2) If* $0 < E(T_{MR}) < \infty$*, then* $\lambda_M = 1/E(T_{MR})$ *and* $P_A = E(T_G)/E(T_{MR})$*. 3) If* $0 < E(T_{MR}) < \infty$ *and* $E(T_G) = 0$*, then* T_{FG} *is always 0. If* $0 < E(T_{MR}) < \infty$ *and* $E(T_G) \neq 0$*, then 3a) for all* $x \in [0, \infty)$*,* $Pr(T_{FG} \leq x) = \int_0^x Pr(T_G > y)dy/E(T_G)$*, 3b)* $E(T_{FG}^k = E(T_G^{k+1})/[(k+1)E(T_G)]$*. In particular, 3c)* $E(T_{FG}) = [1 + V(T_G)/E(T_G)^2]E(T_G)/2$*.*

Process p:
for all $i \geq 1$, at time $\sigma_i = i\eta$, send heartbeat m_i to q;
Process q:
Initialization: $output = S$; {suspect p initially}
for all $i \geq 1$, at time $\tau_i = \sigma_i + \delta$:
  if did not receive m_j with $j \geq i$ **then** $output \leftarrow S$; {suspect p if no fresh message is received}
upon receive message m_j at time $t \in [\tau_i, \tau_{i+1})$:
  if $j \geq i$ **then** $output \leftarrow T$; {trust p when some fresh message is received}

Figure 2: Algorithm NFD-S with parameters η and δ (clocks are synchronized).

The NFD-S algorithm of Chen et al which assumes synchronized clocks is shown in Figure 2. NFD-S has two parameters: η and δ. η is the intersending interval between the heartbeat messages $m_1, m_2, \ldots$ which p sends q. So, σ_i is the sending time of the message m_i. q shifts the σ_i's forward by δ to obtain the sequence of times $\tau_1 < \tau_2 < \ldots$, where $\tau_i = \sigma_i + \delta$, for $i \geq 1$. For $i = 0$, $\tau_0 = 0$. If q has received a heartbeat message m_j with $j \geq i$ within time period $[\tau_i, \tau_{i+1})$, then q trusts p. Otherwise, q suspects p.

2.3 Failure Detectors and QoS

In the literature, the main use of QoS metrics of failure detectors is to evaluate failure detector implementations. Bertier et al. [4, 5] use real data to evaluate their failure detector implementations. Bertier et al. [4] show that their proposed dynamic estimator lead to a lower detection time but a low increase in the number of false suspicions and in the average T_M, compared with Chen et al estimator. Bertier et al. [5] use an adaptation procedure based on the Chen et al configurator for NFD-U algorithm, to calculate the failure detector parameters. Then their evaluation shows that the proposed hierarchical failure detection architecture, when attending the QoS of several processes, can lead to a higher detection time, but with much lower false suspicions and average T_M.

Nunes and Jansch-Pôrto [19] use a pull failure detector and Falai and Bondavalli[10] use a push failure detector to evaluate the failure detector behavior under several message delay predictors (estimators) and safety margins. Both works show that in a scenario with constant safety margin and different predictors, a more accurate predictor does not necessarily imply in a lower detection time and in an improvement of the failure detector accuracy. Their experiments also show that in many times a combination predictor-safety margin does not improve the average T_M and average T_{MR} at the same time. Usually, when the combination leads the failure detector to lower the T_M, also lower the T_{MR}.

Hayashibara et al. [13] utilize QoS metrics only to evaluate their proposed accrual failure detector with Chen et al (NFD-E) and Bertier et al [4] ones. In a wide area network, the accrual and NFD-E failure detectors are similar, but the Bertier et al does not work well (because it was designed for a local network (LAN)). In a LAN and a low detection time (below 1 s), the accrual failure detector overcomes the NFD-E failure detector and is similar to Bertier et al one. However, in a LAN and detection time above 1.4 s, the accrual and NFD-E failure detectors are similar.

These previous works do not show if the required QoS is satisfied from the parameters output by the adaptation procedure. The previous Sotoma and Madeira's papers [24, 25] and this paper configure the failure detector from QoS requirements and then verify if the failure detector really satisfies the requirements. Additionally, these works are the first to address explicitly the occurrence of message loss bursts to build configurators which provide parameters to lead the failure detector to satisfy the QoS requirements.

3 A Model of Loss Bursts for Failure Detectors

This Section uses the Chen et al paper [7] as framework, and only describes shortly the Sotoma and Madeira's paper [25]. Further details and proofs can be found in these works.

3.1 Modified Probabilistic Network Model

The probabilistic network model considered in the proposed model is the same of the Chen et al one (see Section 2.2), except by the following changes:

1) Besides the message loss probability (p_L) and message delay (D), the link between p and q also has the additional probability distribution of loss burst lengths, given by all $p_{L,z}$'s. $p_{L,z} = \frac{o_z}{a}$, where z is the length of a loss burst, o_z is the number of loss bursts of length z, and a is the highest valid heartbeat message received.

2) The *message independence* property is not required. There can be either an independent behavior of any two messages, or a dependent behavior of each message only with its predecessor one.

3.2 The Markov Model for Loss Bursts

The folowing subsections provide an overview of the basic Markov chain model, the NFD-S model for loss bursts, and the NFD-S configurator when message delay distribution is known.

3.2.1 The Basic Markov Chain Model

The Markov model of Sanneck [21] (see Section 2) is the basis for the Definition 2. The NSM-NFD-S configurator (proposed in [25]), which is briefly described in Section 3.2.3, assumes the whole information in Definition 2 is already available when the failure detector configuration starts. This is possible due to approximations of the state and state transition probabilities, as described in [25].

Definition 2.

1. *Z_n is a sequence of random variables with values within space $F = \{0, 1\}$. $Z_n = 0$ means a message was received by q, and $Z_n = 1$ means a message was lost.*

2. *h is the highest loss burst length which has been noted by q until the time at which the configuration takes place.*

3. *$S = [0, h]$, with $S \subseteq \mathbf{N}$, is the set of the possible states in the Markov chain.*

4. *$X_{n+1} = f(X_n, Z_{n+1})$ is the random variable which defines a Markov chain, with $X_n \in S$ and X_0 is the first observed state. If $X_n < h$, then $X_{n+1} = Z_{n+1}X_n + Z_{n+1}$; else if $X_n = h$, then $X_{n+1} = 0$, for $Z_{n+1} = 0$. If $X_n = h$ and $Z_{n+1} = 1$ the markov chain is not defined, according to item 2.*

5. *The meaning of the random variable X_{n+1}, defined in item 4, is as follows: $X_{n+1} = 0$ means no lost message, $X_{n+1} = z$ $(0 < z \leq h)$ means exactly z consecutive lost messages, $X_{n+1} \geq z$ means at least z consecutive lost messages.*

6. *The meaning of state and state transition probabilities, defined in item 4, is as follows. State transitions occur depending on transition probabilities p_{ij}, with $i < j$ (for loss burst lengths lower than or equal to h) or $i \geq j = 0$ (for a message arrival). The state probability of the system for $0 < z \leq h$ is $Pr(X_{n+1} \geq z)$, and for $z = 0$ is $Pr(X_{n+1} = 0)$.*

The Definitions 3 and 4, at next, simplifly the notation for the state transition probabilities of the Markov chain of the Definition 2. The Definition 3 shows the probability of performing state transitions corresponding to a loss burst of length $bs - es + 1$. The Definition 4 shows the probability of performing a transition to state 0 when a message is received after a loss burst with length i, or after another previously received message.

Definition 3. *The probability of forward state transitions, from a state bs to a state $es \geq bs$, is defined as $forw(bs, es) = \prod_{n=bs}^{es-1} p_{n(n+1)}$. From the* Definition 2, *when $n \geq h$, $forw(bs, es) = 0$.*

Definition 4. *The probability of a backward state transition, from a state i to the state 0 is defined as $to0(i) = p_{i0}$, from the* Definition 2. *$p_{i0} = 1 - p_{i(i+1)}$, for $i < h$, and $p_{h0} = 1$, for $i = h$.*

3.2.2 The NFD-S Model to Cope Loss Bursts

The Definition 5 uses the Markov chain of the Definition 2 for QoS of failure detectors in the presence of loss bursts to characterize the probability of an *S-transition* occurs in time τ_i.

Definition 5.

1. *For any $i \geq 1$, let k be the smallest integer such that, for all $j \geq i + k$, m_j is sent at or after time τ_i.*

2. *For any $i \geq 2$, let q_0 be the probability that q receives the message m_{i-1} before time τ_i. In this case, the Markov chain is in state 0.*

3. *For any $i \geq 1$, let $u(x)$ be the probability that q suspects p, by receiving no one of the messages m_{i+j}, for every $0 \leq j \leq k - 1$, at time $\tau_i + x$, for all $x \in [0, \eta)$. This definition assumes the Markov chain is already in state 0 (definitions 5.2 and 5.4). Therefore, from state 0, the Markov chain takes transitions.*

4. *For any $i \geq 2$, let p_s be the probability that an S-transition occurs at time τ_i. This characterizes the whole Markov chain.*

The Proposition 6 mathematically describes the Definition 5 independently of i.

Proposition 6.

1. $k = \lceil \delta/\eta \rceil$.

2. $q_0 = Pr(X_{n+1} = 0)Pr(D < \delta + \eta)$.

3. *For all $x \in [0, \eta)$, and w initially equal to k, $u(x) = u_w(x)$. $u_w(x)$ is defined as follows:*

$$
\begin{aligned}
u_1(x) &= forw(0,1) + to0(0)Pr(D > \delta + x - (k-1)\eta), \text{ for } w = 1;\\
u_w(x) &= to0(0)Pr(D > \delta + x - (k-w)\eta)u_{w-1}(x)\\
&\quad + \sum_{a=1}^{w-2} forw(0,a)to0(a)Pr(D > \delta + x - (a+k-w)\eta)u_{w-(a+1)}(x)\\
&\quad + forw(0,w-1)to0(w-1)Pr(D > \delta + x - (k-1)\eta)\\
&\quad + forw(0,w), \text{ for } w > 1.
\end{aligned}
$$

4. $p_s = q_0 u(0)$.

The Proposition 6.3 is a recursive function which generates all permutations of 0's and 1's, for a given k, by making transitions in the Markov chain according to message losses (1's), and message receipts (0's). The proof of Proposition 6, an example of Proposition 6.3, and further comments of the following propositions can be found in [25].

Proposition 7. *The nondegenerated cases $q_0 > 0$ and $u(0) > 0$, in the proposed model, occur when $Pr(D > \delta) > 0$, $Pr(D < \delta + \eta) > 0$, $0 < Pr(X_{n+1} = 0) < 1$, and $0 < Pr(X_{n+1} \geq 1) < 1$.*

The Definition 8 characterizes the probability for q to suspect p, and the Proposition 9 mathematically describes this probability.

Definition 8. *For any $i \geq 1$, let $v(x)$ be the probability that q suspects p at time $\tau_i + x$, for every $x \in [0, \eta)$. This suspicion occurs when no one of the messages m_{i+j} is received by time $\tau_i + x$, for every $0 \leq j \leq k-1$. Our $u'(x)$ assumes the Markov chain can be in any initial state $s \in S$.*

Proposition 9. *$v(x) = Pr(X_{n+1} = 0)v_{0,k}(x) + \sum_{s=1}^{h} Pr(X_{n+1} \geq s)v_{s,k}(x)$. $u_{w-1}(x)$ and $u_{w-(a+1)}(x)$ use the $u(x)$ definition in the Proposition 6 and the Definition 8. $v_{s,w}(x)$, with w initially equal to k, is defined as follows:*

$$
\begin{aligned}
v_{s,1}(x) &= forw(s,s+1) + to0(s)Pr(D > \delta + x - (k-1)\eta), \text{ for } w = 1;\\
v_{s,w}(x) &= to0(s)Pr(D > \delta + x - (k-w)\eta)u_{w-1}(x)\\
&\quad + \sum_{a=1}^{w-2} forw(s,s+a)to0(s+a)Pr(D > \delta + x - (a+k-w)\eta)u_{w-(a+1)}(x)\\
&\quad + forw(s,s+w-1)to0(s+w-1)Pr(D > \delta + x - (k-1)\eta)\\
&\quad + forw(s,s+w), \text{ for } w > 1.
\end{aligned}
$$

The following theorem summarizes our QoS analysis of the NFD-S, under loss bursts.

Theorem 16. *Consider a system with synchronized clocks, where the probability of message loss p_L, the distribution of message delays $Pr(D \leq x)$, and the probability distribution* *of loss burst lengths are known. The failure detector NFD-S with parameters η and δ has the following properties:*

1. *The detection time is bounded as follows and the bound is tight:* $T_D \leq \delta + \eta$. (1.1)
2. *The average mistake recurrence time is:* $E(T_{MR}) = \frac{\eta}{p_s}$. (1.2)
3. *The average mistake duration is:* $E(T_M) = \frac{\int_0^\eta v(x)d(x)}{p_s}$. (1.3)

3.2.3 The Proposed NSM-NFD-S Configurator to Cope with Loss Bursts

Our goal is to find a configuration procedure, which takes as input the probabilistic behavior of heartbeats and the QoS requirements (T_D^U, T_{MR}^L, T_M^U), and outputs η and δ. Hereafter, we call configurator as a short for configuration procedure. T_D^U is an upper bound on the detection time, T_{MR}^L is a lower bound on the average mistake recurrence time, and T_M^U is an upper bound on the average mistake duration. Then, the QoS requirements are that:

$T_D \leq T_D^U, E(T_{MR}) \geq T_{MR}^L, E(T_M) \leq T_M^U$. (1.4)

From the Theorem 16, the goal can be restated as a mathematical programming problem:

maximize η:

subject to $\delta + \eta \leq T_D^U$ (1.5)

$\frac{\eta}{p_s} \geq T_{MR}^L$ (1.6)

$\frac{\int_0^\eta v(x)dx}{p_s} \leq T_M^U$ (1.7)

where the value of $v(x)$ is given by the Proposition 9, and the value of p_s is given by the Proposition 6. Similar to Chen et al, the problem (1.7), which is hard to solve, was replaced by a simpler and stronger constraint as follows.

Proposition 17. *In the nondegenerated cases of the Proposition 7,* $E(T_M) \leq \frac{v(0)\eta}{q_0 u(0)}$.

From the problem (1.5) and Propositions 6 and 17, we obtain the following Proposition 18, which is used later on by the NSM-NFD-S configurator.

Proposition 18. *Let be* $k' = \lceil T_D^U/\eta \rceil - 1$*. At next,* $v'(0)$ *and* $u'(0)$ *consider, like Chen et al, only the messages* 0 *to* $k-1$. $v'(0) = Pr(X_{n+1} = 0)v_{0,k'}(0) + \sum_{s=1}^{h} Pr(X_{n+1} = s)v'_{s,k'}(0)$. $v'_{s,w}(0)$*, which is based on Proposition 9, is defined as follows:*

$$
\begin{aligned}
v'_{s,1}(0) &= forw(s, s+1) + to0(s)Pr(D > T_D^U - k'\eta), \text{ for } w = 1; \\
v'_{s,w}(0) &= to0(s)Pr(D > T_D^U - (k' - w + 1)\eta)u'_{w-1}(0) \\
&\quad + \sum_{a=1}^{w-2} forw(s, s+a)to0(s+a)Pr(D > T_D^U - (a + k' - w + 1)\eta)u'_{w-(a+1)}(0) \\
&\quad + forw(s, s+w-1)to0(s+w-1)Pr(D > T_D^U - k'\eta) \\
&\quad + forw(s, s+w), \text{ for } w > 1.
\end{aligned}
$$

The terms $u'_{w-1}(0)$ *and* $u'_{w-(a+1)}(0)$ *of* $v'_{s,w}(0)$ *use the following* $u'(0)$ *definition, which is based on Proposition 6:*

$$
\begin{aligned}
u'_1(0) &= forw(0,1) + to0(0)Pr(D > T_D^U - k'\eta), \text{ for } w = 1;\\
u'_w(0) &= to(0)Pr(D > T_D^U - (k' - w + 1)\eta)u'_{w-1}(0)\\
&\quad + \sum_{a=1}^{w-2} forw(0,a)to0(a)Pr(D > T_D^U - (a + k' - w + 1)\eta)u'_{w-(a+1)}(0)\\
&\quad + forw(0, w-1)to0(w-1)Pr(D > T_D^U - k'\eta)\\
&\quad + forw(0,w), \text{ for } w > 1.
\end{aligned}
$$

NSM-NFD-S (New Sotoma and Madeira configurator for NFD-S) Configurator: From the assertions (1.4), (1.5), (1.6), (1.7) and Propositions 17 and 18, we obtain the following configurator, called NSM-NFD-S configurator, to find η and δ:
Step 1: Compute $q'_0 = Pr(X_{n+1} = 0)Pr(D < T_D^U)$ and let $g(\eta) = v'(0)\eta/q'_0u'(0)$, where $v'(0) = v'_{k'}(0)$. If $q'_0u'(0) = 0$, then output *"QoS cannot be achieved"* and stop. Otherwise, find the largest $\eta_{max} \leq T_D^U$ such that $g(\eta_{max}) \leq T_M^U$.
Step 2: Let $f(\eta) = \eta/q'_0u'(0)$, find the largest $\eta \leq \eta_{max}$ such that $f(\eta) \geq T_{MR}^L$.
Step 3: Set $\delta = T_D^U - \eta$ and output η and δ.

Theorem 19. *Consider a system in which clocks are synchronized and the probability of message loss p_L, the distribution of message delays $Pr(D \leq x)$, and the probability distribution of loss burst lengths are known. Suppose we are given a set of QoS requirements as in (1.4). The NSM-NFD-S configurator has two possible outcomes: 1) It outputs η and δ. In this case, with parameters η and δ, the failure detector NFD-S satisfies the given QoS requirements. 2) It outputs "QoS cannot be achieved".*

4 Unknown Delay Distribution and NFD-S

Hereafter we present new contributions to previous works. This Section provides a new configurator to NFD-S which computes η and δ when $Pr(D \leq x)$ is unknown. We use, similarly to Chen et al, the following *One-Sided Inequality*: For any random variable D with finite expectance and finite variance, $Pr(D > t) \leq \frac{V(D)}{V(D)+(t-E(D))^2}$, for all $t > E(D)$. In short, the new configurator to be presented in this Section has more practical use because it needs the mean and variance message delay rather than the knowledge of message delay distribution (e.g. exponential, normal). By applying this inequality on Propositions 6, 9 and 18, we obtain the Proposition 20 and Heuristic 21, at next.

Proposition 20. *Let be $k_0 = \lceil(\delta - E(D))/\eta\rceil$. In the following, $u^{\geq}(0)$ and $v^{\geq}(0)$ consider the messages 0 to $k_0 - 1$, and w is initially k_0. $v^{\geq}(0) = Pr(X_{n+1} = 0)v^{\geq}_{0,k_0}(0) + \sum_{s=1}^{h} Pr(X_{n+1} \geq s)v^{\geq}_{s,k_0}(0)$. $v^{\geq}_{s,w}(0)$, which is based on Proposition 9, is defined as follows:*

$$
\begin{aligned}
v^{\geq}_{s,1}(0) &= forw(s,s+1) + to0(s)(V(D)/(V(D) + (\delta - E(D) - (k_0 - 1)\eta)^2)), \text{ for } w = 1;\\
v^{\geq}_{s,w}(0) &= to0(s)(V(D)/(V(D) + (\delta - E(D) - (k_0 - w)\eta)^2))u^{\geq}_{w-1}(0)\\
&\quad + \sum_{a=1}^{w-2} forw(s,s+a)to0(s+a)(V(D)/(V(D) + (\delta - E(D) - (a + k_0 - w)\eta)^2))u^{\geq}_{w-(a+1)}(0)\\
&\quad + forw(s, s+w-1)to0(s+w-1)(V(D)/(V(D) + (\delta - E(D) - (k_0 - 1)\eta)^2))\\
&\quad + forw(s, s+w), \text{ for } w > 1.
\end{aligned}
$$

The terms $v^{\geq}_{w-1}(0)$ and $v^{\geq}_{w-(a+1)}(0)$ of $v^{\geq}_{s,w}(0)$ use the following definition of $u^{\geq}(0)$, which is based on Proposition 6:

$$\begin{aligned}
u^{\geq}_1(0) &= forw(0,1) + to0(0)(V(D)/(V(D) + (\delta - E(D) - (k_0 - 1)\eta)^2)), \text{ for } w = 1;\\
u^{\geq}_w(0) &= to(0)(V(D)/(V(D) + (\delta - E(D) - (k_0 - w)\eta)^2))u^{\geq}_{w-1}(0)\\
&\quad + \sum_{a=1}^{w-2} forw(0,a)to0(a)(V(D)/(V(D) + (\delta - E(D) - (a + k_0 - w)\eta)^2))u^{\geq}_{w-(a+1)}(0)\\
&\quad + forw(0, w-1)to0(w-1)(V(D)/(V(D) + (\delta - E(D) - (k_0 - 1)\eta)^2))\\
&\quad + forw(0, w), \text{ for } w > 1.
\end{aligned}$$

Proof. For $w = 1$:

$$\begin{aligned}
u_1(0) &= forw(0,1) + to0(0)Pr(D > \delta - (k-1)\eta)(\text{By One-Sided Inequality:})\\
&\leq forw(0,1) + to0(0)(V(D)/(V(D) + (\delta - (k-1)\eta - E(D))^2), \text{ for } k = \lceil \delta/\eta \rceil.
\end{aligned}$$

By One-Sided Inequality definition we have to guarantee that:

$$\delta - (k-1)\eta > E(D) \Leftrightarrow k < \frac{\delta - E(D)}{\eta} + 1 \Rightarrow k_0 = \left\lceil \frac{\delta - E(D)}{\eta} + 1 \right\rceil - 1$$

$$\Leftrightarrow k_0 = \left\lceil \frac{\delta - E(D)}{\eta} \right\rceil.$$

So, by using k_0 we have: $forw(0,1) + to0(0)(V(D)/(V(D) + (\delta - (k_0 - 1)\eta - E(D))^2) = u^{\geq}_1(0)$, for $k_0 = \lceil (\delta - E(D))/\eta \rceil$. For $w > 1$:

$$\begin{aligned}
u_w(0) &= to0(0)Pr(D > \delta - (k-w)\eta)u_{w-1}(0)\\
&\quad + \sum_{a=1}^{w-2} forw(0,a)to0(a)Pr(D > \delta - (a + k - w)\eta)u_{w-(a+1)}(0)\\
&\quad + forw(0, w-1)to0(w-1)Pr(D > \delta - (k-1)\eta)\\
&\quad + forw(0, w)(\text{By One-Sided Inequality:})\\
&\leq to0(0)(V(D)/(V(D) + (\delta - (k-w)\eta - E(D))^2))u_{w-1}(0)\\
&\quad + \sum_{a=1}^{w-2} forw(0,a)to0(a)(V(D)/(V(D) + (\delta - (a + k - w)\eta - E(D))^2))u_{w-(a+1)}(0)\\
&\quad + forw(0, w-1)to0(w-1)(V(D)/(V(D) + (\delta - (k-1)\eta - E(D))^2)\\
&\quad + forw(0, w), \text{ for } k = \lceil \delta/\eta \rceil.
\end{aligned}$$

By One-Sided Inequality definition we have to guarantee that:

$\delta - (k-1)\eta > E(D), \quad \delta - (k-w)\eta > E(D), \quad \delta - (a + k - w)\eta > E(D),$

in other words, it is enough to guarantee that $\delta - (k-1)\eta > E(D)$, which leads to

$k_0 = \left\lceil \dfrac{\delta - E(D)}{\eta} \right\rceil$ as shown before. Therefore, by using k_0 we have:

$$
\begin{aligned}
& to0(0)(V(D)/(V(D)+(\delta-(k_0-w)\eta-E(D))^2))u_{w-1}(0) \\
& +\sum_{a=1}^{w-2} forw(0,a)to0(a)(V(D)/(V(D)+(\delta-(a+k_0-w)\eta-E(D))^2))u_{w-(a+1)}(0) \\
& +forw(0,w-1)to0(w-1)(V(D)/(V(D)+(\delta-(k_0-1)\eta-E(D))^2) \\
& +forw(0,w), \text{ for } k_0=\lceil(\delta-E(D))/\eta\rceil; \\
& \leq to0(0)(V(D)/(V(D)+(\delta-(k_0-w)\eta-E(D))^2))u^{\geq}_{w-1}(0) \\
& +\sum_{a=1}^{w-2} forw(0,a)to0(a)(V(D)/(V(D)+(\delta-(a+k_0-w)\eta-E(D))^2))u^{\geq}_{w-(a+1)}(0) \\
& +forw(0,w-1)to0(w-1)(V(D)/(V(D)+(\delta-(k_0-1)\eta-E(D))^2) \\
& +forw(0,w), \text{ for } k_0=\lceil(\delta-E(D))/\eta\rceil; \\
& =u^{\geq}_{w}(0).
\end{aligned}
$$

By analogy, the proof of $v^{\geq}(0)$ follows directly from the proof of $u^{\geq}(0)$. □

From Proposition 20 we obtain Heuristic 21 at next.

Heuristic 21. *Consider a system with synchronized clocks and assume that $\delta > E(D)$. For the NFD-S algorithm, we have $E(T_{MR}) \geq \eta/\beta$ e $E(T_M) \approx v^{\geq}(0)\eta/\gamma$, where:*

$$
\begin{aligned}
& \beta = Pr(X_{n+1}=0)u^{\geq}(0), \\
& \gamma = Pr(X_{n+1}=0)\left(\frac{(\delta+\eta-E(D))^2}{(V(D)+(\delta+\eta-E(D))^2)}\right)u^{\geq}(0), \\
& k_0=\lceil(\delta-E(D))/\eta\rceil.
\end{aligned}
$$

Remark. The definition of k_0 comes from Proposition 20. We comment first about $E(T_{MR}) \geq \eta/\beta$:

$$
\begin{aligned}
E(T_{MR}) &= \frac{\eta}{p_s} \\
&= \frac{\eta}{Pr(X_{n+1}=0)Pr(D<\delta+\eta)u(0)} \text{(From Theorem 16.2 and Proposition 6)} \\
&\geq \frac{\eta}{Pr(X_{n+1}=0)u(0)} \geq \frac{\eta}{Pr(X_{n+1}=0)u^{\geq}(0)} \text{(From Proposition 20)} \\
&= \eta/\beta.
\end{aligned}
$$

So, if $\eta/\beta \geq T^L_{MR}$, then $E(T_{MR}) \geq T^L_{MR}$. In the following we comment about $E(T_M) \leq v^{\geq}(0)\eta/\gamma$:

$$
\begin{aligned}
E(T_M) &\leq \frac{v(0)\eta}{q(0)u(0)} \text{(From Proposition 17)}\\
&= \frac{v(0)\eta}{Pr(X_{n+1}=0)Pr(D<\delta+\eta)u(0)} \text{(From Proposition 6)}\\
&\text{Then, from Proposition 20, One-Sided Inequality and}\\
&\text{use of } k_0 \text{, we use the following approximation:}\\
&\frac{v^{\geq}(0)\eta}{Pr(X_{n+1}=0)\left(\frac{(\delta+\eta-E(D))^2}{(V(D)+(\delta+\eta-E(D))^2)}\right)u^{\geq}(0)}\\
&= v^{\geq}(0)\eta/\gamma.
\end{aligned}
$$

If $v^{\geq}(0)\eta/\gamma \approx T_M^U$, then $E(T_M) \approx T_M^U$. □

From assertion (1.5) and Proposition 20, we obtain the Proposition 22 at next.

Proposition 22. *Let be* $k_0' = \lceil (T_D^U - E(D))/\eta \rceil - 1$*. At next,* $u^{\geq'}(0)$ *and* $v^{\geq'}(0)$ *consider the messages* 0 *to* $k_0'-1$*, and* w *is initially* k_0'*.* $v^{\geq'}(0) = Pr(X_{n+1}=0)v^{\geq'}_{0,k_0}(0) + \sum_{s=1}^{h} Pr(X_{n+1} \geq s)v^{\geq'}_{s,k_0}(0)$*.* $v^{\geq'}_{s,w}(0)$*, which is based on Proposition 20, is defined as follows:*

$$
\begin{aligned}
v^{\geq'}_{s,1}(0) &= forw(s,s+1) + to0(s)(V(D)/(V(D)+(T_D^U - E(D) - k_0'\eta)^2)), \text{ for } w=1;\\
v^{\geq'}_{s,w}(0) &= to0(s)(V(D)/(V(D)+(T_D^U - E(D) - (k_0'-w+1)\eta)^2))u^{\geq'}_{w-1}(0)\\
&+ \sum_{a=1}^{w-2} forw(s,s+a)to0(s+a)(V(D)/(V(D)+(T_D^U - E(D) - (a+k_0'-w+1)\eta)^2))u^{\geq'}_{w-(a+1)}(0)\\
&+ forw(s,s+w-1)to0(s+w-1)(V(D)/(V(D)+(T_D^U - E(D) - k_0'\eta)^2))\\
&+ forw(s,s+w), \text{ for } w>1.
\end{aligned}
$$

The term $v^{\geq'}_{w-1}(0)$ *and* $v^{\geq'}_{w-(a+1)}(0)$ *of* $v^{\geq'}_{s,w}(0)$ *use the following definition of* $u^{\geq'}(0)$*, which is based on Proposition 6:*

$$
\begin{aligned}
u^{\geq'}_{1}(0) &= forw(0,1) + to0(0)(V(D)/(V(D)+(T_D^U - E(D) - k_0'\eta)^2)), \text{ for } w=1;\\
u^{\geq'}_{w}(0) &= to(0)(V(D)/(V(D)+(T_D^U - E(D) - (k_0'-w+1)\eta)^2))u^{\geq'}_{w-1}(0)\\
&+ \sum_{a=1}^{w-2} forw(0,a)to0(a)(V(D)/(V(D)+(T_D^U - E(D) - (a+k_0'-w+1)\eta)^2))u^{\geq'}_{w-(a+1)}(0)\\
&+ forw(0,w-1)to0(w-1)(V(D)/(V(D)+(T_D^U - E(D) - k_0'\eta)^2))\\
&+ forw(0,w), \text{ for } w>1.
\end{aligned}
$$

NSM-OSI-NFD-S (New Sotoma and Madeira configurator for NFD-S with One-Sided Inequality) Configurator: From assertions (4.4), (4.5), (4.6), (4.7), Heuristic 21 and Proposition 22, we obtain the following configurator, called NSM-OSI-NFD-S configurator, to find η and δ:

Step 1: Make $g(\eta) = v^{\geq'}(0)\eta/\gamma'$, for $\gamma' = Pr(X_{n+1}=0)\left(\frac{(T_D^U - E(D))^2}{(V(D)+(T_D^U - E(D))^2)}\right)u^{\geq'}(0)$. If $\gamma' = 0$, then output *"QoS can not be met"* and stop. Otherwise, find the greatest $\eta_{max} \leq T_D^U$

such that $g(\eta_{max}) \leq T_M^U$.
Step 2: Let be $f(\eta) = \eta/\beta'$, for $\beta' = Pr(X_{n+1} = 0)u^{\geq'}(0)$. Find the greatest $\eta \leq \eta_{max}$ such that $f(\eta) \geq T_{MR}^L$.
Step 3: Assign $\delta = T_D^U - \eta$ and output η and δ.

Theorem 23. *Consider a system with synchronized clocks such that the probability of message loss p_L and the probability distribution of message loss bursts are known, and the message delay distribution $Pr(D \leq x)$ is unknown. Assume that is given a set of QoS requirements as in (4.4). The NSM-OSI-NFD-S configurator has two possible outcomes: 1) It outputs η and δ. In this case, with parameters η and δ, the failure detector NFD-S aproximmates to the QoS requirements. 2) It outputs "QoS can not be met".*

Proof. We prove Theorem 23 in the following two parts:

1. Assume that the NSM-OSI-NFD-S configurator outputs *"QoS can not be met"*. Then, the configuratdor stops in *Step 1* because $\gamma' = 0$. $\gamma' = 0$ implies that $Pr(X_{n+1} = 0) = 0$, $\left(\frac{(T_D^U - E(D))^2}{(V(D) + (T_D^U - E(D))^2)}\right) = 0$, or $u^{\geq'}(0) = 0$. $Pr(X_{n+1} = 0) = 1 - p_L = 0$ implies that when q suspects p, $E(T_M) = \infty$ and so, q fails to met $E(T_M) \approx T_M^U$. $\left(\frac{(T_D^U - E(D))^2}{(V(D) + (T_D^U - E(D))^2)}\right) = 0$ implies that $T_D^U = E(D)$, and $u^{\geq'}(0) = 0$ implies which there is no suspicion. In these two cases, $E(T_M) = g(\eta) = \infty$, which leads q to fail to satisfy $E(T_M) \approx T_M^U$. Therefore, the failure detector q can not satisfy the requested QoS in this case.

2. Assume that the NSM-OSI-NFD-S configurator outputs the parameters η and δ. Then, from *Step 3*, we have $T_D^U = \eta + \delta$. By part 1 of Theorem 16, $T_D \leq T_D^U$ is satisfied. In this case $\gamma' > 0$ because otherwise $g(\eta) = \infty$, and the NSM-OSI-NFD-S configurator would output *"QoS can not be met"* in place of η and δ. From Heuristic 21, Proposition 17 and *Step 1*, $E(T_M) \approx g(\eta) = v^{\geq'}(0)\eta/\gamma' \approx \frac{v(0)\eta}{q_0 u(0)} \leq T_M^U$. So, $E(T_M)$ aproximmates to T_M^U. $f(\eta) = \eta/\beta'$, from restriction (4.2) and from Heuristic 21, $E(T_{MR}) \geq f(\eta) = \eta/\beta' = \eta/\beta$. Then from *Step 2*, $E(T_{MR}) \geq T_{MR}^L$ is satisfied. □

5 Model of Configurator for NFD-U

The NFD-S failure detector and the proposed configurators NSM-NFD-S and NSM-OSI-NFD-S assume synchronized clocks (for example, by using NTP (Network Time Protocol)). This section shows how to build an NSM-NFD-U configurator for the NFD-U failure detector. NSM-NFD-U assumes that the clocks are not synchronized. The Chen et al NFD-U failure detector is shown in Figure 3.

Similarly to Chen et al, the QoS analysis of NFD-U is obtained from replacement of δ by $E(D) + \alpha$ in Proposition 20 and in Heuristic 21. Additionally, the NFD-U failure detector must satisfy the following QoS requirements, the same of Chen et al:

$$T_D \leq T_D^u + E(D), \quad E(T_{MR}) \geq T_{MR}^L, \quad E(T_M) \leq T_M^U. \tag{4.8}$$

The real upper bound on detection time T_D is not T_D^u, but T_D^u plus an *unknown* the message average delay $E(D)$. So, T_D^U is $T_D^u + E(D)$.

By applying $\delta = E(D) + \alpha$ in Proposition 20, we obtain the following Proposition 24.

Proposition 24. *Let be $k_0 = \lceil \alpha/\eta \rceil$. In the following, $u^{\geq}(0)$ and $v^{\geq}(0)$ consider messages 0 to $k_0 - 1$, and w is initially k_0.* $v^{\geq}(0) = Pr(X_{n+1} = 0)v_{0,k_0}^{\geq}(0) + \sum_{s=1}^{h} Pr(X_{n+1} \geq$

Process p: {using p's local clock}
for all $i \geq 1$, at time $i\eta$, send heartbeat m_i to q;
Process q: {using q's local clock}
Initialization:
  $\tau_0 = 0$;
  $l = -1$; {l keeps the largest sequence number in all messages q received so far}
upon τ_{l+1} = current time: {if the current time reaches τ_{l+1}, then none of the received messages is still fresh}
  $output \leftarrow S$; {suspect p since no received message is still fresh at this time}
upon receive message m_j at time t:
  if $j > l$ **then** {received a message with a higher sequence number}
    $l \leftarrow j$;
    $\tau_{l+1} \leftarrow EA_{l+1} + \alpha$; {set the next τ_{l+1} using the expected arrival time of m_{l+1}}
    if $t < \tau_{l+1}$ **then** $output \leftarrow T$; {trust p since m_l is still fresh at time t}

Figure 3: NFD-U algorithm with parameters η and α (unsynchronized clocks, but EA_{l+1}'s are known).

$s)v^{\geq}_{s,k_0}(0)$. $v^{\geq}_{s,w}(0)$, which is based on Proposition 9, is defined as follows:

$$\begin{aligned}
v^{\geq}_{s,1}(0) &= forw(s,s+1) + to0(s)(V(D)/(V(D)+(\alpha-(k_0-1)\eta)^2)), \text{ for } w=1;\\
v^{\geq}_{s,w}(0) &= to0(s)(V(D)/(V(D)+(\alpha-(k_0-w)\eta)^2))u^{\geq}_{w-1}(0)\\
&\quad + \sum_{a=1}^{w-2} forw(s,s+a)to0(s+a)(V(D)/(V(D)+(\alpha-(a+k_0-w)\eta)^2))u^{\geq}_{w-(a+1)}(0)\\
&\quad + forw(s,s+w-1)to0(s+w-1)(V(D)/(V(D)+(\alpha-(k_0-1)\eta)^2))\\
&\quad + forw(s,s+w), \text{ for } w>1.
\end{aligned}$$

The terms $v^{\geq}_{w-1}(0)$ and $v^{\geq}_{w-(a+1)}(0)$ of $v^{\geq}_{s,w}(0)$ use the following definition of $u^{\geq}(0)$, which is based on Proposition 6:

$$\begin{aligned}
u^{\geq}_{1}(0) &= forw(0,1) + to0(0)(V(D)/(V(D)+(\alpha-(k_0-1)\eta)^2)), \text{ for } w=1;\\
u^{\geq}_{w}(0) &= to(0)(V(D)/(V(D)+(\alpha-(k_0-w)\eta)^2))u^{\geq}_{w-1}(0)\\
&\quad + \sum_{a=1}^{w-2} forw(0,a)to0(a)(V(D)/(V(D)+(\alpha-(a+k_0-w)\eta)^2))u^{\geq}_{w-(a+1)}(0)\\
&\quad + forw(0,w-1)to0(w-1)(V(D)/(V(D)+(\alpha-(k_0-1)\eta)^2))\\
&\quad + forw(0,w), \text{ for } w>1.
\end{aligned}$$

By applying $\delta = E(D) + \alpha$ in Heuristic 21, we obtain the following Heuristic 25.

Heuristic 25. *Consider a system with unsynchronized clocks and assume that $\alpha > 0$.*

For the NFD-U algorithm, we have $E(T_{MR}) \geq \eta/\beta$ *and* $E(T_M) \approx v^{\geq}(0)\eta/\gamma$*, where:*

$$\begin{aligned}
\beta &= Pr(X_{n+1}=0)u^{\geq}(0),\\
\gamma &= Pr(X_{n+1}=0)\left(\frac{(\alpha+\eta)^2}{(V(D)+(\alpha+\eta)^2)}\right)u^{\geq}(0),\\
k_0 &= \lceil \alpha/\eta \rceil.
\end{aligned}$$

By using (4.5), $\delta+\eta \leq T_D^U$, $T_D^U = T_D^u + E(D)$, and $\delta = E(D)+\alpha$, we have that $\alpha+\eta = T_D^u$. By using this in Proposition 24 we obtain the following Proposition 26.

Proposition 26. *Let be* $k_0' = \lceil T_D^u/\eta \rceil - 1$*. In the following,* $u^{\geq'}(0)$ *and* $v^{\geq'}(0)$ *consider messages* 0 *to* $k_0'-1$*.* $v^{\geq'}(0) = Pr(X_{n+1}=0)v^{\geq'}_{0,k_0'}(0) + \sum_{s=1}^{h} Pr(X_{n+1}\geq s)v^{\geq'}_{s,k_0'}(0)$*.* $v^{\geq'}_{s,w}(0)$*, which is based on Proposition 9, is defined as follows:*

$$\begin{aligned}
v^{\geq'}_{s,1}(0) &= forw(s,s+1) + to0(s)(V(D)/(V(D)+(T_D^u-k_0'\eta)^2)), \text{ for } w=1;\\
v^{\geq'}_{s,w}(0) &= to0(s)(V(D)/(V(D)+(T_D^u-(k_0'-w+1)\eta)^2))u^{\geq'}_{w-1}(0)\\
&\quad+\sum_{a=1}^{w-2} forw(s,s+a)to0(s+a)(V(D)/(V(D)+(T_D^u-(a+k_0'-w+1)\eta)^2))u^{\geq'}_{w-(a+1)}(0)\\
&\quad+forw(s,s+w-1)to0(s+w-1)(V(D)/(V(D)+(T_D^u-k_0'\eta)^2))\\
&\quad+forw(s,s+w), \text{ for } w>1.
\end{aligned}$$

The terms $v^{\geq'}_{w-1}(0)$ *and* $v^{\geq'}_{w-(a+1)}(0)$ *of* $v^{\geq'}_{s,w}(0)$ *use the following definition of* $u^{\geq'}(0)$*, which is based on Proposition 6:*

$$\begin{aligned}
u_1^{\geq'}(0) &= forw(0,1) + to0(0)(V(D)/(V(D)+(T_D^u-k_0'\eta)^2)), \text{ for } w=1;\\
u_w^{\geq'}(0) &= to(0)(V(D)/(V(D)+(T_D^u-(k_0'-w+1)\eta)^2))u^{\geq'}_{w-1}(0)\\
&\quad+\sum_{a=1}^{w-2} forw(0,a)to0(a)(V(D)/(V(D)+(T_D^u-(a+k_0'-w+1)\eta)^2))u^{\geq'}_{w-(a+1)}(0)\\
&\quad+forw(0,w-1)to0(w-1)(V(D)/(V(D)+(T_D^u-k_0'\eta)^2))\\
&\quad+forw(0,w), \text{ for } w>1.
\end{aligned}$$

NSM-NFD-U Configurator: From assertions (4.4), (4.5), (4.6), (4.7), Heuristic 25 and Proposition 26, we obtain the following configurator, called NSM-NFD-U configurator, to find η and α:
Step 1: Make $g(\eta) = v^{\geq'}(0)\eta/\gamma'$, for $\gamma' = Pr(X_{n+1}=0)\left(\frac{(T_D^u)^2}{(V(D)+(T_D^u)^2)}\right)u^{\geq'}(0)$. If $\gamma'=0$, then output *"QoS can not be met"* and stop. Otherwise, find the greater $\eta_{max} \leq T_D^u$ such that $g(\eta_{max}) \leq T_M^U$.
Step 2: Let be $f(\eta) = \eta/\beta'$, for $\beta' = Pr(X_{n+1}=0)u^{\geq'}(0)$. Find the greater $\eta \leq \eta_{max}$ such that $f(\eta) \geq T_{MR}^L$.
Step 3: Assign $\alpha = T_D^u - \eta$ and output η and α.

Figure 4 presents a scheme of the NFD-U configuration. In this case the EA_{l+1}'s are known. The estimator of the probabilistic behaviour of *heartbeats* provides estimates for p_L, $V(D)$ and $p_{L,z}$'s. The NSM-NFD-U configurator uses these estimates with the QoS requirements and then outputs the parameters η and α.

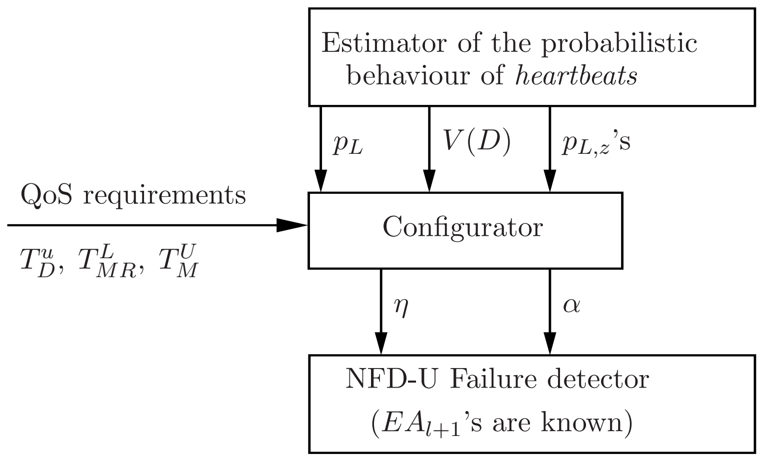

Figure 4: Configuration of NFD-U by NSM-NFD-U.

Theorem 27. *Consider a system with unsynchronized clocks and the message loss probability p_L and the probability distribution of message loss bursts are known, and the message delay probability distribution $Pr(D \leq x)$ is unknown. Assume that a set of QoS requirements is given as is (4.8). The NSM-NFD-U configurator has two possible outcomes: 1) It outputs η and α. In this case, with parameters η and α, the NFD-U failure detector aproximmates to the QoS requirements. 2) It outputs "QoS can not be met".*

A proof for Theorem 27 can be derived directly from the proof of Theorem 23.

Figure 5 presents a scheme of the NFD-E configuration. The only difference with the NFD-U configuration (Figure 4) is that the EA_{l+1}'s are estimated. In the following section, several possible estimators of EA_{l+1} are presented.

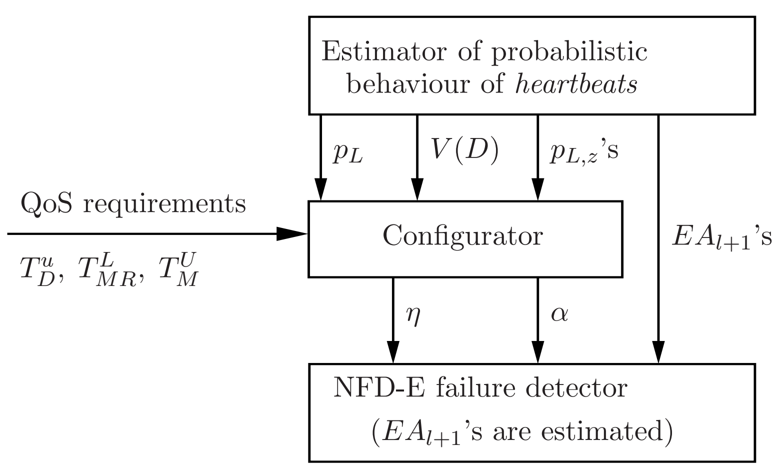

Figure 5: Configuration of NFD-E by NSM-NFD-E.

6 Simulation with Real Data

This section presents: the configuration of the simulation with collected data between two real networks; the analysis of detection time, mistake duration and mistake recurrence times; and an overall discussion of the results.

6.1 Configuration of the Simulation

The simulation has used collected data from the Brazilian link between the `prof.dct.ufms.br` computer at the Departament of Computing and Statistics (DCT) of UFMS (Federal University of Mato Grosso do Sul) and the `belem.ic.unicamp.br` computer of Institute of Computing at UNICAMP (University of Campinas). The gathering was performed in the following days: November 29th and 30th and December 1st, 3rd, 4th, 5th, 7th to 13th in 2005. The packets was sent unidirectionaly from DCT to IC by using the version 0.70 of `rude (http://rude.sourceforge.net/rude)` tool to send packets and `crude` tool to collect them, with the computer `prof` sincronizing its clock with the computer `belem` by `rdate` at each 1 minute. The period used in each day was from 7:01 to 24:00.

Table 1: An overview of collected data.

Day	p_L	h	$E(D)$	V
November 29th	0.101422	1481	0.557797	0.088631
November 30th	0.124808	208	0.561340	0.140205
December 1st	0.109601	250	0.573751	0.100157
December 3rd	0.001832	3	0.552977	0.102306
December 4th	0.007164	12	0.554129	0.109123
December 5th	0.125055	25	0.562738	0.092607
December 7th	0.073340	6	0.550084	0.096384
December 8th	0.030864	468	0.550982	0.111969
December 9th	0.000720	1	0.555134	0.114588
December 10th	0.001995	107	0.563183	0.123667
December 11th	0.006510	115	0.559827	0.095989
December 12th	0.002584	2	0.553556	0.106856
December 13th	0.002241	31	0.553351	0.104977

On top of each graphic the following information is showed: p_L, h, E, V, $min(T_{MR})$ or $min(T_M)$, and $max(T_{MR})$ or $max(T_M)$. The simulation used the NFD-E algorithm with the following delay estimators: a) the average of all message delays perceived until the estimation (MEAN), b) the average of the last n delays perceived until the estimation (WINMEANn), with $n \in \{4, 8, 16, 2, 64, 128, 256\}$, and c) the last delay perceived until the estimation (LAST).

The used T_D^U values varied from 1.6 to 4 s, with steps of 0.1 s. The Chen et al work, varied T_D^U values from 1 to 3.5 s. The initial value of 1.6 was used due to $\eta = 1$, and the message delay was around 0.5 seconds.

The collected data in the WAN showed a high mean in message delays, about 0.5 seconds, a high variance, about 0.09 seconds (which provides a standard deviation about 0.3 seconds), and the occurrence of a wide range of message loss burst lengths, from 1 to 1481. Table 1 shows the characteristics of each observed day.

In the following sections, Chen-NFD-E refers to Chen et al configurator for NFD-E and NSM-NFD-E refers to the NSM-NFD-E configurator defined in the end of Section 5.

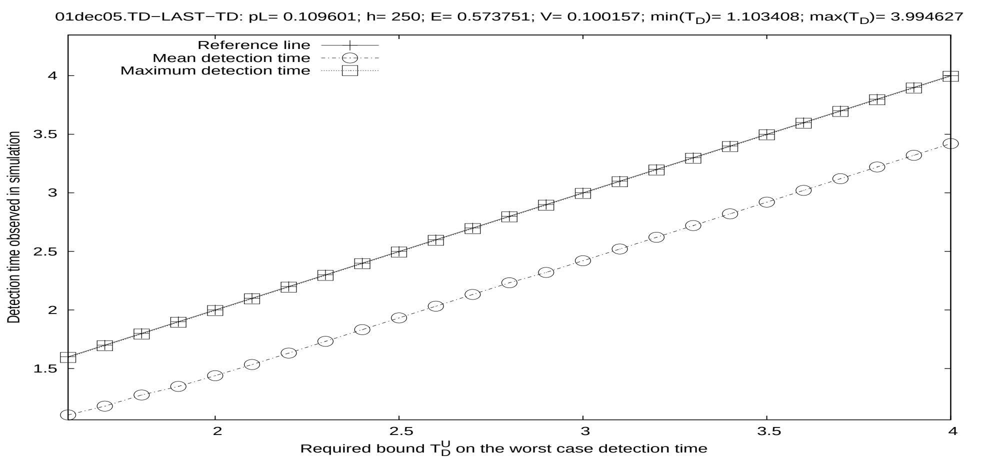

Figure 6: Detection time satisfied by NFD-E when using LAST.

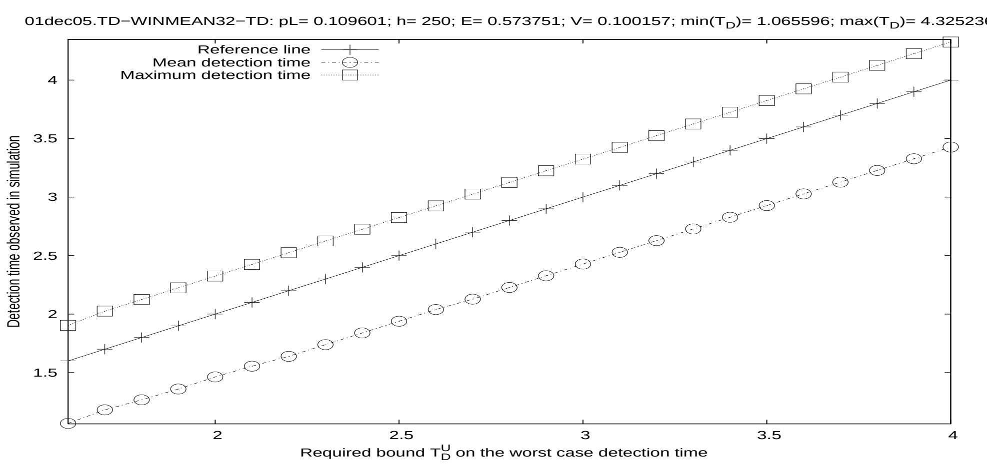

Figure 7: Detection time is not satisfied by NFD-E when using WINMEAN32.

6.2 Analysis of Detection Time

The detection time was verified from the calculation of the greatest detection time observed in 200 runs with arbitrary crash times for each day period and for each value of detection time. The data show that the use of a window in the message delay estimation (WINMEANn) do not lead the NFD-E failure detector to guarantee the detection time restrictions. The estimators WINMEANn always lead NFD-E to cross the upper bound in detection time. In fact, only the estimators MEAN and LAST made NFD-E to satisfy the detection time, practically in an indistinguishable way.

In Figures 6 and 7 at next, the maximum (obtained) detection time, mean (which only considers the values different from zero) and expected maximum (line of reference), of NFD-E, for LAST and WINMEAN32, respectively, are shown.

Figure 6 clearly shows that the detection time is satisfied by NFD-E when using LAST. Because the detection time is satisfied in a very similar way by NFD-E when using MEAN, its figure is not shown. Figure 7 shows that the detection time is not satisfied by NFD-E when using WINMEAN32. The other estimators WINMEANn also made NFD-E to perform similarly to that with WINMEAN32.

6.3 Analysis of Mistake Recurrence Time

The mistake recurrence time was verified from the average of mistake recurrence intervals observed in a day period for each value of detection time. In this section and in the next one, the figures use the term CHEN-analytic to refer to the analytical value generated by the Chen-NFD-E configurator, and the term NSM-analytic to refer to the analytical value generated by the NSM-NFD-E configurator.

The data show that the chosen delay estimator impacts little to NFD-E satisfies the mistake recurrence time requested (represented by the analytical values in the figures at next). The NSM-NFD-E configurator behaves similarly to the Chen-NFD-E configurator and satisfies the T_{MR}^L requirement in all cases. However, the Chen-NFD-E configurator fails in December 04th, as Figure 8 illustrates. This figure shows the behaviour of the NFD-E when using the WINMEAN32 estimator but this behaviour is similar to the other considered estimators. In the case of Figure 8, $p_L = 0.007164$, but $h = 12$, which explains the fact the Chen-NFD-E configurator fails, because it considers only the low p_L. The NSM-NFD-E configurator considers both low p_L and the length of message loss bursts ($h = 12$), and so it does not overestimate the T_{MR} value.

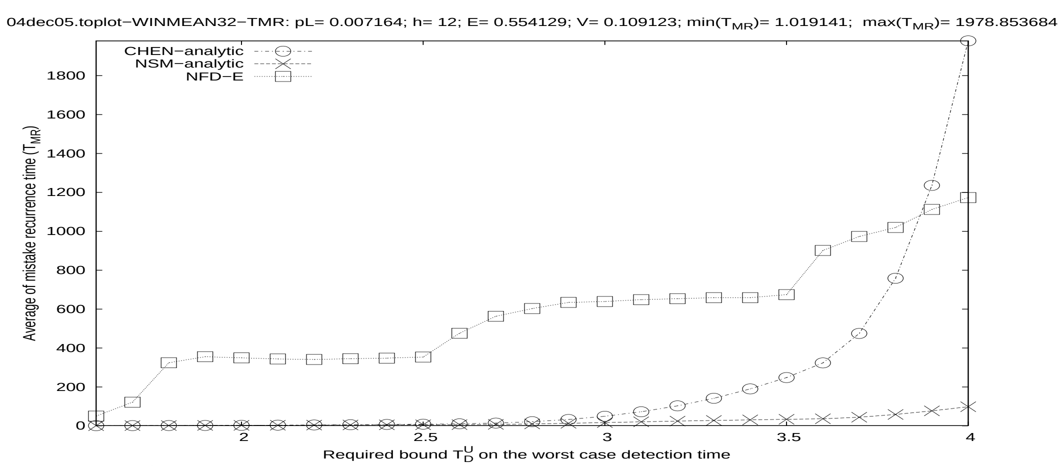

Figure 8: T_{MR}^L is not satisfied by NFD-E when using Chen-NFD-E configurator.

The WINMEANn estimators, with n not so big (from 4 to 32), lead the NFD-E to behave similarly and generate graphics in a well defined step form like in Figure 9. However, WINMEANn estimators with big window (n from 64), for example 256 in Figure 10, and the LAST and MEAN estimators lead the NFD-E to behave similarly and generate graphics smoother and the NFD-E behaviour is closer to the analytical values.

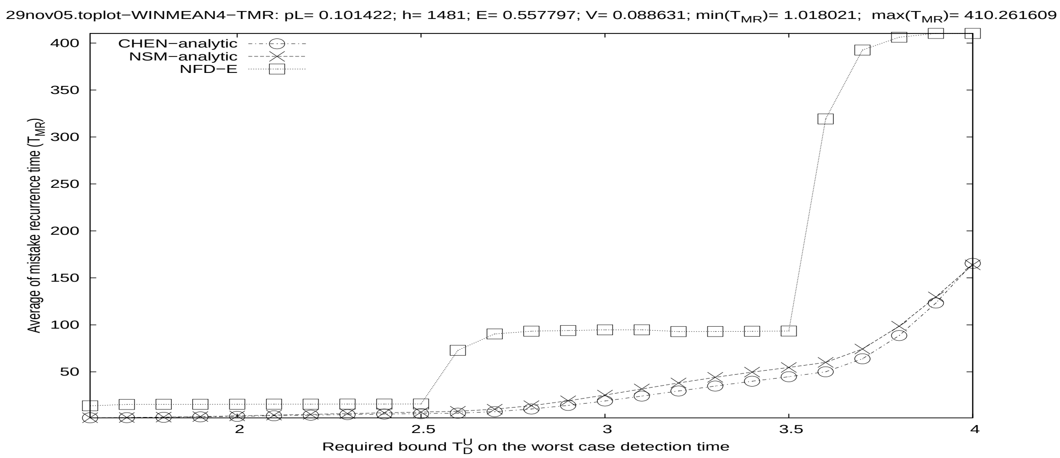

Figure 9: Behaviour of NFD-E in a step form with WINMEAN4.

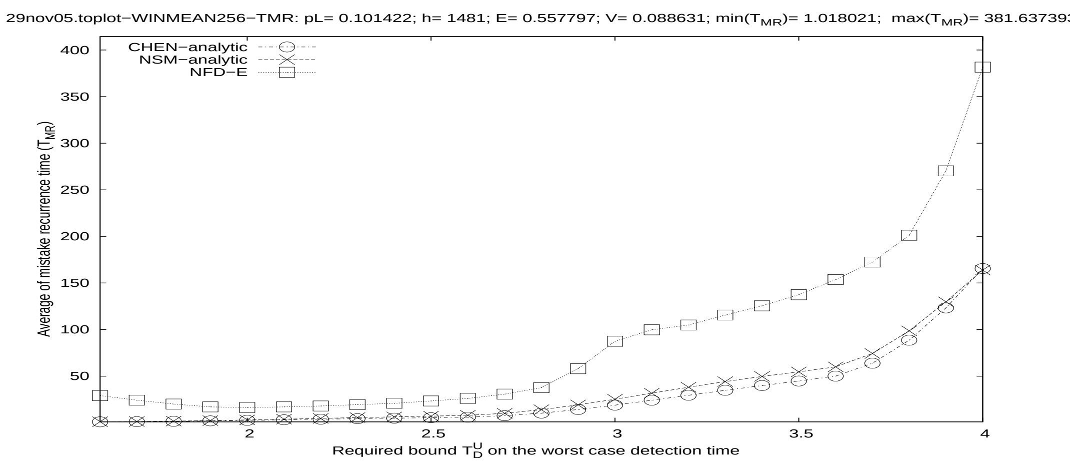

Figure 10: Behaviour of NFD-E is closer to analytical values with WINMEAN256.

6.4 Analysis of Mistake Duration

The mistake duration was verified from the average of mistake duration intervals observed in a day period for each value of detection time. The data show that the chosen delay estimator impacts much to NFD-E satisfies the mistake duration requested (represented by the analytical values in the figures at next). Moreover, the number of intervals of mistake duration related to the maximum loss burst length (h) and/or variance also impacts very much. Both the NSM-NFD-E configurator and Chen-NFD-E configurator do not lead the NFD-E to satisfy the T_M^U requirement in several cases. However, the NSM-NFD-E configurator is clearly better than Chen-NFD-E configurator when there are long loss bursts.

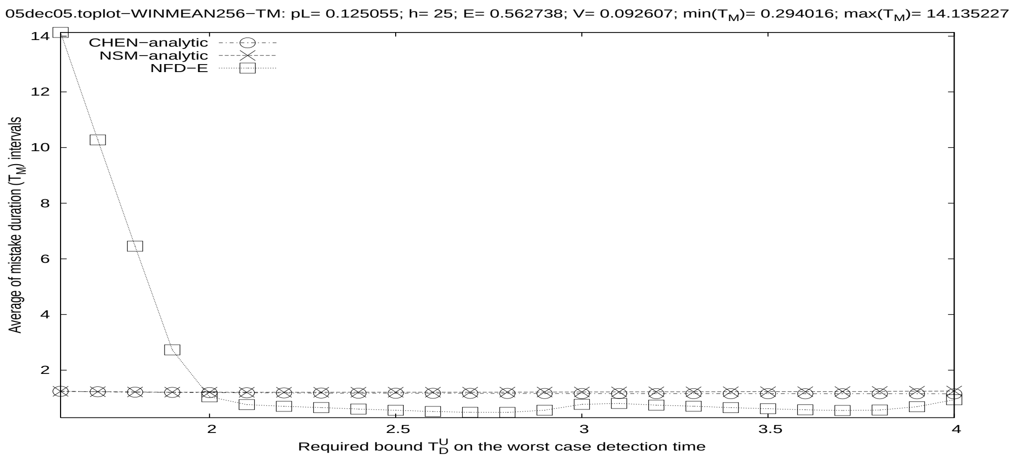

Figure 11: T_M^U is not satisfied for low values of T_D^U.

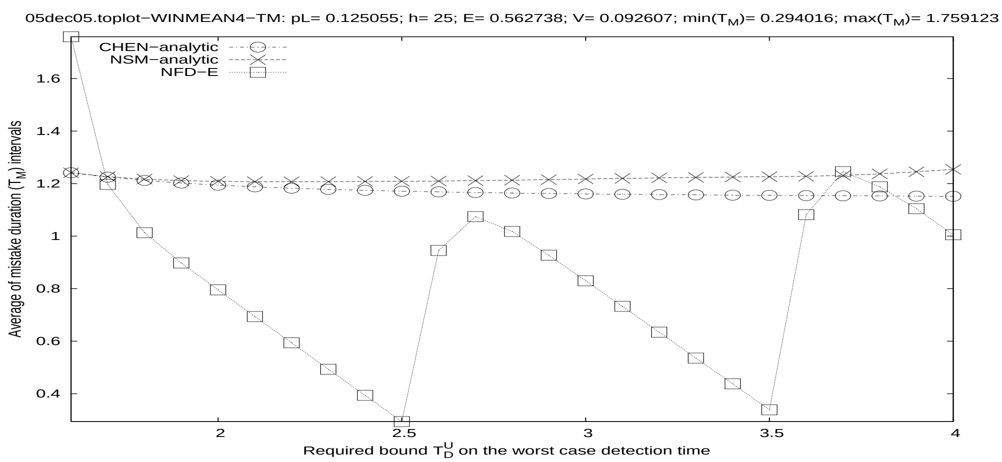

Figure 12: T_M^U is better satisfied for low values of T_D^U.

The Figure 11, with WINMEAN256, represents the behaviour of NFD-E with WINMEAN128, WINMEAN64, LAST and MEAN. NFD-E with the WINMEAN32, WINMEAN16 and WINMEAN8 estimators approximate the behaviour with WINMEAN4 in Figure 12, which offers the best estimate to the NFD-E behaviour for low values of T_D^U. In this case, in which $p_L = 0.125$ is much high and the maximum loss bust length is not so long ($h = 25$), both NSM-NFD-E and Chen-NFD-E configurators perform similarly. Because this pattern of WINMEAN4 (as with WINMEAN8, WINMEAN16 and WINMEAN32) be better than WINMEAN256 repeats in the other days, the following figures show only results with WINMEAN4. In fact, Figure 12 presents a behaviour much similar to that observed in December 7th (not shown), in which the T_M^U analytical values of both NSM-NFD-E and Chen-NFD-E configurators could catch better the behaviour of NFD-E.

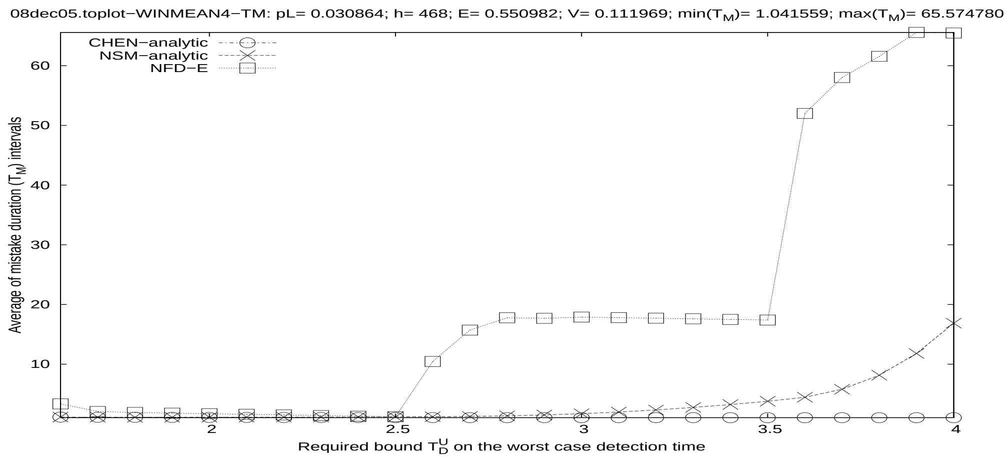

Figure 13: T_M^U is better satisfied for low values of T_D^U.

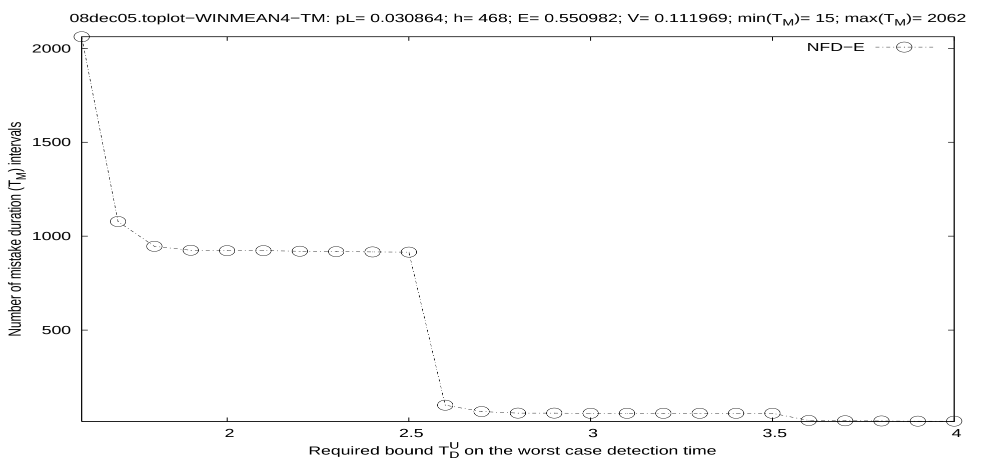

Figure 14: T_M^U is better satisfied for low values of T_D^U.

NFD-E with WINMEAN32, WINMEAN16 and WINMEAN8 estimators approximate the behaviour with WINMEAN4 in Figure 13, which offers the best estimate to NFD-E behaviour for low values of T_D^U. For T_D^U from 2.5 and for T_D^U from 3.5, WINMEAN4 leads NFD-E to behaves very much poorly, but this occurs due a few number of mistake duration intervals (see Figure 14) related to the maximum loss burst length ($h = 468$). In this case, in which $p_L = 0.03$ is high, but $h = 468$ is so long, the NSM-NFD-E configurator provides analytical values better than the Chen-NFD-E configurator, from $T_D^U = 2.5$.

Figure 13 presents a behaviour much similar to that observed in November 29th and 30th and December 1st (not shown), in which p_L's are much high (from 0.03) and h's are also much high (from 208), and the NSM-NFD-E configurator provides analytical values with increasing trend, while the Chen-NFD-E configurator provides analytical values practically constant.

NFD-E with WINMEAN32, WINMEAN16 and WINMEAN8 estimators approximate the behavior with WINMEAN4 in Figure 15, which offers the best estimate to NFD-E behaviour. The loss probability $p_L = 0.006510$ is very low and the variance is high (0.095989), but $h = 115$ is high.

For $T_D^U = 1.6$ the variance impacts so that T_M^U is not satisfied, even though the high number of mistake duration intervals (see Figure 16). For $T_D^U = 1.7$ there is a sharp reduction in the number of mistake duration intervals, but it still is above h (115), which has little impact in T_M. From $T_D^U = 1.8$, the number of mistake duration intervals is lower than h (115), which impacts very much in T_M and leads NFD-E to behaves very much poorly from $T_D^U = 2.6$. However, the NSM-NFD-E configurator improves the prediction from $T_D^U = 3$, while the Chen-NFD-E configurator stays constant. The NSM-NFD-E configurator is better than Chen-NFD-E configurator because it considers p_L, which is considered by Chen-NFD-E, plus the length of loss burst ($h = 115$), which Chen-NFD-E does not consider.

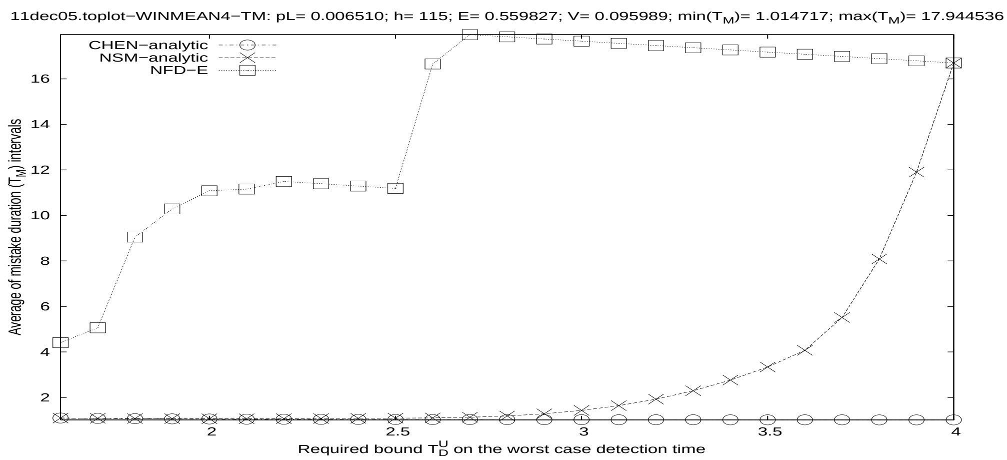

Figure 15: T_M is impacted by high variance and high h.

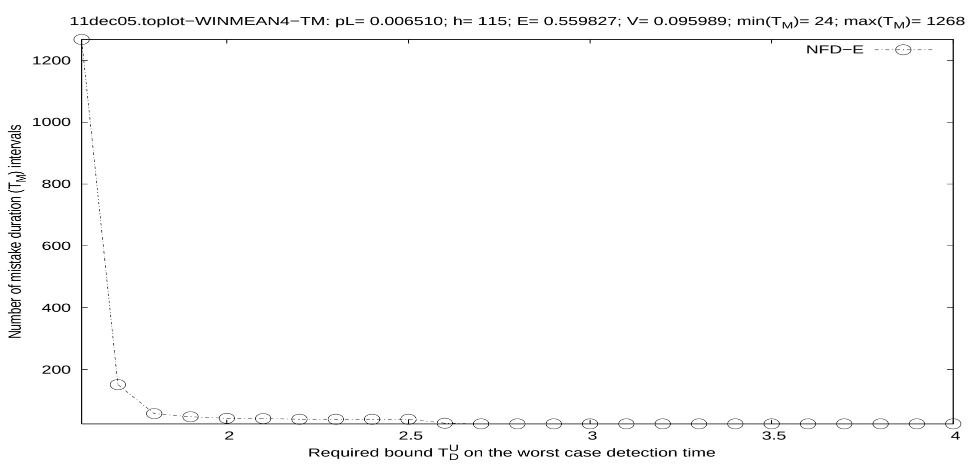

Figure 16: T_M is impacted by high variance and high h.

Figure 15 presents a behaviour much similar to that observed in December 4th, 10th and 13th (not shown), in which p_L's are much low (below 0.008) and h's are medium (12) and high (from 31), and the NSM-NFD-E configurator also provides analytical values with increasing trend, while the Chen-NFD-E configurator provides analytical values practically constant.

NFD-E with WINMEAN32, WINMEAN16 and WINMEAN8 estimators approximate the behaviour with WINMEAN4 in Figure 17, which offers the best estimate to NFD-E behaviour for low values of T_D^U. For T_D^U from 2.5, WINMEAN4 leads NFD-E to behave very much poorly, but this occurs due to a few number of mistake duration intervals (see Figure 18) which is not enough to provide a low average of T_M^U due to the high variance (0.114588), even though $h = 1$ and the loss probability be very low (0.000720). In this case, both NSM-NFD-E and Chen-NFD-E configurators provide similar analytical values, which is natural because both $p_L = 0.000720$ and $h = 1$ are very low.

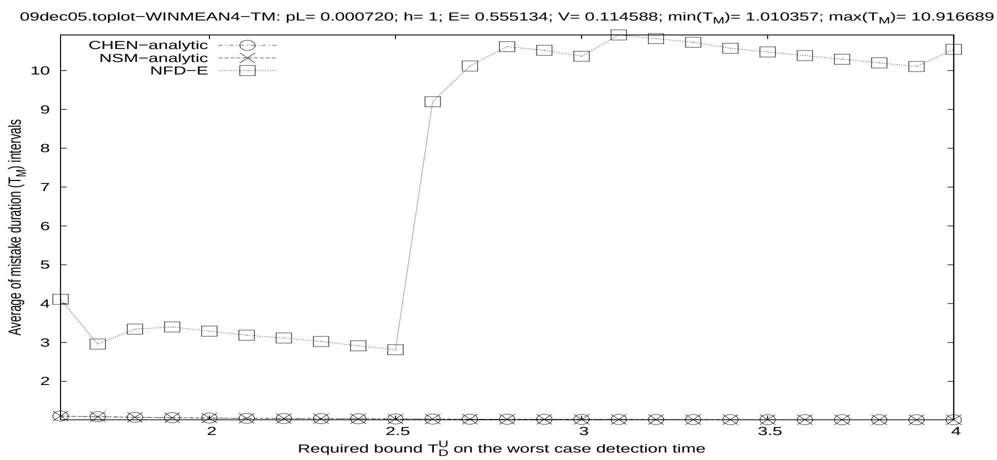

Figure 17: T_M^U is better satisfied for low values of T_D^U.

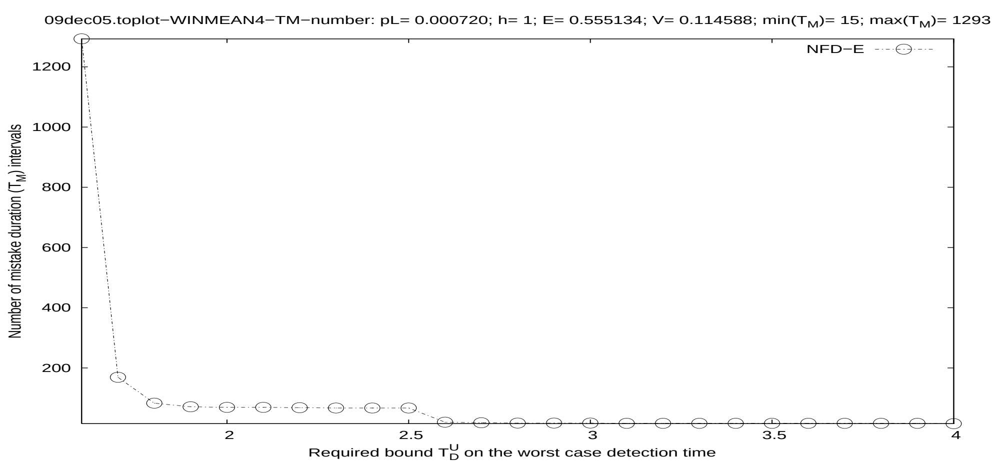

Figure 18: T_M^U is better satisfied for low values of T_D^U.

Figure 17 presents a behaviour much similar to that observed in December 3rd and 12th (not shown), in which p_L's are much low (below 0.003) and h's are much low (below 4), and the NSM-NFD-E configurator provides analytical values similar to Chen-NFD-E configurator analytical values (practically constant).

6.5 Overall Discussion of the Simulation

The data collected from a Brazilian link between two networks show a high message delay, a high variance, and long message loss bursts. The simulation with those data show the message delay estimator impacts so much in the quality of service of the NFD-E failure detector.

In the case of detection time, only the MEAN and LAST estimators lead the NFD-E algorithm to satisfy the required detection time, while all others estimators WINMEANn do not work well. MEAN uses message delay values more stable than WINMEANn, which is an advantage for detection time because temporary periods with delays above the average use a mean in the case of MEAN and a higher value in the case of WINMEANn. Due to LAST uses only the most recent delay as the estimate, the detection time also is timely. Because this behaviour of NFD-E detection time has occurred in the all observed days under variable conditions of message delay (from 0.550084 to 0.573751s), variance (from 0.088631 to 0.140205s^2) and loss burst length (from 1 to 1481), it indicates that NFD-E with LAST or MEAN estimators always tends to satisfy the detection time.

In the case of mistake recurrence times, MEAN and LAST are also good estimators and lead NFD-E to behave in a way closer to the analytical values of Chen-NFD-E and NSM-NFD-E configurators. The WINMEANn estimators with big windows (n from 64) also lead NFD-E to behave like with MEAN and LAST. WINMEANn estimators with little windows (n below 32), like WINMEAN4, are better than MEAN and LAST. The NSM-NFD-E configurator generated analytical values which did not fail in any day (even with wide variability in delay mean, variance, loss probability and loss burst lengths and with any of considered estimators), but the Chen-NFD-E configurator failed in December 04th (for all considered estimators). This failure occurs because Chen-NFD-E configurator does not consider the loss burst length in its model.

In the case of mistake duration, unlike the detection time and the mistake recurrence time, MEAN and LAST are not good estimators, mainly to low values of T_D^U and they lead NFD-E to behave far away from the analytical values of Chen-NFD-E and NSM-NFD-E configurators. For the mistake duration intervals, WINMEAN4 was the best estimator. Both configurators fail to satisfy the mistake duration. Both configurators generate similar values when the loss probability is high or low, and the loss burst lengths are not so long. The long loss bursts or high variance are very harmful for NFD-E algorithm and leads it to not satisfy the T_M analytical values of the configurators. However, the NSM-NFD-E configurator is better than Chen-NFD-E configurator when the loss burst lengths are long. With high variance, even when there are low p_L and low h, both configurators provide bad and similar analytical values which are very far from the NFD-E behaviour.

In short, the NFD-E configurator effectiveness depends so much of the delay estimators, mainly to satisfy detection time and mistake duration requirements. The estimator choice by the NFD-E configurator user would privilege one of these requirements. However, in practice, to satisfy T_D^U with NFD-E we must satisfy $T_D^u + E(D)$. So, the NFD-E configurator user can periodically execute the configurator to obtain a more precise analytical values with the same η by using the most recent estimates (mean, loss probability and loss burst length probability). The main causes of no good results to both configurators about mistake duration are high variance (see Figure 17) and long bursts associated with a number of mistake duration intervals below h (see Figures 13, 14, 15 and 16). Mistake recurrence time is much easier to be satisfied, and the NFD-E configurator can predict well the behaviour of the NFD-E, practically independent of the chosen delay estimator. The results confirm those ones obtained by Nunes and Jansch-Pôrto [19] and Falai and Bondavalli [10] that the choice of the message delay estimator has much impact on the QoS of failure detectors and that is difficult to obtain estimators/predictors which are good to lead the failure detectors to satisfy T_D^U, T_M^U and T_{MR}^L at the same time.

Because detection time is satisfied only with MEAN and LAST estimators, the Table 2 shows an overview of the simulation results for mistake duration (T_M) and mistake recurrence times (T_{MR}).

Table 2: An overview of accuracy results.

	Chen-NFD-E and NFD-E behavior	NSM-NFD-E and NFD-E behavior	Message delay estimator
T_{MR}^L	Satisfied, except in December 04th (see Figure 8).	Always satisfied and NSM-NFD-E is better than Chen-NFD-E when p_L is very low (e.g. 0.007164) and h is not short (e.g. 12) (see Figure 8).	MEAN, LAST, WINMEANn ($n \in [64, 256]$) are good and similar (see Figure 10). WINMEANn ($n \in [4, 32]$) are better and WINMEAN4 is the best (see Figure 9).
T_M^U	Not satisfied, but Chen-NFD-E is similar to NSM-NFD-E when p_L is very low (e.g. 0.000720) and h is 1 (no loss burst) (see Figure 17).	Not satisfied, but NSM-NFD-E is better than Chen-NFD-E when long loss bursts occur (see Figures 12, 13 e 15).	MEAN, LAST and WINMEANn ($n \in [64, 256]$) are much bad estimators (see Figure 11). WINMEANn ($n \in [4, 32]$) are better and WINMEAN4 is the best (see Figure 12).

7 Conclusions

This paper proposed a new NFD-U/NFD-E configurator for Chen et al NFD-U/NFD-E algorithm [7] for unsynchronized clocks. The evaluation was done with data collected from a link between two universities located in different Brazilian States. The proposed configurators work better than the original one when long message loss bursts occur and is similar to Chen et al configurator when there are no loss bursts. The better performance comes from the additional knowledge of distribution of the message loss burst lengths, which is used to build a configurator based on a Markov model. The simulation also confirms previous results about the impact of the message delay estimator on failure detectors QoS.

8 Acknowledgments

The authors would like to thank Elias Procopio Duarte Junior, Luiz Eduardo Buzato, Islene Calciolari Garcia and Joni da Silva Fraga by their useful comments on the work which generated this paper.

9 References

[1] J. Andrén, M. Hilding, and D. Veitch. Understanding End-to-End Internet Traffic Dynamics. *Proceedings of IEEE 1998 Global Communications Conference (GLOBE-COM 98)*, vol. 2, pp. 1118–1122. Sidney, Australia, November 1998.

[2] F. Anjum and L. Tassiulas. Comparative Study of Various TCP Versions Over a Wireless Link With Correlated Losses. *IEEE/ACM Transactions on Networking*, 11(3), pp. 370–383, June 2003.

[3] Ö. Babaoglu, R. Davoli, and A. Montresor. Group Communication in Partitionable Systems: Specification and Algorithms. *IEEE Transactions on Software Engineering*, 27(4):308–336, April 2001.

[4] M. Bertier, O. Marin, and P. Sens. Implementation and performance evaluation of an adaptable failure detector. *Proceedings of the 2002 International Conference on Dependable Systems and Networks (DSN'02)*, pp. 354–363. June 2002.

[5] M. Bertier, O. Marin, and P. Sens. Performance Analysis of a Hierarchical Failure Detector. *Proceedings of the 2003 International Conference on Dependable Systems and Networks (DSN'03)*, pp. 635–644. June 2003.

[6] T. D. Chandra and S. Toueg. Unreliable Failure Detectors for Reliable Distributed Systems. *Journal of the Association for Computing Machinery*, 43(2):225–267, March 1996.

[7] W. Chen, S. Toueg, and M. K. Aguilera. On the Quality of Service of Failure Detectors. *IEEE Transactions on Computers*, 51(5):561–580, May 2002.

[8] G. V. Chockler, I. Keidar, and R. Vitenberg. Group Communication Specifications: A Comprehensive Survey. *ACM Computing Surveys*, 33(4):427–469, December 2001.

[9] C. Dwork, N. Lynch, and L. Stockmeyer. Consensus in the Presence of Partial Synchrony. *Journal of the Association for Computing Machinery*, 35(2):288–323, April 1988.

[10] L. Falai and A. Bondavalli. Experimental Evaluation of the QoS of Failure Detectors on Wide Area Network. *Proceedings of the 2005 International Conference on Dependable Systems and Networks (DSN'05)*, pp. 624–633. June–July 2005.

[11] M. J. Fischer, N. A. Lynch, and M. S. Paterson. Impossibility of Distributed Consensus with One Faulty Process. *Journal of the Association for Computing Machinery*, 32(2):374–382, April 1985.

[12] R. Guerraoui and M. Raynal. The Information Structure of Indulgent Consensus. *IEEE Transactions on Computers*, 53(4):453–466, April 2004.

[13] N. Hayashibara, X. Défago, R. Yared, and T. Katayama. The φ Accrual Failure Detector. *Proceedings of the 23rd International Symposium on Reliable Distributed Systems (SRDS'04)*, pp. 66–78. October 2004.

[14] A. Konrad, B. Y. Zhao, A. D. Joseph, and R. Ludwig. A Markov-Channel Model Algorithm for Wireless Networks. *Wireless Networks*, 9, pp. 189–199, 2003.

[15] M. Larrea, A. Fernandéz, and S. Arévalo. On the Implementation of Unreliable Failure Detectors in Partially Synchronous Systems. *IEEE Transactions on Computers*, 53(7):815–828, July 2004.

[16] H. Lee and P. K. Varshney. Gap-based Modeling of Packet Losses over the Internet. *Proceedings of 10th IEEE International Symposium on Modeling, Analysis and Simulation of Computer and Telecommunications Systems (MASCOTS 2002)*, pp. 507–510. October 2002.

[17] A. Mostefaoui, E. Mourgaya, and M. Raynal. Asynchronous Implementation of Failure Detectors. *Proceedings of the 2003 International Conference on Dependable Systems and Networks (DSN'03)*, pp. 351–360. June 2003.

[18] A. Mostefaoui, D. Powell, and M. Raynal. A Hybrid Approach for Building Eventually Accurate Failure Detectors. *Proceedings of the 10th IEEE Pacific Rim International Symposium on Dependable Computing (PRDC'04)*, pp. 57–65. March 2004.

[19] R. C. Nunes, I. Jansch-Pôrto. QoS of Timeout-based Self-Tuned Failure Detectors: the Effects of the Communication Delay Predictor and the Safety margin. *Proceedings of the 2004 International Conference on Dependable Systems and Networks (DSN'04)*, pp. 753–761. June 2004.

[20] V. Paxson. End-to-end Internet Packet Dynamics. *IEEE Transactions on Networking*, 7(3):277–292, June 1999.

[21] H. Sanneck. *Packet Loss Recovery and Control for Voice Transmission over the Internet*. PhD thesis, Technischen Universität Berlin, Berlin, October 2000.

[22] P. Sarolahti, M. Kojo, and K. Raatikainen. F-RTO: An Enhanced Recovery Algorithm for TCP Retransmission Timeouts. *ACM SIGCOMM Computer Communications Review*, 33(2):51–63, April 2003.

[23] A. Schiper. Failure Detection vs Group Membership in Fault-Tolerant Distributed Systems: Hidden Trade-Offs. *Proceedings of Process Algebra and Probabilistic Methods. Performance Modeling and Verification: Second Joint International Workshop PAPM-PROBMIV 2002*, LNCS 2399, pp. 1–15, July 2002.

[24] I. Sotoma and E. R. M. Madeira. A Markov Model for Quality of Service of Failure Detectors in the Presence of Loss Bursts. *Proceedings of 18th International Conference on Advanced Information Networking and Applications (AINA 2004)*, vol. 2, pp. 62–67. Fukuoka, Japan, March 2004.

[25] I. Sotoma and E. R. M. Madeira. A Markov Model for Providing Quality of Service for Failure Detectors under Message Loss Bursts. *Technical Report IC-06-13*, Institute of Computing, Campinas, September 2006 (Available at http://www.ic.unicamp.br/reltec-ftp/2006/Titles.html).

[26] P. Urbán, I. Shnayderman, and A. Schiper. Comparison of Failure Detectors and Group Membership: Performance Study of Two Atomic Broadcast Algorithms. *Proceedings of the 2003 International Conference on Dependable Systems and Networks (DSN'03)*, pp. 645–654. June 2003.

[27] M. Yajnik, S. B. Moon, J. Kurose, and D. Towsley. Measurement and modeling of the temporal dependence in packet loss. *Proceedings of Eighteenth Annual Joint Conference of the IEEE Computer and Communications Societies (INFOCOM'99)*, vol. 1, pp. 345–352. March 1999.

[28] L. Zhang. Why TCP Timers Don't Work Well. *ACM SIGCOMM Computer Communications Review*, 16(3):397–405, August 1986.

[29] Y. Zhang. *Characterizing End-to-end Internet Performance*. PhD thesis, Cornell University, Cornell, August 2001.