Expressing Performance Requirements using Regular Expressions to specify Stochastic Probes over Process Algebra Models[*]

Ashok Argent-Katwala   Jeremy T. Bradley   Nicholas J. Dingle

Department of Computing, Imperial College London
180 Queen's Gate, London SW7 2BZ, United Kingdom
Email: {a.argent-katwala,jb,njd200}@doc.ic.ac.uk

Abstract

This paper describes how soft performance bounds can be expressed for software systems using stochastic probes over stochastic process algebra models. These stochastic probes are specified using a regular expression syntax that describes the behaviour that must be observed in a model before a performance measurement can be started or stopped. We demonstrate the use of stochastic probes on a $661,960$ state parallel, redundant web server model to verify its passage-time performance characteristics.

1 Introduction

In this paper, we unify the description of soft performance bounds in software systems which have been modelled using stochastic process algebras. Soft performance bounds are an integral part of software and system performance validation. For example, we might have a service-level agreement (SLA) that a particular type of SQL query must return a result within 0.35 seconds 98.7% of the time; this would be derived from a passage-time quantile. Alternatively we might need to assure ourselves that the probability that a just-in-time compiler is running native code exactly 5 seconds after loading a Java applet is at least 0.8; this would be a transient constraint. Finally, we might have to demonstrate that the probability that a particular software failure mode is eventually reached is less that 0.002; this would be a steady-state measure.

In this paper, we unify the specification of all three soft performance criteria using a stochastic probe mechanism. A stochastic probe is a measurement device that defines arbitrary start and end points for such measures. We base the stochastic probe specification on a regular expression syntax (similar to action sequences [1]), which allows the user to describe the behaviour that should be seen in a software model before a performance measure is started or stopped. Unlike [1], however, these stochastic probes are themselves converted into stochastic process algebra components before interrogating the software model through process composition. In this way, we avoid trying to tailor notationally more complex logic specification paradigms, such as LTL or DLTL [2].

Why use stochastic process algebra models at all? Stochastic process algebras or SPAs provide us with a powerful formalism for modelling such software systems. SPAs provide the ability to describe and compose modular software while taking account of potentially many other processes that are concurrently competing for system resources. Additionally, the encapsulation of complex behaviour using random variables provides an abstraction that allows us to obtain soft bounds either through analysis or simulation.

[*]Information about paper length: to ensure readability for refereeing, we have left the paper in single column mode. In the ACM 2-column style the paper is within the required 12 page limit.

The paper is structured as follows: we will briefly introduce the stochastic probe concept and show how it can be used to describe all the three soft performance criteria (Section 2). For this study, we use the stochastic process algebra PEPA (although we could use any SPA in practice) and so we briefly outline PEPA in Section 2.4. Section 2.5 demonstrates some stochastic probe examples on a simple PEPA model. We then define how the regular expression probe specification can be translated into PEPA in Section 3. In Section 4, we outline how the ipc/DNAmaca tool takes a PEPA model, with integrated stochastic probe, through to analysis. Finally, we take a model of a parallel, redundant web server in Section 5 and run several stochastic probe measurements over it to demonstrate its performance behaviour.

2 Stochastic Probes

2.1 Introduction

Using regular expressions over action names to probe system behaviour is not a new concept. In particular, work with Propositional Dynamic Logic (PDL) introduced by [3, 4], and related dynamic logics has extended from reasoning about the decidability and complexity of systems to dealing with action predicates and fluents—notably in [5] using a PDL-based logic and [6] using LTL. In [5], De Giacoma and Lenzerini use regular expressions as state formulae in the modalities of PDL to describe possible or necessary flows of actions.

DLTL [2] is an extension of LTL to incorporate the dynamic aspects of PDL, which uses regular expressions to express programs. The modalities have been used in [1] to express fluents of actions which can in turn be used to pose purely functional queries of programs.

As an example of regular expressions being used in a very different stochastic context, stochastic regular expressions were introduced in [7] where the expressions themselves encode the stochastic behaviour. These expressions are then probabilistically pattern matched against strings. Our approach differs somewhat, in that our stochastic probes are non-interfering—all of the stochastic behaviour is already in the system and we merely want to define and observe certain measures.

2.2 Summary of the Stochastic Probe Definition Syntax

We represent stochastic probe definitions as regular expressions (as introduced by Kleene in [8]). In this case however, the atoms of the regular expression are action names drawn from the alphabet of the underlying process algebra model. This is a significant generalisation over [13], where the modeller is only allowed a fixed-structure two state probe. As such a probe definition, R, has the following elements:

Operator	Description	Operator	Description
a	action label	R, R	sequence operator
$R \mid R$	choice operator	$R?$	zero-or-one operator
$R\{n\}$	iterative operator	$R\{m, n\}$	range operator
R^+	positive closure operator	R^*	closure operator

a is an action label that matches a label in the system. Any action specified in the probe has to be observed in the model before the probe can advance a state. An action, a, can also be distinguished as a *start* or *stop* action in a probe and signifies an action which will start or stop a measurement, respectively.

R_1, R_2 is the sequential operator. R_1 is matched against the system's operation before R_2 is matched.

$R_1 \mid R_2$ is the choice operator. Either R_1 or R_2 is matched against the system being probed.

$R\{n\}$ is the iterative operator. n sequential copies of R are matched against the system e.g. $R\{3\}$ is simply shorthand for R, R, R.

$R\{m, n\}$ is the range operator. Between m and n copies of R are matched against the system's operation.

R^* is the closure operator, where zero or more copies of R are matched against the system.

R^+ is the positive closure operator, where one or more copies of R are matched against the system. It is syntactic sugar for R, R^*.

$R?$ is the optional operator, matching zero or one copy of R against the system. It is syntactic sugar for $R\{0, 1\}$.

2.3 Performance Analysis with Stochastic Probes

A stochastic probe is a fragment of stochastic process algebra, for our purposes PEPA, which describes the start and end points of a measurement that a modeller wishes to make of a system. Typically a modeller will wish to make one of the following types of measurement of a given system and will need to specify the start and end points for such a measurement:

steady-state measures are used to calculate the probability that the system ends up in the given set of states S eventually (independent of start state). In this case, a modeller would use a single probe to partition the state space into S and $\neg S$, The start actions, from the probe regular expression, define the entering transitions to S and the stop transitions, the leaving transition from S.

transient measures give the probability that a system is in a given set of states S_2 at time t, having started from one of the start states in S_1. In this case, we use the probe in slightly a slightly different way from when we were specifying a steady-state quantity. Here S_1 is taken to be the set of states that the system (and probe) are in after a start transition, and S_2 is similarly defined to be the set of states that occur after the stop transitions.

passage-time measures depict the probability that the system completes a passage from a given set of states S_1 to states in a given set S_2 in time t. As with transient measures, we take the intuitive approach of letting S_1 be defined by the start actions and S_2 by the stop actions of the probe. Passage-time analysis can produce either the probability density or cumulative distribution function of the amount of time taken. These are used most frequently for soft bounds on event-duration.

2.4 PEPA

We use the stochastic process algebra PEPA to represent our stochastic probes and models. PEPA [9] is a parsimonious stochastic process algebra that can describe compositional stochastic models. These models consist of components whose actions incorporate random exponential delays.

In fact, stochastic probes are SPA-independent and we could have tailored them to any suitable SPA. PEPA suited our needs here as it has an uncomplicated syntax which lends itself to describing the underlying ideas. As we will see, the probe is expressed as a single PEPA component, so that it can then be combined with the model being queried.

The syntax of a PEPA component, P, is represented by:

$$P \quad ::= \quad (a, \lambda).P \;\big|\; P + P \;\big|\; P \underset{S}{\bowtie} P \;\big|\; P/L \;\big|\; A \tag{1}$$

$(a, \lambda).P$ is a prefix operation. It represents a process which does an action, a, and then becomes a new process, P. The time taken to perform a is described by an exponentially distributed random variable with parameter λ. The rate parameter may also take the value $\top$, which makes the action passive in a cooperation (see below).

$P_1 + P_2$ is a choice operation. A race is entered into between components P_1 and P_2. If P_1 evolves first then any behaviour of P_2 is discarded and vice-versa.

$P_1 \underset{S}{\bowtie} P_2$ is the cooperation operator. P_1 and P_2 run in parallel and synchronise over the set of actions in the set S. If P_1 is to evolve with an action $a \in S$, then it must first wait for P_2 to reach a point where it is also capable of producing an a-action, and vice-versa. In an active cooperation, the two components then jointly produce an a-action with a rate that reflects the slower of the two components (usually the minimum of the two individual a-rates). In a passive cooperation, where P_1, say, can evolve with an $(a, \top)$-transition, the joint a-action inherits its rate from the P_2 component alone.

P/L is a hiding operator where actions in the set L that emanate from the component P are rewritten as silent τ actions (with the same appropriate delays). The actions in L can no longer be used in cooperation with other components.

A is a constant label and allows, amongst other things, recursive definitions to be constructed.

While on the subject of PEPA, it is worth noting that Gilmore and Hillston [10] have developed their own *feature interaction* logic which explores a PEPA model, assigning reward to component states for use in steady-state and transient-state analysis. This is an alternative technique to what we are trying to achieve here. Instead of using a logic to interrogate a model, we use the language's own cooperation operator to observe the key events that we wish to measure. By selectively sampling a model's behaviour in this way, we can simplify the task of picking the states that are relevant to our measure.

Our method has the benefit of not requiring the user to learn an entirely different paradigm (it being based on the process algebra that the model is described in). At the moment, it has the downside that, being observationally-based, in cannot distinguish actions that are generated by different copies of the same component. This is possible in a logic setting such as that set up by Gilmore et al [11] and Clark [12] for steady-state measure specification.

2.5 Stochastic Probe Examples

We give a few examples of stochastic probes as specified by regular expressions over a simple PEPA model.

$$\begin{aligned}
\mathrm{Light}_1 &\overset{\mathit{def}}{=} (\mathrm{red}, r_1).(\mathrm{green}, r_2).(\mathrm{blue}, r_3).\mathrm{Light} + (\mathrm{fail}, r_4).\mathrm{Light}_2 \\
\mathrm{Light}_2 &\overset{\mathit{def}}{=} (\mathrm{repair}, r_1).\mathrm{Light}_1
\end{aligned}$$

This describes a simple oscillating light that can periodically fail and then repair itself. We look at the following probe definitions:

Example 1: $(\mathrm{red}:\mathsf{start}, \mathrm{fail}:\mathsf{stop})$ creates a probe which waits for a red action, then absorbs any other behaviour before changing state again on seeing a fail. It could be used to measure the time to failure (passage-time). It could be used to identify all the states that make up the *correct* functioning of the light for a steady-state measure. It could be used to ascertain the probability of failing at a particular time, t, if used in a transient measure.

Example 2: $((\mathrm{fail}, \mathrm{red}:\mathsf{start}, \mathrm{red}:\mathsf{stop}) \mid (\mathrm{red}:\mathsf{start}, \mathrm{red}:\mathsf{stop}))$ creates a probe which could be used to analyse the time between successive red-actions after a fail or in normal operation.

Example 3: $(\mathrm{green}:\mathsf{start}, \mathrm{blue}\{3\}, \mathrm{blue}:\mathsf{stop})$ creates a probe which looks for the occurrence of a green action followed by four blue actions.

It is important to realise that the probe will never block the behaviour of the model it is synchronising with. As described in the next section, the probe will absorb behaviour (without altering state) which it sees that is not part of its next specified action. In LTL terms, a modeller can only ask a question, “will I ever see this behaviour?” and not “will I see exactly this behaviour next?”. Asking the latter question with a probe would potentially change the functional behaviour of the model, and as we have said, probes are meant to be neutral in the effect they have on the model.

3 Converting a Probe Specification to PEPA

Formally, a stochastic probe has the following syntax:

$$\begin{array}{rcl} R & ::= & a \mid T,T \mid (S) \\ S & ::= & T \mid T|S \\ T & ::= & R \mid R\{n\} \mid R\{m,n\} \mid R^{+} \mid R^{*} \mid R? \\ a & ::= & act \mid act:start \mid act:stop \end{array} \tag{2}$$

In diagrammatic terms, Fig. 1 shows the conversion of the individual regular expression elements to process algebra components.

Symbolically, we first build a parse tree, during which the syntactic sugar is removed, converting R^{+} to R, R^{*} and $R\{n\}$ to $R, R, \ldots, R$ to give n explicit copies of R. At each stage of the conversion, only the topmost element in the tree is considered, and the subtrees are handled recursively. At each call we take as input the name of the as-yet-undefined component we are to define and the component name we are to end at. Calling these labels P and Q respectively, we translate the six remaining elements as follows.

Action a: is always a leaf node and translates to $\mathrm{P} \overset{def}{=} (\mathrm{a}, \top).\mathrm{Q}$

Choice $R_1 \mid R_2$: translates to $\mathrm{P} \overset{def}{=} \mathrm{N}_1 + \mathrm{N}_2$ and the algorithm is repeated for the sub-trees with R_1 running from N_1 to Q and R_2 running from N_2 to Q.

Sequence R_1, R_2: translates to R_1 running from P to N_1 and R_2 from N_1 to Q.

Closure R^{*}: becomes $\mathrm{P} \overset{def}{=} \mathrm{N}_1 + \mathrm{Q}$, where R is translated running from N_1 to P.

Zero-or-one $R?$: becomes $\mathrm{P} \overset{def}{=} \mathrm{N}_1 + \mathrm{Q}$, where R is translated running from N_1 to Q.

Range $R\{n,m\}$: can be rewritten as $(R\{n\}, R\{0, m-n\}) \equiv (R\{n\}, (R, R\{0, m-n-1\})?)$, where $R\{0,1\} \equiv R?$

After the complete regular expression has been translated, the last labels in the probe description are set to point at the beginning of the probe. Thus the probe is made cyclic and, for a well-specified probe, irreducibility is preserved in the whole system.

To make the probe non-interfering when composed with the system under analysis, we ensure all actions in the cooperation set are enabled in every state of the probe. We do this by adding null loops in each state which absorb any action not immediately available from that state. We do not discriminate between the untagged and tagged forms of action labels—i.e. a, a:start and a:stop are all considered equal and the loops are only ever the plain, untagged version of the action. Thus, the probe will never change the observable behaviour of any system it is composed with.

This procedure gives us a valid PEPA fragment which will behave properly as a probe. However, for our analysis with ipc/DNAmaca we also need to be able to tell, purely from the state of the probe whether it

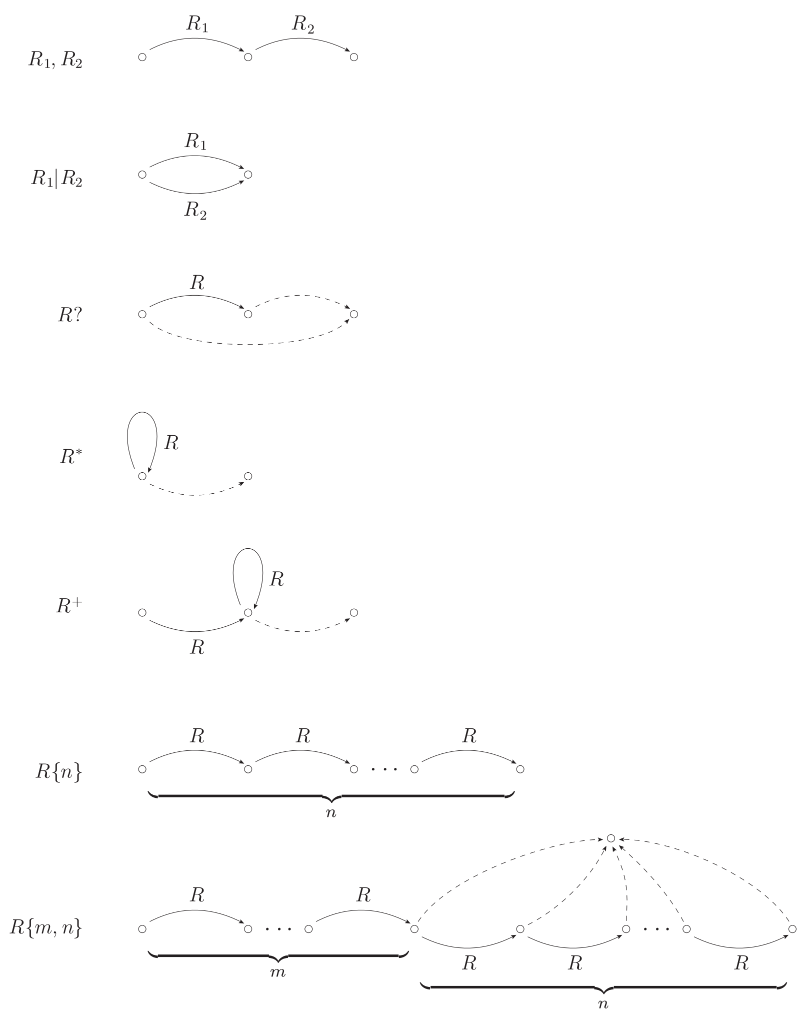

Fig. 1. The representation of the distinct regular expression terms as PEPA components

is running or stopped. The procedure above will not do this for any probe which contains a choice where there is a mixture of plain actions and tagged ones.

So, at a cost of at most doubling the state space of the probe, we make an observationally equivalent probe which has this partition in its state-space which we will need later in the analysis pipeline. The new process has the same state space as the cooperation over all the tagged actions of the probe built above with a tracker with two states—Running and Stopped.

For example, for the expression $(\text{foo}, (\text{foo}{:}\texttt{start} \mid \text{bar}), (\text{foo}{:}\texttt{stop} \mid \text{bar}))$, the approach above would generate this process, depicted in Fig. 3 with the PEPA description of Fig. 2. Plainly, the probe can bypass the tagged foo actions whenever they are presented so we cannot know whether the probe is running or not based on the state alone.

Our two-state model is simply the one used in [13] that keeps track of whether we have started or not, and silently consumes start actions when running and stop actions when stopped.

$$
\begin{aligned}
\mathrm{Probe} &\overset{\mathit{def}}{=} (\mathrm{foo},\top).\mathrm{Probe1} + (\mathrm{foo}{:}\mathsf{start},\top).\mathrm{Probe}\\
&\quad + (\mathrm{bar},\top).\mathrm{Probe} + (\mathrm{foo}{:}\mathsf{stop},\top).\mathrm{Probe}\\
\mathrm{Probe1} &\overset{\mathit{def}}{=} \mathrm{Probe2} + \mathrm{Probe3} + (\mathrm{foo},\top).\mathrm{Probe1} + (\mathrm{foo}{:}\mathsf{stop},\top).\mathrm{Probe1}\\
\mathrm{Probe2} &\overset{\mathit{def}}{=} (\mathrm{foo}{:}\mathsf{start},\top).\mathrm{Probe}\\
\mathrm{Probe3} &\overset{\mathit{def}}{=} (\mathrm{bar},\top).\mathrm{Probe4} + (\mathrm{foo},\top).\mathrm{Probe3}\\
&\quad + (\mathrm{foo}{:}\mathsf{start},\top).\mathrm{Probe3} + (\mathrm{foo}{:}\mathsf{stop},\top).\mathrm{Probe3}\\
\mathrm{Probe4} &\overset{\mathit{def}}{=} \mathrm{Probe5} + \mathrm{Probe6} + (\mathrm{foo},\top).\mathrm{Probe4} + (\mathrm{foo}{:}\mathsf{start},\top).\mathrm{Probe4}\\
\mathrm{Probe5} &\overset{\mathit{def}}{=} (\mathrm{foo}{:}\mathsf{stop},\top).\mathrm{Probe}\\
\mathrm{Probe6} &\overset{\mathit{def}}{=} (\mathrm{bar},\top).\mathrm{Probe}
\end{aligned}
$$

Fig. 2. The PEPA description of the probe specification $(\mathrm{foo}, (\mathrm{foo}{:}\mathsf{start} \mid \mathrm{bar}), (\mathrm{foo}{:}\mathsf{stop} \mid \mathrm{bar}))$.

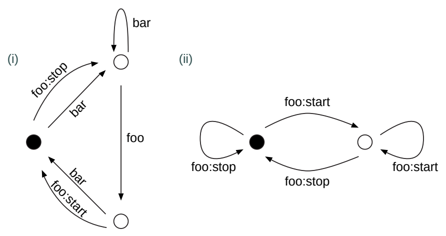

Fig. 3. (i) Probe and (ii) two-state running/stopped tracker

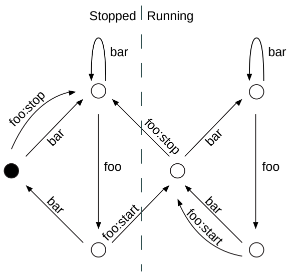

Fig. 4. Full probe, made by cooperating the 3-state probe and the tracker

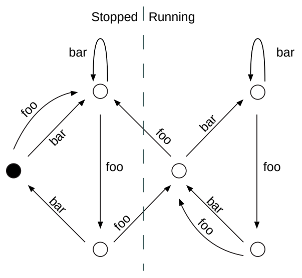

Fig. 5. Complete, relabelled probe

Now the full probe,shown in Fig. 4, is made from:

$$\begin{aligned}
\mathrm{Stopped} &\overset{\mathit{def}}{=} (\mathrm{foo}\!:\!\mathsf{start}, \top).\mathrm{Running} + (\mathrm{foo}\!:\!\mathsf{stop}, \top).\mathrm{Stopped} \\
\mathrm{Running} &\overset{\mathit{def}}{=} (\mathrm{foo}\!:\!\mathsf{start}, \top).\mathrm{Running} + (\mathrm{foo}\!:\!\mathsf{stop}, \top).\mathrm{Stopped} \\
\mathrm{FullProbe} &\overset{\mathit{def}}{=} \mathrm{Stopped} \underset{L_p}{\bowtie} \mathrm{Probe}
\end{aligned}$$

where $L_p = \{\mathrm{foo}\!:\!\mathsf{start}, \mathrm{foo}\!:\!\mathsf{stop}\}$.

Note that at this level we are treating foo: start and foo: stop as ordinary action names that happen to end in ":start" and ":stop", so the plain foo above is not involved in the synchronisation.

If we used relabelling like that in CCS [14], where the output actions are rewritten only on output, to convert foo: start and foo: stop into plain foo, we would be done. PEPA, however, does not include relabelling, but we can still explicitly build a process that would have the same state-space and behaviour as FullProbe but with all the foo: start and foo: stop actions relabelled to plain foo on output, as shown in Fig. 5.

4 The ipc/DNAmaca Tool Chain

4.1 ipc: The Imperial PEPA Compiler

ipc performs a translation from PEPA to a stochastic Petri net formalism (similar to that described in [15]), and also incorporates any extra logic necessary for expressing the passage-time, transient or steady-state query derived from Section 3. The ipc compiler consists of:

1. `.pepa` file parser
2. PEPA normal form translator
3. component state space explorer
4. DNAmaca component linker and `.mod` file generator
5. PEPA-measurement specifier
6. command line parser for passage-time and steady-state queries

The `.pepa` file format allows, for instance, arbitrary numbers of sequential prefixes and also arbitrary numbers of summation and cooperation operations attributed to a single constant label. The normal form in question strictly enforces the binary summation and cooperation, insisting on constant labels after each operation; it therefore also has to take care of the unique labelling of all components' states.

The component linker takes the DNAmaca description of the individual PEPA components and creates shared transitions with appropriate preconditions and actions to represent the cooperation over shared actions.

The PEPA-measurement specifier augments the input PEPA model with the process algebra probe (see Section 3) and adds the command for the requisite measurement specification (passage-time, transient or steady-state) to the DNAmaca file.

4.2 DNAmaca: Markovian Analyser

DNAmaca [16] is a modelling language for Markov and semi-Markov chains. Previous publications [17, 18, 19, 20, 21, 22] describe the mathematical foundation for the calculation of steady-state, transient and passage-time distributions in such models, so we do not dwell on the details here. Rather in Section 4.3, we will briefly describe the theory behind the *uniformization* technique [23, 24] used by the HYDRA release [25, 26] of DNAmaca to calculate passage-time quantities in Markov models.

The DNAmaca interface language, to which ipc compiles, is described in [16], with the passage-time specification syntax in [19]. ipc uses this to describe the performance measurement to be carried out on the stochastic probe, and thus on the model itself.

4.3 Passage-time Calculation

PEPA models reduce to an underlying continuous-time Markov chain (CTMC), so we consider an n state CTMC with $n \times n$ generator matrix $Q = q_{ij}$. Solving the linear system $\pi Q = 0$ subject to $\sum \pi_i = 1$ gives us the steady state vector, π. We calculate passage-time densities from many source states $\vec{i}$ to many target states $\vec{j}$ by means of an efficient uniformization-based analysis.

Uniformization [23, 24] transforms a CTMC into one in which all states have the same mean holding time $1/q$, by allowing *invisible* transitions from a state to itself. After normalisation of the generator matrix rows with an associated Poisson process of rate q, we obtain a one-step DTMC transition matrix P, given by:

$$P = Q/q + I \tag{3}$$

where $q > \max_i |q_{ii}|$ (to ensure that the DTMC is aperiodic).

While uniformization is normally used for transient analysis, it can also be employed for the calculation of response-time densities and quantiles [27, 28]. We add an extra, absorbing state to our uniformized chain, which is the sole successor state for all target states (thus ensuring we calculate the *first* passage-time density). We denote by P' the one-step transition matrix of the modified, uniformized chain. Remembering that the time taken to traverse a path with n hops in this chain will have an Erlang distribution with parameters n and q, the density of the time taken to pass from a set of source states $\vec{i}$ into a set of target states $\vec{j}$ is given by:

$$f_{\vec{i}\vec{j}}(t) = \sum_{n=1}^{\infty} \frac{q^n t^{n-1} e^{-qt}}{(n-1)!} \sum_{k \in \vec{j}} \pi_k^{(n)} \tag{4}$$

where:

$$\pi^{(n+1)} = \pi^{(n)} P' \qquad \text{for } n \geq 0$$

with:

$$\pi_k^{(0)} = \begin{cases} 0 & \text{for } k \notin \vec{i} \\ \pi_k / \sum_{j \in \vec{i}} \pi_j & \text{for } k \in \vec{i} \end{cases} \tag{5}$$

and in which $\boldsymbol{\pi}$ is any non-zero solution to $\pi = \pi P$. The corresponding passage-time cumulative distribution function is given by:

$$F_{\vec{i}\vec{j}}(t) = \sum_{n=1}^{\infty} \left\{ \left(1 - e^{-qt} \sum_{k=0}^{n-1} \frac{(qt)^k}{k!} \right) \sum_{k \in \vec{j}} \pi_k^{(n)} \right\} \tag{6}$$

Truncation is employed to approximate the infinite sum in Eq. (4) (and Eq. (6)), terminating the calculation when the Erlang term drops below a specified threshold value. Concurrently, when the convergence criterion

$$\frac{||\pi^{(n+1)} - \pi^{(n)}||_\infty}{||\pi^{(n)}||_\infty} < \epsilon \tag{7}$$

is met, for given tolerance ϵ, the steady state probabilities of P' are considered to have been obtained with sufficient accuracy and no further multiplications with P' are performed.

5 Example: High-availability Web Server

In this section, we demonstrate performance and reliability analysis of a high-availability distributed web server. Using stochastic probes to pick out key system passage-times, we make soft performance and reliability statements about the number of requests that the system can handle within a given time, as well as the probability of complete failure within the first minutes or hours of operation.

A typical application scenario for such a system is a web-based current events news feed which must meet strict quality-of-service requirements on availability and response-time. Regardless of whether the underlying technology is web-based or not, such systems require careful performance engineering to achieve peak efficiency [29]. In part, the strict QoS requirements are met by skewing the prioritisation for *fast* reads over writes so that writes are buffered and only processed at times of low read load. The consequence of this is that there is no guarantee that a reader will see the latest version of the site although high availability is maintained.

The system is made up of a cluster of servers, each of which can fail and be repaired independently. If all of the servers fail then a special recovery mechanism can restart them all. We now proceed to describe the components of the system.

5.1 The server model

The server receives read requests each of which incurs a read lookup before the server is available to serve the next request (states Server and $\mathrm{Server\_read}$ below). A successful write request requires that none of the servers are in the process of performing a read access, so that all servers can be simultaneously updated (therefore no $\mathrm{s\_write}$ action is allowed in the $\mathrm{Server\_read}$ state). The server may fail and while failed (the $\mathrm{Server\_fail}$ state) read requests are not intercepted (no $\mathrm{s\_read\_request}$ activities) whereas write requests ($\mathrm{s\_write}$) are absorbed without action. It is assumed that server resynchronisation occurs during recovery. The write costs are different in the functioning and failure states (the costs are quantified by variables r_{s_w} and $r_{s_{fsw}}$ respectively). Failed servers may be repaired individually ($\mathrm{s\_fail\_recover}$) or collectively ($\mathrm{s\_fail\_recover\_all}$).

$$
\begin{aligned}
\mathrm{Server} &\overset{\mathit{def}}{=} (\mathrm{s\_read\_request}, \top).\mathrm{Server\_read} \\
&\quad + (\mathrm{s\_fail}, r_{s_f}).\mathrm{Server\_fail} \\
&\quad + (\mathrm{s\_write}, r_{s_w}).\mathrm{Server} \\
\mathrm{Server\_read} &\overset{\mathit{def}}{=} (\mathrm{s\_read\_lookup}, r_{s_{rl}}).\mathrm{Server} \\
\mathrm{Server\_fail} &\overset{\mathit{def}}{=} (\mathrm{s\_fail\_recover}, r_{s_{fr}}).\mathrm{Server} \\
&\quad + (\mathrm{s\_fail\_recover\_all}, \top).\mathrm{Server} \\
&\quad + (\mathrm{s\_write}, r_{s_{fsw}}).\mathrm{Server\_fail}
\end{aligned}
$$

5.2 Server groups

The recovery of the servers is co-ordinated by a server group manager. The responsibility of this component is to witness server failures and recoveries from failures. When the server number, S, is reached and all servers have failed, the server group is restarted with all failures recovered simultaneously (as a result of a high-priority repair).

$$\begin{aligned}
\mathrm{Server\_group_0} &\overset{\mathit{def}}{=} (\mathrm{s\_fail}, \top).\mathrm{Server\_group_1} \\
\mathrm{Server\_group_i} &\overset{\mathit{def}}{=} (\mathrm{s\_fail}, \top).\mathrm{Server\_group_{i+1}} \\
&+ (\mathrm{s\_fail\_recover}, \top) \\
&\quad .\mathrm{Server\_group_{i-1}} : 1 \leq \mathrm{i} < \mathrm{S} \\
\mathrm{Server\_group_S} &\overset{\mathit{def}}{=} (\mathrm{s\_fail\_recover\_all}, r_{s_{gsfra}}) \\
&\quad .\mathrm{Server\_group_0}
\end{aligned}$$

Taking, for example, $S = 4$, the process instantiation expression for the server cluster is as shown below:

$$\begin{aligned}
\mathrm{Servers} \overset{\mathit{def}}{=} &(\mathrm{Server} \underset{\mathcal{L}}{\bowtie} \mathrm{Server} \underset{\mathcal{L}}{\bowtie} \mathrm{Server} \underset{\mathcal{L}}{\bowtie} \mathrm{Server}) \\
&\underset{\mathcal{L}'}{\bowtie} \mathrm{Server\_group_0}
\end{aligned}$$

where $\mathcal{L} = \{\mathrm{s\_write}, \mathrm{s\_fail\_recover\_all}\}$ and $\mathcal{L}' = \mathcal{L} \cup \{\mathrm{s\_fail\_recover}\}$.

5.3 Buffered writes

The write buffer manages the promotion of buffered writes to server write actions. To ameliorate the relative infrequency of all the servers being available simultaneously to perform an s_write, write requests are necessarily buffered until the buffer capacity, B, is reached. When an s_write action does occur, the entire write buffer is executed and emptied.

$$\begin{aligned}
\mathrm{Write\_buffer_0} &\overset{\mathit{def}}{=} (\mathrm{b\_write}, \top).\mathrm{Write\_buffer_1} \\
\mathrm{Write\_buffer_i} &\overset{\mathit{def}}{=} (\mathrm{b\_write}, \top).\mathrm{Write\_buffer_{i+1}} \\
&+ (\mathrm{s\_write}, \top).\mathrm{Write\_buffer_0} \\
&\qquad : 1 \leq i < B \\
\mathrm{Write\_buffer_B} &\overset{\mathit{def}}{=} (\mathrm{s\_write}, \top).\mathrm{Write\_buffer_0}
\end{aligned}$$

5.4 Web authors and browsers

Web authors (writers) issue web content to the system. Web browsers are the readers in our system. It is assumed that there are multiple writers and, comparatively, a much larger number of web browsers (readers). This accounts for the prioritisation of read access over writes.

Writers may perform only buffered writes. They have no capacity to perform a server write directly.

$$\begin{aligned}
\mathrm{Writer} &\overset{\mathit{def}}{=} (\mathrm{b\_write}, r_{w_{bw}}).\mathrm{Writer\_writ} \\
\mathrm{Writer\_writ} &\overset{\mathit{def}}{=} (\mathrm{w\_reset}, r_{w_r}).\mathrm{Writer}
\end{aligned}$$

With three writers ($W = 3$) the process instantiation expression for the writers is as shown below:

$$\mathrm{Writers} \overset{\mathit{def}}{=} \mathrm{Writer} \underset{\emptyset}{\bowtie} \mathrm{Writer} \underset{\emptyset}{\bowtie} \mathrm{Writer}$$

Readers send read requests and then await the response from the server.

$$\begin{aligned}
\mathrm{Reader} &\stackrel{\mathit{def}}{=} (\mathrm{s\_read\_request}, r_{r_{srr}}) \\
&\quad .(\mathrm{s\_read\_lookup}, \top).\mathrm{Reader\_read} \\
\mathrm{Reader\_read} &\stackrel{\mathit{def}}{=} (\mathrm{r\_reset}, r_{r_r}).\mathrm{Reader}
\end{aligned}$$

If we model three readers ($R = 3$), for instance, we get:

$$\mathrm{Readers} \stackrel{\mathit{def}}{=} \mathrm{Reader} \underset{\emptyset}{\bowtie} \mathrm{Reader} \underset{\emptyset}{\bowtie} \mathrm{Reader}$$

5.5 The system equation

The system is built by composing the behaviours of the simpler component to form the behaviour of the model as a whole:

$$\begin{aligned}
\mathrm{Environment} &\stackrel{\mathit{def}}{=} \mathrm{Writers} \underset{\emptyset}{\bowtie} \mathrm{Readers} \\
\mathrm{Web\_cluster} &\stackrel{\mathit{def}}{=} \mathrm{Servers} \underset{\{\mathtt{s\_write}\}}{\bowtie} \mathrm{Write\_buffer}_0 \\
\mathrm{System} &\stackrel{\mathit{def}}{=} \mathrm{Environment} \underset{\mathcal{N}}{\bowtie} \mathrm{Web\_cluster}
\end{aligned}$$

where $\mathcal{N} = \{\mathrm{b\_write}, \mathrm{s\_read\_request}, \mathrm{s\_read\_lookup}\}$.

6 Results

In this section we describe three distinct stochastic probes which we use to extract passage-time distributions from the underlying Markov model via the ipc/DNAmaca tool. We focus on passage-times as this is still a relatively rare piece of analysis within stochastic process algebra modelling.

Probe	S	B	W	R	States	Transitions
A	5	4	3	3	362,240	4,061,440
B	5	4	3	3	581,200	6,516,356
C	5	4	3	5	661,960	9,440,988

Tab. 1. Size of analysed model for given probes and model configurations

Tab. 1 shows the size of the Markovian model to be analysed once the stochastic probe query has been incorporated. The probes used are:

Probe A: measures the length of time to process at least 4 read events fully, i.e. observe 4 s_read_lookup actions.

$$\Pi_A = \mathrm{s\_read\_request}:\mathsf{start}, (\mathrm{s\_read\_lookup}\{4\}), \mathrm{r\_reset}:\mathsf{stop}$$

The measurement is started on seeing the first s_read_request and terminates, after the passage of at least 4 s_read_lookup actions, upon completion of the next r_reset. The probability density function of this passage is show in Fig. 6. From this we can see that all 4 read actions will have occurred by time $t = 3$.

Probe B: observes the length of time it takes either to observe 4 completed reads or 2 completed writes.

$$\begin{aligned}\Pi_B \;&=\; (\text{s\_read\_request}:\mathsf{start}, (\text{s\_read\_lookup}\{4\}), \text{r\_reset}:\mathsf{stop}) \\ &\quad \mid (\text{b\_write}:\mathsf{start}, \text{b\_write}\{2\}, \text{w\_reset}:\mathsf{stop})\end{aligned}$$

The probe is started by either seeing a s_read_request or a b_write. It then either looks for 4 s_read_lookup actions or 2 further b_write actions, before terminating on an appropriate read or write reset. Fig. 7 and Fig. 8 show the probability density function and cumulative distribution function for the passage, respectively. From the cumulative function, we can deduce the soft bound, that the system will complete either 4 reads or 2 writes within 4 time units with probability 0.98856.

Probe C: extracts the passage-time to first full system failure.

$$\Pi_C = (\text{s\_read\_request}:\mathsf{start} \mid \text{b\_write}:\mathsf{start}), \text{s\_fail\_recover}:\mathsf{stop}$$

A full system failure is said to have occurred if all the servers have independently failed and none has managed to recover. The probe starts the measurement on encountering a read request from a reader or a buffer write event from a writer. The measurement is only stopped when a full fail event (s_fail_recover) has occurred. This is a good example of a rare event which would be hard to simulate as the absolute number of complete failures is going to be very small for a given execution trace. Fig. 9 and Fig. 10 show the density function and cumulative distribution for this passage, respectively, and demonstrate how infrequent this occurrence is. From the cumulative distribution we can ascertain the soft reliability bound that the system will fail within the first 90 time units is 0.00559.

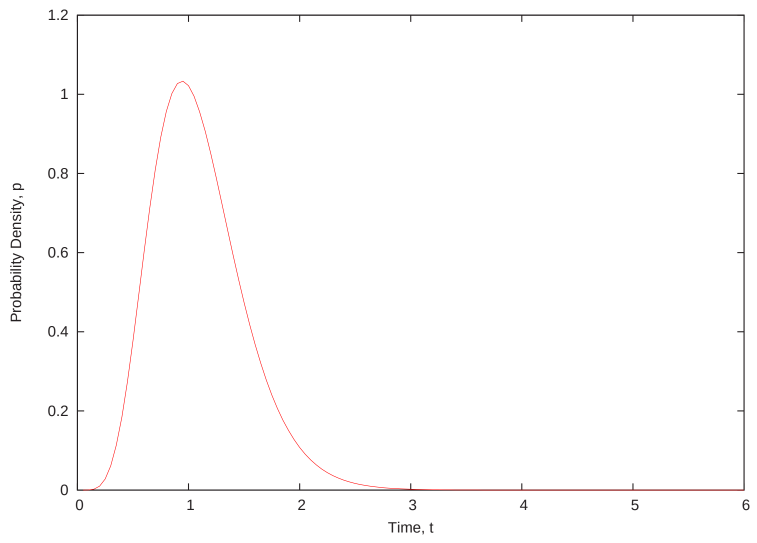

Fig. 6. Probability density function for Probe A: a $362,240$ state system

7 Conclusion

In this paper, we have introduced stochastic probes as a means of expressing three types of performance analysis in a unified manner. We presented a regular expression language which specifies the stochastic probe and is then converted into a stochastic process algebra component. We showed how this could be combined with the original SPA model and analysed to provide soft performance and reliability bounds.

We finished by analysing a $661,960$ state PEPA model of a parallel web server for a number of soft performance and reliability bounds.

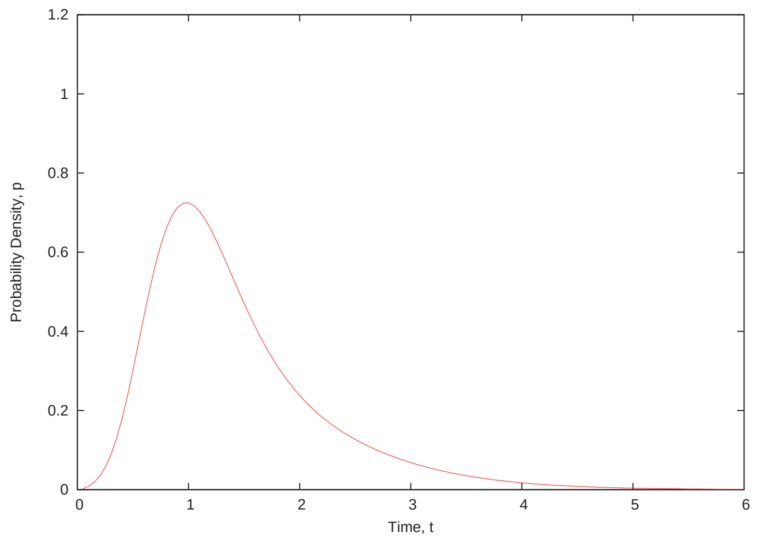

Fig. 7. Probability density function for Probe B: a $581,200$ state system

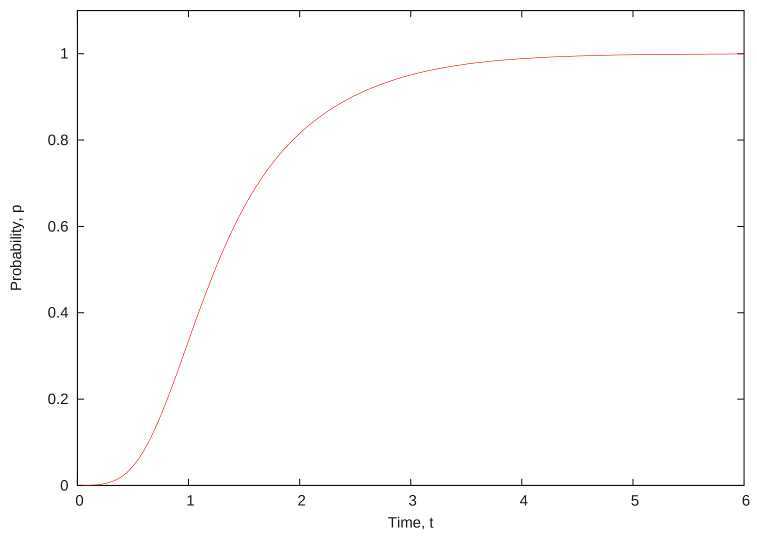

Fig. 8. Cumulative distribution function for Probe B: a $581,200$ state system

Acknowledgements

The authors would like to thank Stephen Gilmore, Tony Field and Helen Wilson for helpful remarks on the ideas behind this paper.

References

[1] L. Giordano, A. Martelli, and C. Schwind, “Reasoning about actions in dynamic linear time temporal logic,” *Logic Journal of the IGPL*, vol. 9, no. 2, pp. 273–287, 2001.

[2] J. G. Henriksen and P. S. Thiagarajan, “Dynamic linear time temporal logic,” *Annals of Pure and Applied Logic*, vol. 96, no. 1–3, pp. 187–207, 1999.

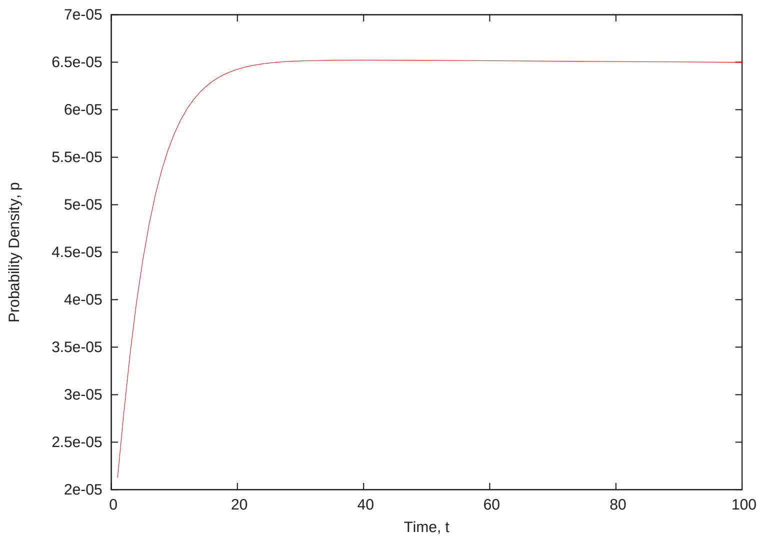

Fig. 9. Probability density function for Probe C: a $661,960$ state system

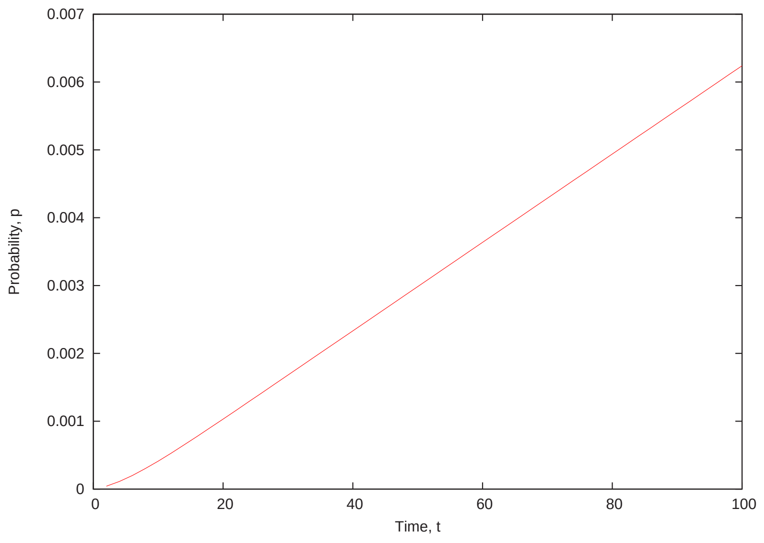

Fig. 10. Cumulative distribution function for Probe C: a $661,960$ state system

[3] M. J. Fischer and R. E. Ladner, “Propositional dynamic logic of regular programs,” *Journal of Computer and System Sciences*, vol. 18, no. 2, pp. 194–211, 1979.

[4] D. Harel, A. Pnueli, and J. Stavi, “Propositional dynamic logic of non-regular programs,” *Journal of Computer and System Sciences*, vol. 26, no. 2, pp. 222–243, 1983.

[5] G. De Giacomo and M. Lenzerini, “PDL-based framework for reasoning about actions,” in *AI*IA’95, Proceedings of the 4th Conference of the Italian Association for Artificial Intelligence*, vol. 992 of *Lecture Notes in Computer Science*, pp. 103–114, Springer-Verlag, 1995.

[6] D. Calvanese, G. De Giacomo, and M. Y. Vardi, “Reasoning about actions and planning in LTL action theories,” in *KR’02, Proceedings of the 8th International Conference on Principles of Knowledge Representation and Reasoning* (D. Fensel, F. Giunchiglia, D. L. McGuinness, and M.-A. Williams, eds.), (Toulouse), pp. 593–602, Morgan Kaufmann, April 2002.

[7] B. J. Ross, “Probabilistic pattern matching and the evolution of stochastic regular expressions,” *International Journal of Applied Intelligence*, vol. 13, pp. 285–300, Nov–Dec 2000.

[8] S. C. Kleene, “Representation of events in nerve nets and finite automata,” in *Automata Studies* (C. E. Shannon and J. McCarthy, eds.), pp. 3–41, Princeton, New Jersey: Princeton University Press, 1956.

[9] J. Hillston, *A Compositional Approach to Performance Modelling*, vol. 12 of *Distinguished Dissertations in Computer Science*. Cambridge University Press, 1996. ISBN 0 521 57189 8.

[10] S. Gilmore and J. Hillston, “Feature interaction in PEPA,” in *Process Algebra and Performance Modelling Workshop* (C. Priami, ed.), pp. 17–26, Università Degli Studi di Verona, Nice, September 1998.

[11] S. Gilmore, J. Hillston, and G. Clark, “Specifying performance measures for PEPA,” in *Proceedings of the 5th International AMAST Workshop on Real-Time and Probabilistic Systems*, vol. 1601 of *Lecture Notes in Computer Science*, (Bamberg), pp. 211–227, Springer-Verlag, 1999.

[12] G. Clark, *Techniques for the Construction and Analysis of Algebraic Performance Models*. PhD thesis, Department of Computer Science, University of Edinburgh, Edinburgh EH9 3JZ, UK, 1994. CST–107–94.

[13] J. T. Bradley, N. J. Dingle, S. T. Gilmore, and W. J. Knottenbelt, “Derivation of passage-time densities in PEPA models using ipc: the Imperial PEPA Compiler,” in *MASCOTS’03, Proceedings of the 11th IEEE/ACM International Symposium on Modeling, Analysis and Simulation of Computer and Telecommunications Systems* (G. Kotsis, ed.), (University of Central Florida), IEEE Computer Society Press, October 2003. To appear.

[14] R. Milner, *A Calculus of Communicating Systems*, vol. 92 of *Lecture Notes in Computer Science*. Springer-Verlag, 1980.

[15] M. Ribaudo, “Stochastic Petri net semantics for stochastic process algebras,” in *PNPM’95, Proceedings of 6th Int. Workshop on Petri Nets and Performance Models*, (Durham, North Carolina), pp. 148–157, October 1995.

[16] W. J. Knottenbelt, “Generalised Markovian analysis of timed transitions systems,” MSc thesis, University of Cape Town, South Africa, July 1996.

[17] J. T. Bradley, N. J. Dingle, P. G. Harrison, and W. J. Knottenbelt, “Distributed computation of passage time quantiles and transient state distributions in large semi-Markov models,” in *Performance Modelling, Evaluation and Optimization of Parallel and Distributed Systems*, (Nice), IEEE Computer Society Press, April 2003.

[18] J. T. Bradley, N. J. Dingle, W. J. Knottenbelt, and P. G. Harrison, “Performance queries on semi-Markov stochastic Petri nets with an extended Continuous Stochastic Logic,” in *PNPM’03, Proceedings of Petri Nets and Performance Models*, 2003. (To appear).

[19] N. J. Dingle, P. G. Harrison, and W. J. Knottenbelt, “Response time densities in Generalised Stochastic Petri Net models,” in *Proceedings of the 3rd International Workshop on Software and Performance (WOSP’2002)*, (Rome), pp. 46–54, July 2002.

[20] P. G. Harrison and W. J. Knottenbelt, “Passage-time distributions in large Markov chains,” in *Proceedings of ACM SIGMETRICS 2002* (M. Martonosi and E. d. S. e Silva, eds.), pp. 77–85, Marina Del Rey, USA, June 2002.

[21] W. J. Knottenbelt, *Parallel Performance Analysis of Large Markov Models*. PhD thesis, Imperial College, London, United Kingdom, February 2000.

[22] W. J. Knottenbelt and P. G. Harrison, “Distributed disk-based solution techniques for large Markov models,” in *NSMC’99, Proceedings of the 3rd Intl. Conference on the Numerical Solution of Markov Chains*, (Zaragoza), pp. 58–75, September 1999.

[23] W. Grassman, “Means and variances of time averages in Markovian environments,” *European Journal of Operational Research*, vol. 31, no. 1, pp. 132–139, 1987.

[24] A. Reibman and K. Trivedi, “Numerical transient analysis of Markov models,” *Computers and Operations Research*, vol. 15, no. 1, pp. 19–36, 1988.

[25] N. J. Dingle, P. G. Harrison, and W. J. Knottenbelt, “Uniformization and hypergraph partitioning for the distributed computation of response time densities in very large Markov models,” *Submitted to the Journal of Parallel and Distributed Computing*, 2002.

[26] N. J. Dingle, W. J. Knottenbelt, and P. G. Harrison, “HYDRA: HYpergraph-based Distributed Response-time Analyser,” in *PDPTA’03, Proceedings of the 2003 International Conference on Parallel and Distributed Processing Techniques and Applications* (H. R. Arabnia and Y. Man, eds.), vol. 1, (Las Vegas, NV), pp. 215–219, June 2003.

[27] B. Melamed and M. Yadin, “Randomization procedures in the computation of cumulative-time distributions over discrete state Markov processes,” *Operations Research*, vol. 32, pp. 926–944, July–August 1984.

[28] J. K. Muppala and K. S. Trivedi, “Numerical transient analysis of finite Markovian queueing systems,” in *Queueing and Related Models* (U. N. Bhat and I. V. Basawa, eds.), pp. 262–284, Oxford University Press, 1992.

[29] A. Carmona, L. Domingo, R. Macau, R. Puigjaner, and F. Rojo, “Performance experiences of the Barcelona Olympic games computer system,” in *Computer Performance Evaluation, Modeling Techniques and Tools* (G. Haring and G. Kotsis, eds.), vol. 794 of *Lecture Notes in Computer Science*, pp. 52–75, Springer, 1994.