Stability properties of linear file-sharing networks

[Invited Presentation, Extended Abstract]

Lasse Leskelä
Helsinki University of Technology
Department of Mathematics and Systems Analysis
PO Box 1100, 02015 TKK, Finland
lasse.leskela@iki.fi

Philippe Robert and Florian Simatos
INRIA Paris-Rocquencourt
Domaine de Voluceau
BP 105, 78153 Le Chesnay, France
{philippe.robert, florian.simatos}@inria.fr

1. INTRODUCTION

File-sharing networks are distributed systems used to disseminate information among a subset of the nodes of the Internet. The general principle is the following: once a node of the system has retrieved a file it becomes a server for this file. The advantages of this scheme are numerous, and it has been used for some time now in peer-to-peer systems such as BitTorrent or Emule.

An improved version of this principle consists in splitting the original file into several pieces (called "chunks") so that a given node can retrieve simultaneously several chunks of the same file from different servers. The rate to get a given file thus increases significantly, as well as the global capacity of the file-sharing system since a node becomes a server of a chunk as soon as it has retrieved it. This improvement has interesting algorithmic implications since each node has to establish a matching between chunks and servers.

The efficiency of these systems can be considered from a transient and stationary perspective. In a transient scenario, a new file is typically owned by one node at time 0, and N other nodes are interested by it. One can look at the time needed so that a given node, or a fraction α of the N nodes, retrieves it. See for instance Yang and de Veciana [4] and Simatos *et al.* [2]. In stationarity, a constant flow of requests enters, and the first question is whether the capacity of the file-sharing system is sufficient to cope with this flow.

2. MODEL DESCRIPTION

The stationary behavior of a file-sharing system is investigated in a stochastic context: arrival times are random as well as chunk transmission times. For related models, see Susitaival *et al.* [3] and references therein. A simple strategy to disseminate chunks is considered: chunks are retrieved sequentially and a given node can be the server of only the last chunk it got. See Parvez *et al.* [1] for a detailed motivation of this situation.

When the file is cut into $n \geq 1$ chunks, the system is modeled by an $(n+1)$-dimensional (continuous-time) Markov process $(X^{(n)}(t) = X_0^{(n)}(t), \ldots, X_n^{(n)}(t))$, where for $0 \leq k \leq n$, $X_k^{(n)}(t)$ represents the number of nodes with k chunks are time t. Since chunks are retrieved sequentially, $X_k^{(n)}(t)$ represents the number of nodes with the k first chunks.

Model in the Single-Chunk Case. When $n = 1$, we analyze the two-dimensional Markov process $(X^{(1)}(t))$ with Q-matrix Q_1 given, for any $x = (x_0, x_1) \in \{0, 1, \ldots\}^2$, by

$$\begin{cases} Q_1(x, x+e_0) = \lambda, \\ Q_1(x, x-e_0+e_1) = \mu r(x)(x_1 \vee 1)\chi(x_0 > 0), \\ Q_1(x, x-e_1) = \nu x_1, \end{cases}$$

where $e_0 = (1, 0)$ and $e_1 = (0, 1)$ and χ is the indicator function. This Q-matrix has the following meaning: new nodes arrive according to a Poisson process with intensity λ; once a node has the file, it leaves after an exponential random variable with parameter ν; finally, the function $r(\cdot)$ with values in $[0, 1]$ represents an interaction term.
If $r(x) \equiv 1$, the $X_0(t)$ nodes which do not have the file are served by the $X_1(t)$ which have it; the boundary condition $x_1 \vee 1 = \max(1, x_1)$ prevents the second coordinate to end up in the absorbing state 0. One can consider $r(x) = 1 \wedge (x_0/x_1)$ to account for the fact that two nodes cannot serve one node simultaneously. The case $r(x) = x_0/(x_0 + x_1)$ models the system from a contact process perspective, where each node tries to get the file from a randomly chosen node at times of a Poisson process.

Model in the Multi-Chunk Case. When $n > 1$, we analyze the $(n+1)$-dimensional Markov process $(X^{(n)}(t))$ with Q-matrix Q_n given, for any $x = (x_0, \ldots, x_n) \in \{0, 1, \ldots\}^{n+1}$, by

$$\begin{cases} Q_n(x, x+e_0) = \lambda, \\ Q_n(x, x-e_{k-1}+e_k) = \mu_k(x_k \vee 1)\chi(x_{k-1} > 0), \\ Q_n(x, x-e_n) = \nu x_n \end{cases}$$

where k ranges from 1 to n.

The interpretation of this Q-matrix should be clear in view of the above discussion in the case $n = 1$. When $n = 1$, the system is characterized by the parameters (λ, μ, r, ν), whereas for $n > 1$, it is characterized by $(\lambda, \mu_1, \ldots, \mu_n, \nu)$. In both cases we are interested in determining the stability region, i.e., the set of parameters for which the corresponding Markov process is positive recurrent or transient.

The main technical difficulty to prove stability/instability results for this class of stochastic networks is that, except for the input, the Markov process has unbounded jump rates.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

VALUETOOLS 2009, October 20-22, Pisa, Italy

Copyright © 2009 ICST 978-963-9799-70-7

DOI 10.4108/ICST.VALUETOOLS2009.7815

3. STABILITY ANALYSIS

Fluid Limits. Classically, to analyze the stability properties of stochastic networks, one can use the limits of a scaling of the Markov process, the so-called fluid limits. The scaling consists in speeding up time by the norm $\|x\|$ of the initial state x, by scaling the state vector by $1/\|x\|$ and by letting $\|x\|$ go to infinity. This scaling is, however, better suited to "locally additive" processes, that is, Markov processes that behave locally as random walks. When the transition rates are unbounded, it may occur that the corresponding fluid limits have discontinuities; this complicates a lot the analysis of a possible limiting dynamical system.

A "fluid scaling" is nevertheless available for file-sharing networks. A description for a possible candidate $(x_i(t))$ for this limiting picture would satisfy the following differential equations,

$$\begin{cases} \dot{x}_0(t) = \lambda - \mu_1 x_1(t), \\ \dot{x}_k(t) = \mu_k x_k(t) - \mu_{k+1} x_{k+1}(t), \quad 1 \le k \le n-1, \\ \dot{x}_n(t) = \mu_n x_n(t) - \nu x_n(t). \end{cases}$$

This has been, up to now, one of the main tools to investigate mathematical models of file-sharing networks. In the (closely related) context of loss networks, an analogous limiting picture can be rigorously justified when the input rates and buffer sizes are scaled by some N and the state variable by $1/N$. This scaling is not useful here, since the problem is precisely of determining the values of λ for which the associated Markov is ergodic whereas in the above scaling λ is scaled. From this point of view the above dynamical system is quite informal.

Interacting Branching Processes. Since scaling techniques do not apply here, one needs to resort to different techniques to study stability: coupling the linear file-sharing network with interacting branching processes is a key idea. For $k \ge 1$, without the departures the process $(X_k(t))$ would be a branching process where individuals give birth to one child at rate μ_k. This description of such a file-sharing system as a branching process is quite natural. It has been used to analyze the transient behavior of these systems, see for instance Yang and de Veciana [4] and Simatos *et al.* [2]. A departure for $(X_k(t))$ can be seen as a death of an individual of class k and at the same time as a birth of an individual of class $k{+}1$. The file-sharing network can thus be described as a system of interacting branching processes with a constant input rate λ.

4. SUMMARY OF RESULTS

Heuristically, when $\mu < \nu$, then the second coordinate $X_1^{(1)}$ of $(X^{(1)}(t))$ is "stable", since it essentially behaves as a stable birth-and-death process. In this case the first coordinates $X_0^{(1)}$ is close to a queue with input rate λ and x^* servers which work at rate μ, where x^* is the expectation of $X_1^{(1)}$ under its stationary distribution. Although quite informal, this argument gives the following result.

Proposition 1 (Case $n = 1$). *Note $\rho = \mu/\nu$ and*

$$\lambda^* = \frac{\mu}{(1-\rho)(1-\log(1-\rho))}$$

when $\rho < 1$.
If $\rho < 1$ and $\lambda > \lambda^*$, then $X^{(1)}$ is transient.
For the converse, assume that for each $x_1 \ge 0$,

$$\lim_{x_0 \to \infty} r(x_0, x_1) = 1.$$

Under this assumption, $X^{(1)}$ is positive recurrent if $\rho \ge 1$, or if $\rho < 1$ and $\lambda < \lambda^*$.

Although a complete classification is obtained in the single-chunk case $n = 1$, the situation is more complex for $n > 1$. The next result proves the analog of the "good" case $\rho \le 1$ when $n = 1$.

Proposition 2 (Infinite stability region). *Fix $n \ge 1$, and assume that*

$$\mu_1 > \mu_2 > \ldots > \mu_{n-1} > \mu_n - \nu > 0.$$

Then $X^{(n)}$ is positive recurrent for any $\lambda > 0$.

Intuitively, what happens when the above condition fails is that a bottleneck appears at the first k such that $\mu_k < \mu_{k+1}$: similarly as in the case $n = 1$, there is a certain subsystem which is "stable", and therefore only offers a finite throughput to the previous nodes. We prove this result in the first non-trivial case $n = 2$. For the next proposition we need the Markov process $(X^S(t) = X_1^S(t), X_2^S(t))$ whose Q-matrix Q^S is defined by

$$\begin{cases} Q^S\big((x_1,x_2),(x_1+1,x_2)\big) = \mu_1(x_1 \vee 1), \\ Q^S\big((x_1,x_2),(x_1-1,x_2+1)\big) = \mu_2(x_2 \vee 1)\chi(x_1 > 0), \\ Q^S\big((x_1,x_2),(x_1,x_2-1)\big) = \nu x_2. \end{cases}$$

Proposition 3 (First bottleneck). *Fix $n = 2$, and assume that $\mu_2 - \nu > \mu_1$: then $(X^S(t))$ is ergodic. If π is its stationary distribution and $\lambda^* = \nu \mathbb{E}_\pi(X_2^S)$, then $(X^{(2)}(t))$ is ergodic for $\lambda < \lambda^*$ and transient for $\lambda > \lambda^*$.*

5. REFERENCES

[1] N. Parvez, C. Williamson, A. Mahanti, and N. Carlsson. Analysis of bittorrent-like protocols for on-demand stored media streaming. In *Proceedings of SIGMETRICS 2008*, pages 301–312, New York, NY, USA, 2008. ACM.

[2] F. Simatos, P. Robert, and F. Guillemin. A queueing system for modeling a file sharing principle. In *Proceedings of SIGMETRICS'08*, pages 181–192, New York, NY, USA, 2008. ACM.

[3] R. Susitaival, S. Aalto, and J. Virtamo. Analyzing the dynamics and resource usage of P2P file sharing by a spatio-temporal model. In *International Workshop on P2P for High Performance Comptutational Sciences*, 2006.

[4] X. Yang and G. de Veciana. Service capacity of peer to peer networks. In *Proceedings of IEEE Infocom'04*, pages 2242–2252. ACM, 2004.