Analytic models for tree communication protocols

Philippe Flajolet, Philippe Jacquet

INRIA

UNITE DE RECHERCHE
INRIA-ROCQUENCOURT

Institut National
de Recherche
en Informatique
et en Automatique

Domaine de Voluceau
Rocquencourt
B.P.105
78153 Le Chesnay Cedex
France
Tél.(1) 39 63 55 11

Rapports de Recherche

N° 648

ANALYTIC MODELS FOR TREE COMMUNICATION PROTOCOLS

Philippe FLAJOLET
Philippe JACQUET

Mars 1987

ANALYTIC MODELS FOR TREE COMMUNICATION PROTOCOLS

Philippe Flajolet and *Philippe Jacquet*

Abstract:† *The tree protocol for local area networks, together with a number of its variants, can be exactly analysed under a Poisson arrival model. This note surveys some of the evaluations that have been obtained for characteristic parameters including delay, session length or probability of immediate access to the channel. The mathematical techniques involved are: functional equations and Mellin transforms.*

MODELES ANALYTIQUES POUR LES PROTOCOLES DE COMMUNICATION EN ARBRE

Philippe Flajolet et *Philippe Jacquet*

Résumé: Le protocole en arbre pour réseaux locaux, ainsi que nombre de ses variantes, peut être exactement analysé sous un modèle d'arrivées Poissoniennes. Cette note présente une synthèse d'évaluations obtenues pour les paramètres caractéristiques tels: le délai, la longueur de session ou les probabilités d'accès immédiat au canal. Les techniques mathématiques en jeu comprennent les équations fonctionnelles et la transformation de Mellin.

---

† Invited lecture at the NATO Advanced Study Institute on *"Flow Control of Congested Networks"*, Capri (Oct. 1986); Proceedings published by Springer, New-York (1987)

ANALYTIC MODELS FOR TREE COMMUNICATION PROTOCOLS

Philippe Flajolet and *Philippe Jacquet*
INRIA, Rocquencourt
78150 Le Chesnay (France)

Abstract: The tree protocol for local area networks, together with a number of its variants, can be exactly analysed under a Poisson arrival model. This note surveys some of the evaluations that have been obtained for characteristic parameters including delay, session length or probability of immediate access to the channel. The mathematical techniques involved are: functional equations and Mellin transforms.

1. Introduction

Protocols for regulating access to a channel shared by several stations were first designed in the sixties, and started with the ALOHA concept: Each station transmits as soon as it has a message to send; the message is broadcast, and picked up by its intended receiver unless two (or more) stations collide; in that case, every station detects the collision, and senders schedule a later retransmission. The rule for retransmitting is precisely the *communication protocol.*

Stations' "feedback" from the channel is thus limited to a ternary information: ACK (acknowledgement, i.e. successful transmission); LACK (lack of transmission, i.e. silence); NACK (no ACK, i.e. collision). Since stations are not distinguishable from each other in general, a key idea to resolve collisions is to let a *random* component enter their retransmission policy. If the protocol is suitably designed, one may hope for the best, namely expect the channel to succesfully transmit messages with reasonable delays as long as the traffic load is not too high.

In this paper, we consider the case of a *slotted time channel* where transmissions start at discrete instants $0, 1, 2, \dots$ and messages are calibrated so that their duration does not exceed one slot. Basically, the ALOHA protocol is the following simple rule:

A. In case of a collision, wait a random amount of time (i.e. slots) uniformly distributed over the interval $[1..\delta_0]$ before attempting a retransmission. Parameter δ_0 is a design parameter whose value is fixed and common to all stations, its value being chosen based on the network configuration.

It was soon realised (Fayolle et al.; Kleinrock et al.) that the ALOHA protocol is unstable: If messages arrive according to a Poisson process with intensity $\lambda > 0$, then with probability 1, the "backlog" (messages awaiting retransmission because of previous collisions) tends to infinity. Intuitively, the protocol maintains a virtual time window of a fixed size that, sooner or later, is doomed to become saturated.

The next idea, which gave rise to the Ethernet protocol, was to use a "sliding" parameter δ whose value changes dynamically with stations and time, and whose control is meant to have the protocol adjust to traffic variations. The simple idea is for a station to retransmit randomly in the interval $[1..\delta]$ where:

E. Initially, upon a message arrival, δ is 1. After each collision experienced by its message, the station doubles it own value of δ.

It took some time and effort (Aldous) to realise that Ethernet is itself unstable: in practice, this may mean fairly suboptimal channel utilisation, delay inefficiencies or poor response to bursts of traffic. Meanwhile, the *tree protocol* was invented circa 1977 by Capetanakis, based on the following elegant idea:

T. If a group G of stations collide ($|G| \geq 2$), that group is split by coin flippings into two subgroups G_0 and G_1. The stations in G_0 *first* recursively resolve their collisions. *Then* the group G_1 resolve their collisions independently.

The interest of this protocol is to use a dichotomy to separate colliders, and an execution is simply described by a tree. It is not immediately clear however that it can be implemented without having the stations communicate some extraneous informations about their coin flippings. A decentralised formulation (providing a practical implementation) of that protocol was arrived at independently by Tsybakov et al. and there, each station manages a *stack.*

Partial stability regions were characterised in the initial papers, and we now know that, under the Poisson model, the protocol is stable until an arrival rate of about 35%. Actually, going to details, there are two ways of implementing the protocol:

- *Free Access:* The immediate access rule (in the style of ALOHA) is enforced. Thus the resolution of collisions by a group may involve newly arrived messages; the system operates in a somewhat last-in first-out fashion.
- *Blocked Access:* There collisions are resolved in sessions. Newly arrived stations wait until a collision resolution "session" is over (in case any is taking place) before they are allowed to enter the competition, and start a new session.

The purpose of this paper is to present the analytic methods involved in the evaluations, putting into perspective some of our own results [FFHJ 1985], [FFH 1986], [MF1985], [Jacquet 1987] and [JR 1986]. It is quite interesting that there is a fairly rich mathematical structure behind the tree protocol, and many relevant parameters can be exactly analysed. We shall try to illustrate the mathematical methods at stake here.

Note: The reader can refer to the survey paper by Massey [Massey 1981] and to a special issue of the IEEE Transactions [Massey 1985] for detailed references on the subject that are not duplicated here.

2. Blocked Access: A Basic Analysis

Strangely enough, the basic tree protocol with blocked access had been analysed *before* it was invented, by Knuth (1973). The reason is the generality of the recursive splitting process based on random choices that turns out to be the exact model for the trie data structure, and for a variety of searching methods in computer science. Below is a revised presentation of Knuth's result.

Let L_N be the random variable (RV) denoting the time taken to resolve N collisions using the tree method, assuming no further arrivals. If the group of size N is split into subgroups of size K and $N-K$ $(N \geq 2)$ then:

$$L_N = 1 + L_K + L_{N-K} \qquad \text{with} \qquad L_0 = L_1 = 1. \tag{1}$$

When a fair coin is used, the "splitting" probabilities that the RV K has value k is:

$$\pi_{N,k} = \frac{1}{2^N}\binom{N}{k} \tag{2}$$

so that, taking expectations of (1):

$$l_N \equiv E[L_N] = 1 + \sum_{k=0}^{N} \pi_{N,k}(l_k + l_{N-k}), \qquad (N \geq 2) \tag{3a}$$

a relation that permits to compute inductively each of the l_N. The form of (3a) suggests using exponential generating functions (egf's). If $l(z) = \sum_{N \geq 0} l_N \frac{z^N}{N!}$, then (3a) becomes:

$$l(z) = e^z - 2 - 2z + 2e^{z/2} l(\frac{z}{2}), \tag{3b}$$

a *difference equation*. From there two routes are possible:

- Direct solution: Set $\Lambda(z) = e^{-z} l(z)$ and determine the relation satisfied by it. It is of the form: $\Lambda(z) = 2\Lambda(\frac{z}{2}) +$ a known function. Thus the coefficients of Λ can be exactly recovered, whence by convolution with those of e^z, the coefficients of $l(z)$. This provides a finite sum for l_N, which involves exponential cancellations and does not yield easily to asymptotic analysis (though the so-called "Rice integrals" method from the calculus of finite differences can be used).
- Iterative solution: A functional equation like (3b) is of the form:

  $$\phi(z) = \alpha(z) + \beta(z)\phi(\gamma(z))$$

  with ϕ the unknown function, and it can be solved by iteration. If this is done here, and the solution is expanded, one finds a well-conditioned sum:

  $$l_N = 1 + 2\sum_{k=0}^{\infty} 2^k \left[1 - (1 - \frac{1}{2^k})^N - \frac{N}{2^k}(1 - \frac{1}{2^k})^{N-1}\right] \tag{4}$$

  an expansion that however reserves some surprises!

It is interesting *per se* to consider approximations of l_N, and it is not difficult to conjecture that $l_N \approx \frac{2N}{\log 2}$. A good physical reason for interest in this asymptotic problem is that, if $l_N \sim cN$, then the constant c is (asymptotically) an average service time. Thus the protocol should be stable for arrival rates λ such that $c\lambda < 1$, which suggests a maximum admissible throughput for the tree protocol of $\lambda = \log 2/2 = 0.34657$.

That conjecture is almost true, but not as simple as it looks. First, using in (4) the approximation $(1-a)^N \approx e^{-aN}$ –which is easy to justify– we find:

$$l_N = 2F(N) + O(\sqrt{N}) \qquad \text{with} \qquad F(x) = \sum_{k\geq 0} 2^k[1-(1+\frac{x}{2^k})e^{-x/2^k}]. \tag{5}$$

That sum is a so-called harmonic sum and the best way to treat it is to determine its Mellin transform defined as:

$$F^*(s) = \int_0^\infty F(x)x^{s-1}\,dx. \tag{6}$$

Here, we find:

$$F^*(s) = \frac{(s+1)\Gamma(s)}{1-2^{s+1}} \qquad \text{for } s \text{ in the strip } -2 < \Re(s) < -1. \tag{7}$$

From there, using the Mellin inversion formula, and computing the integral by residues, we get:

$$F(x) = \frac{1}{2i\pi}\int_{-3/2-i\infty}^{-3/2+i\infty} F^*(s)x^{-s}\,ds \sim \sum_{\Re(s)\geq -1} \mathrm{Res}[F^*(s)x^{-s}], \tag{8}$$

where the last sum has the character of an asymptotic expansion in non-increasing powers of x. However (There is the rub!), $F^*(s)$ has poles with a non-zero imaginary part, and since $x^{it} = e^{it\log x}$, these correspond to *periodic fluctuations*. Thus l_N is not a "smooth" function of N, though the periodic fluctuations are of limited importance since their amplitude is $< 10^{-5}$. We have arrived at:

THEOREM 1: *(i).* [Knuth 1973] *Quantity l_N satisfies asymptotically:*

$$l_N = \frac{2}{\log 2}N + NP(\log_2 N) + O(\sqrt{N})$$

where $P(.)$ is a periodic function with amplitude $< 10^{-5}$.
(ii). The supremum $\lambda_{\max}$ of arrival rates λ for which the tree protocol with blocked access is stable satisfies:

$$|\lambda_{\max} - \frac{\log 2}{2}| < 10^{-5}.$$

Therefore $\lambda_{\max} \approx 0.34657$.

3. Free Access: A Functional Equation...

It is natural to try and extend the previous analysis to the case of the free access version of the protocol. As we shall see, the functional equations that appear are of a different form, and they lead to some analytic difficulties. However, from the analysis, the maximum admissible throughput can be exactly determined.

Consider thus the free access tree protocol with a Poisson rate of arrival λ. The starting point is a recursive relation on RV's that extends Eq (1). Since arrivals keep coming in, the basic equation is:

$$L_N = 1 + L_{K+X} + L_{N-K+Y} \qquad \text{with} \qquad L_0 = L_1 = 1 \tag{9}$$

where X, Y are Poisson RV's. At this stage, also consider the possibility for the coin flippings to be biased, with probabilities p and $q = 1 - p$ for head and tail. Introduce the egf $l(z)$ of the expectations l_N and the modified egf $\Lambda(z) = e^{-z} l(z)$. These quantities now depend on λ and $l_N = l_N(\lambda)$ etc., thus the l_N for blocked access coincide with $l_N(0)$. It is not too difficult to see that $\Lambda(z)$ satisfies an equation of the form:

$$\Lambda(z) = \Lambda(\lambda + pz) + \Lambda(\lambda + qz) + \alpha(z) \tag{10}$$

where $\alpha(z)$ is a known function. That functional equation ceases to be "local": it relates the values of Λ around 0 to the values around λ, and this reflects the fact that, due to arrivals, each l_N is related by an infinite recursion to all the other l_j.

A first step is thus to consider functional equations of the general form (where ϕ is the unknown function):

$$\phi(z) = \alpha(z) + u\phi(\sigma_1(z)) + v\phi(\sigma_2(z)) \qquad \text{with} \quad \sigma_1(z) = \lambda + pz \text{ and } \sigma_2(z) = \lambda + qz. \tag{11a}$$

Solutions by iteration of (11a) involve a sum over the "iteration semigroup" H defined as the set of all compositions of σ_1 and σ_2 (observe that H is non commutative in general):

$$\phi(z) = \mathbf{S}[\alpha(.); z] = \sum_{\tau \in H} (u; v)^\tau \alpha(\tau(z)) \qquad \text{where} \quad (u; v)^\tau \equiv u^{|\tau|_{\sigma_1}} v^{|\tau|_{\sigma_2}}. \tag{11b}$$

We have thus at our disposal a summation operator $\mathbf{S}$ to solve non local difference equations of the form (11a). Applying it to the equation giving Λ yields explicit expressions, again involving sums indexed by H that are easy to evaluate numerically. More important, we see that $\Lambda(z)$, as well as the l_N, become infinite when $\lambda \to \lambda'_{\max}$ where $\lambda'_{\max}$ is determined as the solution of a certain transcental equation. Thus the maximum admissible throughput of the tree protocol with free access is precisely determined. We state here the easier case where $p = q = \frac{1}{2}$:

THEOREM 2: [FFH 1986] *The supremum $\lambda'_{\max}$ of arrival rates for which the tree protocol with free access using fair coins is stable, is the smallest positive root of the* *equation:*

$$1+\frac{2\exp(-2\lambda)}{1-2\lambda}\sum_{i\geq 0}2^i\exp(2\lambda/2^i)\Big[\exp(-\lambda/2^i)(1-\lambda/2^i)-1-2(\lambda/2^i)^2+2(\lambda/2^i)\Big]=0. \quad (12)$$

Numerically, $\lambda'_{\max} = 0.360177147$.

Thus, free access accepts a slightly er traffic rate than blocked access before destabilising. Viewed from the stations, free access is also easier to implement since there is no need of a continuous monitoring of the channel during inactivity periods. Therefore it seems to be the method of choice in this class of methods, especially after further optimisations to be discussed later are applied.

The asymptotic analysis is trickier than before. The Mellin transform of a harmonic sum involves a Dirichlet series related to amplitudes and frequencies:

$$\int_0^\infty \left(\sum_k a_k f(b_k x)\right) x^{s-1}\,dx = \omega(s)\int_0^\infty f(x)x^{s-1}\,dx \qquad \text{with } \omega(s)=\sum_k a_k b_k^{-s}. \quad (13)$$

In the case examined in Section 2, we just had $\omega(s) = (1-2^{s+1})^{-1}$, by summation of a geometric progression. Now, there appears sums indexed by semi-group H, of the form:

$$\omega(s)=\sum_{\tau\in H} r(\tau(0))(p^s;q^s)^\tau \quad (14)$$

where $r(u)$ is a standard C^∞ function (a combination of exponentials). To carry out the asymptotic analysis requires determining the singularities of such an $\omega(s)$.

There is an interesting topological property of semi-group H: the images of a given point z_0 under H are *dense* over the real interval $[\frac{\lambda}{p},\frac{\lambda}{q}]$ determined by the fixed points of σ_1,σ_2. Furthermore, these images are in a sense asymptotically *uniformly distributed*, and this fact yields the poles of $\omega(s)$, from which the asymptotic analysis can be completed.

THEOREM 3: [FFH 1986] *For the tree protocol with free access under a Poisson flow of arrivals of parameter* λ *with* $\lambda < \lambda'_{\max}$, *we have, neglecting small fluctuations:*

$$l_N \approx c(\lambda)N + o(N) \qquad \text{as } N\to\infty. \quad (15)$$

Thus, the burst response of the protocol is also fully characterised.

4. Free Access: Delay and Other Parameters

The previous section has demonstrated how to analyse the session (or collision resolution interval, CRI) lengths when there are N initial colliders –both exactly and asymptotically– characterising in passing the stability region. In particular, we introduced in (11b) a rather powerful summation operator **S** over semi-groups.

The problem now is to determine the steady state behaviour of important parameters of the protocol like delay etc. The basic approach is in two steps:

1. For parameters that are inductively defined on the tree structure, generating functions of the form (11a) can be set up. Thus, using the solution method (11b), we could also obtain expressions for their expectations over a session conditioned to be with N initial colliders.
2. The *values* of the generating functions, like $\Lambda(z)$, at $z = \lambda$ have a probabilistic interpretation, and they yield unconditional expectations of the parameters.

The new point is 2 above. To see it in the case of session length, observe that, by definition:

$$\Lambda(\lambda) = \sum_{N \geq 0} l_N e^{-\lambda} \frac{\lambda^N}{N!}. \tag{16}$$

Now, the coefficient of l_N in the above is nothing but the Poisson probability. But sessions start at "random" times, where the number of colliders also obey the Poisson distribution. Thus, the weighting in (16) yields the unconditional expectation of a CRI (or interval between two returns to the empty state, with no station backlogged).

That argument can be adapted to a delay analysis:

(i). The expectation of the cumulated delay experienced by all stations in a session, conditioned upon the number of initial colliders N, has a modified generating function $D(z)$ which satisfies an equation of the form (11a).

(ii). Over a large number s of sessions, by the "law of large numbers", the total delay will tend to $sD(\lambda)$. The total session length will tend to $s\Lambda(\lambda)$, with asymptotically $s\lambda\Lambda(\lambda)$ arrivals. Therefore, the unconditional (steady state) mean delay per message is equal to the quotient $D(\lambda)/(\lambda\Lambda(\lambda))$.

We shall summarise (see [FFHJ 1985] for detailed expressions) the previous discussion by:

THEOREM 4: [FFHJ 1985] *The steady state expectation of the delay experienced by a station under the free access tree protocol has an explicit expression in terms of the summation operator* S *of (11b) applied to standard functions.*

Similar results hold true for the variance of delay, the top-of-the-stack occupancy distribution etc. and the expressions obtained lend themselves to easy numerical evaluation. When $p = q = \frac{1}{2}$ they further simplify and somewhat resemble expressions found in Theorem 2. As an example of extensive numerical estimates, when $\lambda = 0.25$ and fair coins are used, the expected delay is only 4.79180 with a standard deviation of 11.2; there is probability 0.619 that a newly arriving message will be delivered immediately. Also, a detailed "low traffic" analysis can be conducted: We find that using a baised coin with $p = 2 - \sqrt{2} = 0.586$ slightly optimises delay for low traffic.

Finally, this type of analysis applies *mutatis mutandis* to other versions of the tree protocol. An unexpected result [MF 1985] is that using ternary instead of binary splittings in the basic protocol improves some of the characteristics by about 10%, at no extra cost of implementation: In the stack formulation, simply go down by two levels in the stack when a collision is encountered.

THEOREM 5: [MF 1985] *The tree protocol with free access and ternary splittings has a maximum admissible throughput of*

$$\lambda''_{\max} = 0.401599.$$

5. The Deterministic Tree Protocol

So far, the tree idea has been used, in accordance with the ALOHA principles, in a probabilistic manner. However, as already noticed by Capetanakis, it can also be implemented in a non randomised way, when the number of stations is fixed in advance.

Assume a network is configured to operate with a maximum number K of stations and consider the blocked access version of the protocol. If K is taken a power of 2, $K = 2^k$, then assign to each station a binary string of length k that uniquely identifies it. In each session, a station can use its predetermined string (instead of a random sequence) to participate in the splittings and schedule its retransmissions. This is the so-called ***deterministic tree protocol.***

Quite clearly, the maximum length of a session is now $2K - 1$, corresponding to a full binary tree of height k that develops when all K stations are active; the maximum delay for a message will always be less than $4K$. Thus, it seems, a little gain occurs from this worst case guarantee at the expense of a little loss in flexibility. We shall see later that there is actually more to it!

We wish to analyse the deterministic protocol under the assumption that each station has messages arriving at a Poisson rate of λ/K, resulting in a global rate that is Poisson(λ).

A not too surprising phenomenon is that, when $\lambda < \lambda_{\max}$, the deterministic protocol behaves basically like the standard randomised version. More strikingly, however, stability is retained when $\lambda > \lambda_{\max}$, provided λ remains less than $\frac{1}{2}$. We shall call such a region the ***hyperstable region*** and there is a sort of "phase transition" taking place at $\lambda_{\max}$.

Simulations reveal -and analysis confirms- that good approximations, in the hypersable region, (say within a few percent when $K = 256$) are obtained by letting $K \to \infty$. In particular, the queueing phenomena at the stations can be characterised in the limit.

The starting point for evaluations is relations that parametrise with k (or $K = 2^k$) the analysis in Eqs. (1-4). Let $l_N^{(k)}$ denote the expected length of a session starting with N initial colliders in a universe of $K = 2^k$ stations. Quantities $l_N^{(k)}$ are also relevant to the analysis of tries in computer algorithms and had been determined earlier by Trabb Pardo (1977):

THEOREM 6: *(i).* [Trabb Pardo] *The expected length of a session for the deterministic protocol, with N initial colliders and a universe of $K = 2^k$ stations, is:*

$$l_N^{(k)} = 1 + 2^{k+1} \sum_{j=1}^{k} \left[2^{-j} \left(1 - \frac{\binom{2^k - 2^j}{N}}{\binom{2^k}{N}}\right) - \frac{\binom{2^k - 2^j}{N-1}}{\binom{2^k}{N}} \right]. \tag{17}$$

(ii). When $k \to \infty$, each expectation $l_N^{(k)}$ tends monotonically to $l_N^{(k)}$: $l_N^{(k)} \to l_N$ with $l_N^{(k)} < l_N$.

The second assertion follows by simple computations. One can actually prove [Jacquet 1987] that, in the sense of Markov chains (i.e. the transition probabilities), the deterministic protocol converges to the probabilistic protocol provided $\lambda < \lambda_{\max}$. Accordingly, the queueing phenomena at the stations are asymptotically negligible. More interesting phenomena occur when $\lambda > \lambda_{\max}$:

PROPOSITION 7: *The system formed by queues at the stations coupled via the deterministic tree protocol is stable for $\lambda < \lambda^{\circ}_{\max}$, where:*

$$\lambda^{\circ}_{\max} = \frac{K}{2K - 1}.$$

This follows from the observation that the K stations receive service in at most $2K - 1$ slots. For large K, $\lambda^{\circ}_{\max}$ tends to $\frac{1}{2}$.

To continue the analysis, we observe that parameters of interest (session length, delay) should be re-normalised by K, if we want asymptotically meaningful quantities. We start with a simplified session model whose analysis is closely related to (17):

- Assume the channel is idle and that, at a beginning of a new session, each station becomes active with probability x. Then the expected length $l^{(k)}(x)$ of that session satisfies:

$$\frac{l^{(k)}(x)}{2^k} = 2^{-k} + 2 \sum_{j=0}^{k} \left[2^{-j} (1 - (1 - x)^{2^j}) - x (1 - x)^{2^j - 1} \right].$$

Essentially, $l^{(k)}(x)$ is a generating function of the $l_N^{(k)}$. As $k \to \infty$, the $2^{-k} l^{(k)}(x)$ converge to $L(x)$ given by:

$$L(x) = 2 \sum_{j=0}^{\infty} \left[2^{-j} (1 - (1 - x)^{2^j}) - x (1 - x)^{2^j - 1} \right]. \tag{18}$$

This function $L(x)$ plays an essential role in our subsequent analysis:

THEOREM 8: [Jacquet 1987] *Asymptotically for large K, the RV describing the queue length at any station, under the deterministic tree protocol with arrival rate λ in the hyperstable region, has generating function:*

$$q(z) = (1 - \mu) \frac{1 - z}{1 - z \exp(\mu(1 - z))}, \tag{19}$$

where μ is determined from λ by the equilibrium equation:

$$\mu = \lambda L(\mu)$$

and $L(x)$ is given by Eq. (18).

The reader will have recognised in (19) the generating function of an $M/D/1$ process (Markovian, i.e. Poisson arrivals/ deterministic service time): An $M/D/1$ queuing process with rate μ is a discrete time process $Q(t)$ such that:

$$Q(t+1) = |Q(t) - 1|^{+} + A(t)$$

where $A(t)$ is a Poisson variable with parameter μ. Thus Theorem 8 expresses the following fact: in the hyperstable region, the queuing system behaves asymptotically like a collection of independent $M/D/1$ processes.

Let us give a quick intuition about the probabilistic phenomena at stake. In the steady state, a fraction μ ($0 \leq \mu \leq 1$) of the population will be active, resulting in a session of length $\sim KL(\mu)$ during which there arrive in turn $K\lambda L(\mu)$ new messages. Whence the equilibrium equation: $\mu = \lambda L(\mu)$. Since the "server" (i.e. channel) is periodically available, each time a new session is started, the system, once normalised, resembles in the limit an $M/D/1$ process. It can, in effect, be proved rigourously that the state transition probabilities corresponding to finite values of K converge to the transition probabilities of the $M/D/1$ process.

6. Limit Distributions

We conclude our review of analytic results on the tree protocol with a few results obtained by Jacquet and Régnier [JR 1986]. Let X_N be a parameter (random variable) of the blocked access tree protocol, like session length (i.e. size of the associated tree), path length or height of the tree, when a session is started with N initial colliders. As N increases, those parameters X_N have complicated (exact) distributions that however tend to limiting distributions of a simple form.

THEOREM 9: [Jacquet, Regnier 1986] *Let S_N be the random variable representing the length of a session of the blocked access tree protocol in its randomised version, started with N initial colliders. As $N \to \infty$, the distribution of the random variable S_N tends to a limiting Gaussian distribution.*

Let $p_{N,k}$ be the probability that the modified session length $X_N = (S_N - 1)/2$ be equal to k, and introduce the bivariate generating function

$$P(z,u) = \sum_{N,k} p_{N,k}\, e^{-z} \frac{z^N}{N!} u^k. \qquad (20)$$

Function P is also a Poisson generating function when z is fixed: It represents the probability generating function (pgf) of X_N when N is itself Poisson with parameter z. The rather difficult proof proceeds in stages:

1. First use the recursive nature of the tree process to set up a ***non-linear difference equation*** satisfied by $P(z,u)$, like what has been done before:

$$P(z,u) = uP^2(\frac{z}{2},u) + (1-u)(1+z)e^{-z}. \qquad (21)$$

2. The problem is to obtain a good asymptotic approximation for $P(z,u)$ for fixed u and large z. Setting $L(z,u) = \log P(z,u)$, L satisfies a ***quasi-linear difference equation***. From there, the growth of $P(z,u)$ as well as its moments $P_u(z,1)$ and $P_{uu}(z,1)$ can be determined using Mellin transform techniques.
3. The characteristic function $P(z,e^{it})$ –after normalisation using mean and variance estimates from Point 2– converges as $z \to \infty$ to the characteristic function of a normally distributed variable, namely $e^{-t^2/2}$.
4. There now remains to translate the previous limit result under a Poisson model with parameter z to the case where N is fixed but large (The latter is the so-called Bernoulli model). What is needed here is an argument with which, if a_N is a sequence of numbers and

$$A(z) = \sum_N a_N \, e^{-z} \frac{z^N}{N!},$$

then $a_N \sim A(N)$. A general "semi-Tauberian" theorem (requiring the estimates to be valid in some region of the complex plane for z values) is given in [Jacquet, Régnier 1986] to that effect. By this device, results can be transfered from the Poisson to the Bernoulli case.

Several parameters, including various notions of height, can be analysed in this fashion. Also the method is general enough to accomodate the case of biased coins.

7. A Local Area Network Realisation

Besides being theoretically analysable, the tree protocol has many practical advantages. Its simplicity is comparable to that of Ethernet, since stations only need to maintain an integer index (representing their stack level): It is the way that index is managed that differs. Implemented in an asynchronous mode (unslotted time), it is compatible with the IEEE norm 802.3 . It is also fairly resistant to misinterpretations of channel feedback by stations. Last but not least, the deterministic version offers good worst-case guarantees on message delay when real-time constraints are present, and increased throughput results from the hyperstability phenomenon discussed earlier.

For those reasons, the SCORE project at INRIA has designed a prototype realisation, called LYNX, of a real time local area network based on the deterministic tree protocol, which is currently under industrial development. At present, the network consists of 14 stations: It necessitates only a simple modification of the Ethernet coupling boards, in accordance with what has been said earlier concerning norm compatibility.

Extensive measures have been conducted on LYNX, as well as with a network software emulator for 256 stations. In LYNX (as in any Ethernet network), collision slots being of a shorter duration, the observed performances are actually better than our previous computations imply. It is not difficult to "tune" the mathematical models to take this fact into account. (We have only refrained from doing so to keep the discussion simple). Measures and simulations amply confirm the analyses given in this paper. For instance, it appears clearly on that configuration that Ethernet will destabilise when $\lambda \approx 70\%$ while this occurs only at $\lambda \approx 90\%$ for the deterministic tree protocol. The rejection rate of messages, due to real time constraints not being satisfied, is also appreciably lowered by this change of protocol.

8. References

The general references we have used for this paper are:

[Massey 1981]: J. Massey, "'Collision resolution algorithms and random access communications", in *Multi-User Communication Systems*, G. Longo Ed., CISM Courses and Lectures **235**, Springer, New-York (1981), pp. 73-137.

[Massey 1985]: J. Massey Editor, *IEEE Transactions on Information Theory*, Special Issue on Random Access Communications **31** (1985).

[Knuth 1973]: D.E. Knuth, *The Art of Computer Programming: Sorting and Searching*, Addison-Wesley, Reading (1973).

Papers whose results are surveyed here are:

[FFH 1986]: G. Fayolle, P. Flajolet, M. Hofri, On a functional equation arising in the analysis of a protocol for a multi-access broadcast channel, *Adv. Appl. Prob.* **18** (1986), pp. 441-472.

[FFHJ 1985]: G. Fayolle, P. Flajolet, P. Jacquet, M. Hofri, Analysis of a stack algorithm for random multiple access communication, *IEEE Transactions on Information Theory*, Special Issue on Random Access Communications **31** (1985), pp. 244-254.

[MF 1985]: P. Mathys, P. Flajolet, *Q*-ary collision resolution algorithms in random-access systems with free or blocked channel access, *IEEE Transactions on Information Theory*, Special Issue on Random Access Communications **31** (1985), pp. 217-243.

[JR 1986]: P. Jacquet, M. Régnier, Trie partitioning process: limiting distributions, in *CAAP'86*, Lecture Notes in Computer Science **214** (1986), pp. 196-210.

[Jacquet 1987]: P. Jacquet, Evaluation of the queues in stations for the deterministic tree protocol, INRIA Res. Rep., in print (1987).

References contained in the above papers are not duplicated here.

Imprimé en France

par

l'Institut National de Recherche en Informatique et en Automatique