
Avant-propos

J'ai découvert, un jour de 1986, l'ouvrage de G. Fishman [FIS 73] sur la *simulation* au centre de documentation de l'INRIA¹ à Rocquencourt. J'ai été aussitôt attiré par ce procédé numérique et par son utilisation pour l'évaluation de performances, en raison de sa puissance en termes de capacité de modélisation, et également par son assimilation relativement aisée sachant mes connaissances en informatique. En le pratiquant, j'ai su très vite que, derrière la façade informatique, se cache en réalité un outil expérimental scientifiquement encadré par la théorie des probabilités et de la statistique. En le côtoyant au fil des ans, j'ai compris que c'est aussi un art. Ce manuscrit transcrit ce que j'ai appris et retenu comme étant les leçons de base, à travers mon expérience aussi bien en recherche qu'en enseignement.

Ce manuscrit s'adresse en premier lieu aux élèves-ingénieurs des filières en informatiques et/ou en télécommunications ainsi qu'aux étudiant(e)s du cycle magistral (Master 1, Master 2) du même thème. La matière traitée dans ce manuscrit trouve sa place aussi bien dans un centre de recherche et de développement industriel que dans un laboratoire universitaire. De ce fait, ce manuscrit peut également être utilisé par des étudiants en cycle doctoral et des ingénieurs désireux d'approfondir leurs connaissances et d'élargir leur expertise.

Ce manuscrit a pour domaine d'applications les réseaux informatiques qui occupent une place de plus en plus importante dans la société moderne.

1. Institut national de recherche en informatique et en automatique.

Plus particulièrement, ce manuscrit se focalise sur un problème crucial qui est la maîtrise des performances des réseaux informatiques. Ces derniers sont devenus des systèmes de plus en plus complexes, et leur présence et importance ne cessent de croître dans tous les secteurs. La maîtrise de ces systèmes est devenue ainsi un enjeu majeur et un souci commun pour tous les acteurs : opérateurs, fournisseurs de services, fournisseurs d'accès, équipementiers, etc.

Nous nous intéressons ici plus particulièrement à la maîtrise des aspects *quantitatifs* de ces systèmes avec ses conséquences multiples, en supposant acquise la maîtrise des aspects fonctionnels. A titre d'exemple, imaginons que l'on vous confie la mission de créer un site Web pour promouvoir un produit. Pour mener à bien cette mission, vous devez bien sûr posséder une bonne maîtrise des technologies adéquates (HTML, PHP, MySQL, etc.), mais vous devez également choisir le matériel sur lequel tournera votre site. L'aspect quantitatif intervient alors inévitablement. Il faut en effet choisir un matériel *adéquat* qui répond aux attentes des utilisateurs, les attentes actuelles tout comme celles à venir. On voit apparaître les deux facettes duales :

- les performances à atteindre : c'est-à-dire les seuils de satisfaction des usagers relatifs à des indicateurs, par exemple, le temps de réaction du site ;
- le dimensionnement de ressources : c'est-à-dire l'estimation de l'ensemble des moyens physiques à mettre en œuvre pour atteindre ces performances, par exemple le type de serveur ou la bande passante à choisir dans le cas de l'hébergement de ce site par un *Data Center*.

Un surdimensionnement induit un coût supplémentaire inutile, alors qu'un sous-dimensionnement risquerait de faire fuir des clients potentiels. Dans les deux cas, il y a un impact négatif évident sur la compétitivité. Il faut donc établir une bonne *adéquation* entre les performances recherchées et les ressources à commettre. Ce *compromis* n'est pas facile à atteindre et cela passe par une maîtrise des aspects quantitatifs du système. C'est l'objectif de ce manuscrit.

Ce manuscrit présente une série de résultats théoriques et de procédés pratiques qui constitue une panoplie d'outils complémentaires destinés à l'étude quantitative des réseaux informatiques. La complémentarité de ces outils est pour l'auteur un élément indispensable si l'on veut mener à bien des études quantitatives sur les réseaux informatiques, tant les systèmes à étudier

sont complexes. Ce manuscrit a été conçu dans cet esprit et c'est pourquoi il comprend deux facettes :

- la simulation informatique avec ses langages et ses pratiques ;
- la modélisation mathématique avec la théorie des files d'attente, et, plus généralement, les processus de Markov.

Ce manuscrit se veut pragmatique et consacre une part importante à la simulation qui présente un double avantage :

- c'est un outil permettant une modélisation fine et proche de la réalité ;
- cet outil est accessible moyennant un minimum de connaissances en informatique et sans exigence d'expertise *avancée* en mathématiques.

L'auteur introduit deux logiciels de simulation OMNeT++ et OPNET (dans sa déclinaison IT GURU version académique). Le premier est présenté plus en détails, alors que le deuxième plus sommairement. Ces deux outils ont été choisis non seulement en raison de leur richesse fonctionnelle et de leur popularité, mais également en raison de leur accessibilité : tous les deux étant exploitables librement (sous certaines conditions) et téléchargeables depuis Internet. OMNeT++ est, de surcroît, un logiciel *open source*. En plus, le couple <OPNET, OMNeT++> permet de comprendre, de manière complémentaire, le fonctionnement d'un outil de simulation dans ses moindres détails et la puissance de modélisation qu'il offre.

Une étude par simulation, aussi détaillée qu'elle puisse être, est fondamentalement une étude statistique, c'est-à-dire une conclusion tirée d'une observation ponctuelle, limitée aussi bien dans le temps que par les circonstances sous lesquelles la simulation s'est déroulée. La modélisation analytique, quant à elle, offre des résultats solidement bâtis à partir d'un cadre théorique. Ces résultats théoriques sont à la fois généraux et paramétrables, donc valables pour différents *scenarii*. Cependant, les contraintes de calculabilité mathématique conduisent souvent à un modèle beaucoup moins fin que celui permis par un simulateur. De ce fait, la modélisation analytique constitue, avec la simulation, deux outils complémentaires pour l'étude quantitative des systèmes.

Ce manuscrit a choisi les réseaux informatiques comme champ d'applications cible. Cependant, les outils présentés sont génériques, ils sont applicables

aux systèmes informatiques en général, et plus largement à de nombreux autres domaines d'application (gestion de stock, par exemple).

L'auteur a pour souci également de rendre accessible les fondements théoriques de l'évaluation de performances à un public de niveau mathématique moyen. Pour cette raison, nous avons choisi de toujours mettre le formalisme au service de l'application et de l'assimilation des concepts, au prix, éventuellement, d'une certaine liberté prise par rapport à la pureté formelle. Pour cette raison également, nous avons choisi d'inclure un rappel des principaux concepts et résultats issus de la théorie des probabilités et de la statistique qui sont utilisés, implicitement ou explicitement, dans ce manuscrit.

Ces considérations sur les auditoires, les thèmes et les domaines d'applications conduisent à une structuration de ce manuscrit qui est la suivante :

– nous commençons par une brève introduction au problème d'évaluation de performances dans des réseaux informatiques. Nous introduisons ensuite les principes généraux de l'évaluation des performances et les principaux outils que sont la modélisation analytique et la simulation logicielle. Nous divisons la suite du manuscrit en trois parties, qui sont consacrées respectivement à la simulation, aux modèles analytiques ainsi qu'au cadre théorique sous-jacent (probabilités, statistique) ;

– la première partie (les chapitres 1 à 6, p. 31 à p. 152) se focalise sur la simulation. Le principe et la pratique de la réalisation logicielle d'un programme de simulation occupe bien entendu une place importante. Nous insistons également sur le fait que la simulation est la réalisation d'une expérience statistique virtuelle. De ce fait, la conception et la conduite d'une simulation, ainsi que la collecte et l'interprétation des données qui en sont issues, doivent respecter des règles précises que nous présenterons. Cette partie sera illustrée à travers des exemples réalisés en OMNeT++ ainsi qu'en OPNET ;

– la deuxième partie (les chapitres 7 à 11, p. 155 à p. 222) est consacrée à la modélisation par la théorie de files d'attente, et, plus généralement, les processus stochastiques. Nous y présentons les résultats théoriques de base, pour les files $M/M/1$ et leurs dérivés, les files $M/G/1$ et leurs dérivés, ainsi que pour les réseaux de files d'attente ;

– dans la troisième partie (les chapitres 12 à 14, p. 225 à p. 304), une brève introduction à la théorie des probabilités et de la statistique est fournie. En effet, aussi bien la modélisation analytique que la simulation procèdent

en réalité à une modélisation du système à étudier comme un système stochastique. Ces deux techniques font donc appel très largement à la théorie des probabilités et de la statistique. Cette partie contient aussi une introduction aux processus markoviens. Cette partie a été conçue pour faciliter d'éventuels besoins de rafraîchissement de certaines notions mathématiques, rendant ainsi ce document autocontenant tout en minimisant les redites.

Au cours de la rédaction de ce manuscrit, l'auteur s'est efforcé de faciliter les renvois et les références croisées, à la manière d'un document Html. En effet, un manuscrit *imprimé* est par définition un document *sérialisé*, alors que les matières traitées dans ce manuscrit comportent beaucoup d'imbrications mutuelles.