Performance Bound Hierarchies for Queueing Networks

DEREK L. EAGER and KENNETH C. SEVCIK
University of Toronto

In applications of queueing network models to computer system performance prediction, the computational effort required to obtain an exact equilibrium solution of a model may not be justified by the accuracy actually required. In these cases, there is a need for approximation or bounding techniques that can provide the necessary information with less computational effort. This paper presents a new technique that yields performance bounds for single-class separable queueing networks consisting of fixed-rate and delay service centers. Unlike previous approximation or bounding techniques, there is a smooth trade-off between computational effort and accuracy. Any level of accuracy (including the exact solution) can be guaranteed by investing the necessary computational effort. Performance bounds that are sufficiently tight for most practical purposes may be obtained with a fraction of the effort required for the exact solution. Since bounds are produced, as opposed to approximations, guarantees about the accuracy of a model solution can be provided.

Categories and Subject Descriptors: D.4.8 [**Operating Systems**]: Performance—*modeling and prediction; operational analysis; stochastic analysis*

General Terms: Measurement, Performance

Additional Key Words and Phrases: Asymptotic analysis, balanced job bounds, bounding analysis, product form networks

1. INTRODUCTION

Queueing network models are widely used as aids in computer system performance prediction. As a result of the infeasibility of computing the exact equilibrium solution of the most general form of queueing networks, a restricted class known as separable networks [12] has been used in practice. In this paper, separable networks with a single customer class and with fixed-rate and delay service centers are considered. Fixed-rate centers model system components that have a single fixed rate at which customers are processed. Delay centers model system

The research described in this paper was supported by the Natural Sciences and Engineering Research Council of Canada.

This paper was originally presented at the 1982 ACM/SIGMETRICS Conference on Measurement and Modeling of Computer Systems (Seattle, Washington, August 30–September 1, 1982). The paper was selected for publication in the Systems Modeling and Performance Evaluation department of the *Communications of the ACM*. In the interim, however, TOCS began publication. The authors kindly agreed to publish the paper in this journal.

Authors' address: Computer Systems Research Group, Sandford Fleming Building, University of Toronto, Toronto, Ontario M5S 1A4, CANADA.

components at which queueing does not occur (such as at a collection of terminals).

The solution of separable networks requires substantially less computation than does the solution of those that are not separable. Even with separable networks, however, it may be the case that the accuracy required does not justify the computational effort of an exact solution. In these cases, approximation or bounding techniques can provide the necessary information with substantially less computational effort.

Bounding can be used in conjunction with an approximation technique, or it can be used independently. In the former case, bounding provides a simple means of determining a nontrivial upper bound on the error in the approximation, an ability not found in most current approximation techniques. If the bounds are sufficiently tight, or the accuracy requirements of the model solution are sufficiently minimal, bounds can be used independently. In this case, bounds are preferable to approximations since they provide more reliable information about the location of the exact solution.

To be useful, a bounding technique should be analytically simple, effective with a considerable reduction in computational effort, and reasonably accurate. Ideally, there should be a trade-off between computational effort and accuracy.

Asymptotic bound analysis (ABA) [4, 6, 9] produces bounds on mean system residence time and system throughput by considering the extremes of system behavior: either no queueing delay takes place, or one or more centers operate at capacity. ABA bounds are, in general, very loose. They have the advantages of applying to a larger class of queueing networks than that considered here, and of being analytically and computationally simple.

Balanced job bounds (BJBs) [15] are derived by considering related queueing networks whose performance measures bound those of the original network. These related networks have the property that all customers exhibit balanced resource usage, a property that produces computationally and analytically simple results. The bounds produced are bounds on mean system residence time and system throughput. Bounds on individual center performance measures that cannot be obtained directly from system measures have also been derived, in part by using ABA and BJB analysis [5].

In Figure 1, ABA bounds and BJBs are shown for a sample model. The model has fifty fixed-rate centers with loadings as follows: 1 center at 20/417, 2 at 19/417, 5 at 18/417, 5 at 15/417, 5 at 10/417, 8 at 7/417, 8 at 5/417, 8 at 4/417, and 8 at 2/417. The loading of a center is defined as the product of the mean service time per visit and the average number of visits per customer; for simplicity of notation, the unit of time measurement used in this paper is such that the sum of the fixed-rate center loadings equals one. A model with no delay centers has been chosen since the BJB technique does not efficiently treat models with this type of center. Another major disadvantage of BJBs (and ABA bounds) is that there is no trade-off between computational effort and accuracy. If the accuracy provided by BJBs or ABA bounds is insufficient, either the technique must be abandoned or the model must be decomposed. ABA bounds, BJBs, or the bounds that will be presented here can be computed for a submodel, and the result can be combined with characteristics of the remaining centers to obtain an overall solution of the model. In general, this solution will be more accurate than that

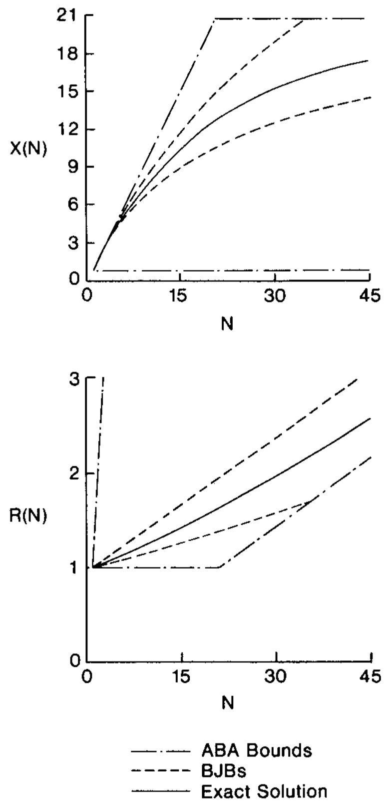

Fig. 1. Asymptotic and balanced job bounds on system throughput (X(N)), and mean system residence time (R(N)), as functions of the number of customers (N).

obtained by applying the respective bounding technique directly to the entire model.

In this paper, *performance bound hierarchies* (PBHs) are developed. A PBH consists of a hierarchy of successively more accurate upper or lower bounds on a system performance measure, with the exact solution as its limit. By utilizing a pair of PBHs, one on each side of the exact solution, any required accuracy level can be attained with a corresponding level of computational effort. The technique is also analytically simple, being based on the Mean Value Analysis (MVA) solution algorithm for separable networks [10]. In fact, PBHs can be viewed as a link between the exact solution algorithm and the single-class versions of approximate MVA methods such as those in [1, 2, 13, 14].

2. PERFORMANCE BOUND HIERARCHIES

2.1 The PBH Approach

For a fixed-rate center k, the central equation in the MVA solution algorithm is

$$R_k(N) = L_k[1 + \bar{n}_k(N - 1)], \tag{1}$$

where $R_k(N)$ denotes the mean residence time of a customer at center k (in all of its visits to center k) when there are N customers in the network, $\bar{n}_k(N - 1)$ is the mean queue length at center k when there are $N - 1$ customers in the network, and L_k is the loading of center k [10]. For computational purposes, all of the delay centers in a model may be aggregated into a single delay center with a loading equal to the sum of the individual loadings [11]. The loading of an aggregate delay center will be denoted by Z. Since there is no queueing at a delay center, the mean residence time at an aggregate delay center will also be equal to Z. (In applying PBHs to a system with no delay centers, the value of Z is simply zero in the following development.)

Indexing the fixed-rate centers of a queueing network model as centers 1 through K, and the delay center as center $K + 1$, the mean system residence time $R(N)$ will be defined as

$$R(N) = \sum_{k=1}^{K} R_k(N). \tag{2}$$

As motivated by the usual application of delay centers to the modeling of terminal systems, $R(N)$ excludes the residence time at the delay center. Little's equation [8] and the Forced Flow Law [4] yield, for each fixed-rate center k,

$$\bar{n}_k(N-1) = \frac{R_k(N-1)}{Z + R(N-1)}(N-1). \tag{3}$$

Finally, eq. (3) may be substituted into eq. (1) to produce

$$R_k(N) = L_k\left[1 + \frac{R_k(N-1)}{Z + R(N-1)}(N-1)\right]. \tag{4}$$

Equations (2) and (4) capture the relationship between mean residence times with some given population and those with one fewer customer. The exact MVA technique uses eqs. (2) and (4) in iterating from a population level of one up to the population level for which performance statistics are desired. Approximate MVA techniques utilize an approximation of this exact relationship; for example, in [13], it is assumed that $R_k(N-1)/(Z + R(N-1)) = R_k(N)/(Z + R(N))$. Approximate MVA iterations are usually confined to adjacent population levels, with the iteration continuing until the inexact approximating relationship produces consistent level N and $N - 1$ performance measures (by eqs. (2) and (4)). Note that since approximate MVA is based in part on an inexact relationship, any convergence is not, in general, to the exact solution.

In contrast, the PBHs that are presented here form a sequence of successively tighter optimistic and pessimistic bound pairs that converge to the exact solution. The members of the pessimistic (or optimistic) PBH are stated as upper (or lower) bounds on mean system residence time, but corresponding bounds on system throughput and center utilizations can be obtained in a straightforward manner by using Little's equation and the Forced Flow Law. PBHs are, like exact MVA, based on the exact relationship between the performance measures of adjacent population levels. Like approximate MVA, PBHs involve a variable number of iterations; in particular, a level i PBH bound may be computed with i iterations using eqs. (2) and (4). Unlike approximate MVA, since PBHs are based only on exact relationships, any level of accuracy can be attained by performing a sufficient number of iterations (although the number of iterations performed must be chosen in advance of the computation).

The defining equations for the level i mean system residence time bound $R^{(i)}(N)(i \geq 1)$ in a PBH are

$$R_k^{(i)}(N) = L_k\left[1 + \frac{R_k^{(i-1)}(N-1)}{Z + R^{(i-1)}(N-1)}(N-1)\right] \tag{5}$$

and

$$R^{(i)}(N) = \sum_{k=1}^{K} R_k^{(i)}(N). \tag{6}$$

In these equations, the level i bound with N customers is derived from the level $i-1$ bound with $N-1$ customers. Note that the level $i-1$ bound cannot be used to derive a tighter performance bound for the same population level, but only for a population level with one additional customer. Assuming that N is greater than i, to calculate the level i bound with N customers it is necessary first to calculate the level 1 bound with $N-i+1$ customers (using some choice of $R_k^{(0)}(N-i)$ values), then to calculate the level 2 bound with $N-i+2$ customers, until finally, continuing in this manner, the level i bound with N customers is calculated. (For $N \leq i$, the bound computation reduces to exact MVA.)

Whatever $R_k^{(0)}(N-i)$ values are chosen, $R_k^{(i)}(N)$ is exact for $N \leq i$, and as i increases, increasingly accurate approximations are produced, in general, for $N > i$. However, the choice of the $R_k^{(0)}(N-i)$ values determines whether the hierarchy members are optimistic bounds, pessimistic bounds, or merely approximations. In the next two sections, initialization choices that produce the desired bound hierarchies will be presented. It will be seen that the iterations of eqs. (5) and (6) can be carried out symbolically to yield expressions that are often much simpler analytically and computationally than the iterations themselves. A PBH computation algorithm that was derived from this symbolic iteration process is presented in Section 3.

2.2 The Optimistic Hierarchy

Each optimistic hierarchy member will be denoted with a subscript *opt*. The initialization required is

$$R_{k\,\mathrm{opt}}^{(0)}(N) = \frac{1}{K}\max[NL_b - Z, 1], \tag{7}$$

where b is the index of a bottleneck center (a fixed-rate center with a loading greater than or equal to that of any other fixed-rate center). The optimistic initialization yields

$$R_{\mathrm{opt}}^{(0)}(N) = \max[NL_b - Z, 1], \tag{8}$$

which corresponds exactly to the optimistic ABA bound (remembering that the time unit chosen here is such that loadings at the fixed-rate centers sum to unity). The individual center residence time choices spread this residence time evenly among the K fixed-rate centers; this division is optimistic for highly loaded centers and pessimistic for lightly loaded centers. A proof that this choice yields a hierarchy of optimistic bounds is given in Appendix A.

Upon denoting $\max[NL_b - Z, 1]$ by $a(N)$, eqs. (5), (7), and (8) yield

$$R_{k\,\mathrm{opt}}^{(1)}(N) = L_k\left[1 + \frac{1}{K}\left(\frac{a(N-1)}{Z + a(N-1)}\right)(N-1)\right]. \tag{9}$$

From eq. (6),

$$R_{\mathrm{opt}}^{(1)}(N) = 1 + \frac{1}{K}\left(\frac{a(N-1)}{Z + a(N-1)}\right)(N-1). \tag{10}$$

When $Z = 0$, corresponding to a model with no delay centers, this bound is exactly the BJB optimistic bound.

The level 2 optimistic PBH member is given by

$$R_{\mathrm{opt}}^{(2)}(N) = 1 + S\left(\frac{1 + \frac{1}{K}\left(\frac{a(N-2)}{Z + a(N-2)}\right)(N-2)}{Z + 1 + \frac{1}{K}\left(\frac{a(N-2)}{Z + a(N-2)}\right)(N-2)}\right)(N-1), \tag{11}$$

where $S = \sum_{k=1}^{K} L_k^2$. When $Z = 0$, this bound reduces to

$$R_{\mathrm{opt}}^{(2)}(N) = 1 + S(N-1). \tag{12}$$

Any higher level optimistic bound can be similarly expressed analytically. In practice, the optimistic asymptotic bound $a(N)$ (i.e., $R_{\mathrm{opt}}^{(0)}(N)$) is used in conjunction with each optimistic PBH bound, since each PBH bound (of a level higher than zero) will eventually cross the asymptotic bound as N increases. When the asymptotic bound is greater than the PBH bound, the asymptotic bound is utilized. Note that if the asymptotic bound is not used in this manner, the optimistic hierarchy is not strictly nested, in that for some values of N a lower level bound (the level zero bound, for example) may be tighter than a higher level bound. However, there is evidence suggesting that when each bound is used in conjunction with the asymptotic bound, the hierarchy is strictly nested.

These lower bounds on mean system residence time may also be expressed as upper bounds on system throughput or center utilizations.

2.3 The Pessimistic Hierarchy

Each pessimistic hierarchy member will be denoted with a subscript *pess*. The initialization required is

$$R_{k\,\mathrm{pess}}^{(0)}(N) = \begin{cases} N & \text{for } k = b \\ 0 & \text{for } k \neq b. \end{cases} \tag{13}$$

The pessimistic initialization yields

$$R_{\mathrm{pess}}^{(0)}(N) = N, \tag{14}$$

which corresponds to the trivial ABA pessimistic bound. The individual center residence time choice is pessimistic for the bottleneck center and optimistic for all other centers. A proof that this choice guarantees a hierarchy of pessimistic bounds is given in Appendix B. This proof also shows that the hierarchy is nested, in that the level i bound is guaranteed to be at least as tight as the level $i - 1$ bound.

Equations (5), (13), and (14) yield

$$R_{k\,\text{pess}}^{(1)}(N) = \begin{cases} L_b\left(1 + \left(\dfrac{N-1}{Z+N-1}\right)(N-1)\right) & \text{for } k = b \\ L_k & \text{for } k \neq b. \end{cases} \tag{15}$$

From eq. (6),

$$R_{\text{pess}}^{(1)}(N) = 1 + L_b\left(\frac{N-1}{Z+N-1}\right)(N-1). \tag{16}$$

When $Z = 0$, this bound corresponds exactly to the BJB pessimistic bound.

The level 2 pessimistic PBH bound is

$$R_{\text{pess}}^{(2)}(N) = 1 + \left(\frac{S + L_b^2 \dfrac{(N-2)^2}{(Z+N-2)}}{Z + 1 + L_b \dfrac{(N-2)^2}{(Z+N-2)}}\right)(N-1). \tag{17}$$

If $Z = 0$, this reduces to

$$R_{\text{pess}}^{(2)}(N) = 1 + \left(\frac{S + L_b^2(N-2)}{1 + L_b(N-2)}\right)(N-1). \tag{18}$$

Any higher level pessimistic bound can be similarly derived. These upper bounds on mean system residence time may also be expressed as lower bounds on system throughput or center utilizations. Optimistic and pessimistic PBH members are shown in Figure 2 for the model of Figure 1. Figure 3 shows PBH bounds for the model as modified by the addition of an aggregate delay center, with loading 4000/417.

3. HIERARCHY PROPERTIES

3.1 Convergence

Given level i optimistic and pessimistic PBH bounds, the width of the interval defined by these bounds can be explicitly calculated. If the tightness of the bounds is not sufficient, higher level PBH members can be utilized. (If the bound tightness indicates that the level $i + j$ bounds might be appropriate, for example, it would be necessary to start over from $R_k^{(0)}(N - i - j)$ values.) To minimize the computational expense, it is useful to have information about the tightness of a pair of bounds prior to computing those bounds. In this section, some preliminary results on the speed of convergence of the optimistic and pessimistic hierarchies to the exact solution are presented.

It is first necessary to choose a measure of bound tightness. The measure chosen here is the magnitude of the percent relative error in the optimal (with regard to minimizing the worst-case relative error) bound-based approximation. In the case of mean system residence time, this approximation can be shown to be

$$\frac{2R_{\text{pess}}^{(i)}(N)R_{\text{opt}}^{(i)}(N)}{R_{\text{pess}}^{(i)}(N) + R_{\text{opt}}^{(i)}(N)}, \tag{19}$$

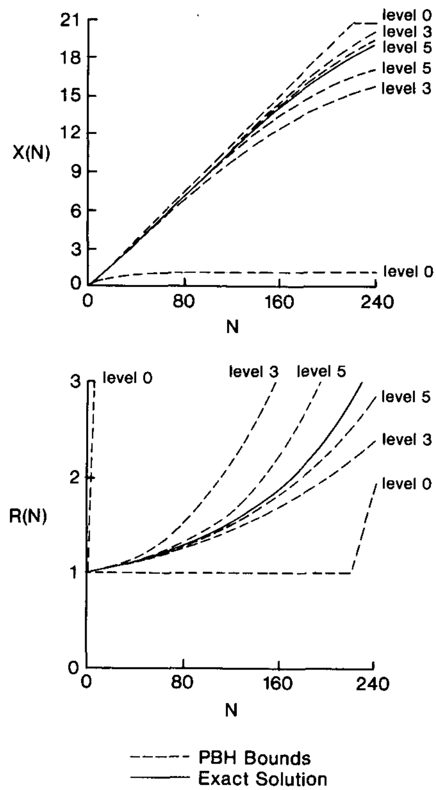

Fig. 2. PBH bounds on system throughput and mean system residence time as functions of the number of customers.

which has an associated worst-case (or maximal) relative error magnitude of

$$\frac{R_{\text{pess}}^{(i)}(N) - R_{\text{opt}}^{(i)}(N)}{R_{\text{pess}}^{(i)}(N) + R_{\text{opt}}^{(i)}(N)} \times 100\%. \tag{20}$$

This maximal error would be attained only if one of the bounds coincided with the exact solution. Similar expressions hold for other performance measures such as system throughput. (However, the optimal throughput approximation and the optimal residence time approximation are not, in general, related by Little's equation.)

The results that have been attained to date concern the supremum of the maximal errors associated with the level i bounds, considering all possible queueing networks of the type being treated and all possible population levels. This supremum is, in practice, overly pessimistic, as it has been found to occur as the number of centers and customers in the model tend to infinity. In fact, experimental results and analytical verification for small i have indicated that the supremum occurs as $L_b \to 0$ and $L_k/L_b \to 0$ for all $k \neq b$ (which means that the number of centers must tend to infinity, since $\sum_{k=1}^{K} L_k = 1$). Similar evidence shows that the population level that yields the supremum is that at which the level i optimistic bound intersects the asymptotic bound (past which point the asymptotic bound is used); this population level can be shown to tend to $(Z + 1)/L_b$ as $L_b \to 0$ and $L_k/L_b \to 0$.

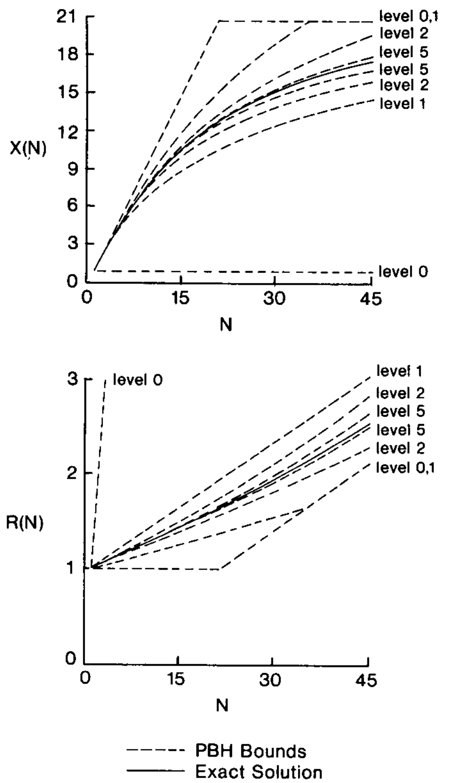

Fig. 3. PBH bounds on system throughput and mean system residence time in a system including a delay center as functions of the number of customers.

For this limit, each of the following limiting values hold:

$$
\begin{array}{lll}
R_{k\,\mathrm{opt}}^{(i)}(N) \rightarrow L_k & R_{k\,\mathrm{pess}}^{(i)}(N) \rightarrow L_k & \text{for } k \neq b \\[6pt]
R_{b\,\mathrm{opt}}^{(i)}(N) \rightarrow iL_b & R_{b\,\mathrm{pess}}^{(i)}(N) \rightarrow \dfrac{Z+1}{i} & \\[6pt]
R_{\mathrm{opt}}^{(i)}(N) \rightarrow 1 & R_{\mathrm{pess}}^{(i)}(N) \rightarrow 1 + \dfrac{Z+1}{i}. &
\end{array}
\tag{21}
$$

Also, since $R_{\mathrm{pess}}^{(i)}(N) \geq R(N)$ for all i, $R(N)$ must tend to 1. Since $R_{\mathrm{opt}}^{(i)}(N)$ also tends to 1, the maximal error at this limit must be achieved, implying that the supremum of the maximal relative errors is also the supremum of the actual relative errors. It then follows from eq. (20) and its equivalent for system throughput that the supremum of the relative errors in mean system residence time when using the level i bounds is $(Z+1)/(2i+Z+1) \times 100\%$, while that for system throughput is $1/(2i+1) \times 100\%$. These results are worst-case results and therefore only provide an upper bound on the error encountered in practice; in Figure 4, the actual and maximal relative errors for the model of Figure 3 and a population level of 240 customers are shown as functions of i. (Note that although the *actual* error magnitude for mean system residence time is not monotonically decreasing with i in this case, the *bound* on the magnitude is.)

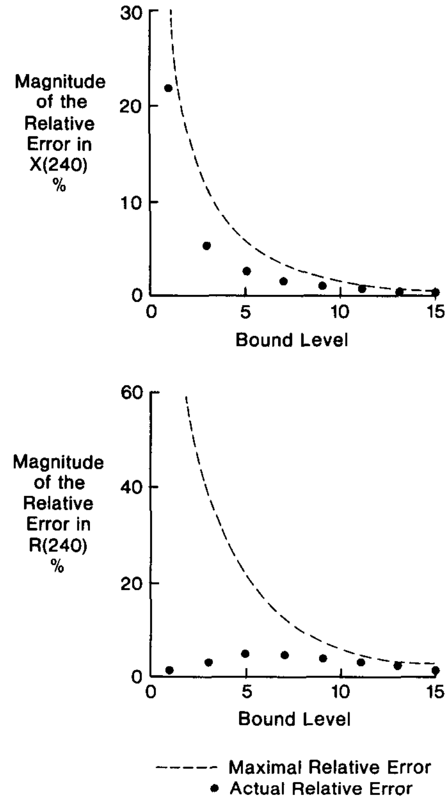

Fig. 4. Error measures of approximations based on PBHs as functions of the bound level utilized.

3.2 Computational Efficiency

The straightforward method of calculating PBH bounds is to make direct use of the defining iterative equations. With this approach, each level of iteration requires approximately $4K$ arithmetic operations (the same as one iteration of MVA). Thus, the level i member of the optimistic or pessimistic PBH can be obtained in $4Ki$ operations. If i is chosen to be N, the exact solution is obtained in $4KN$ operations (as with exact MVA). When the approximate forms of MVA are used, convergence often requires ten iterations or more. Thus, with the same computational effort, at least the fifth level members of the optimistic and pessimistic PBH can be calculated.

One way to improve the computational efficiency of PBH bounds is to utilize tighter hierarchy initializations. For example, a tighter pessimistic hierarchy initialization is obtained by using eqs. (15) and (16) (from the level 1 hierarchy member) in place of eqs. (13) and (14). Essentially, one MVA-like iteration is replaced by a more computationally complex initialization.

However, PBH bounds can be calculated still more efficiently by not performing any MVA-like iterations at all. After calculating $L^{(p)} = \sum_{k=1}^{K} L_k^p$ for $p = 1, 2, \ldots, i$ (which requires about $2Ki$ operations), the level i member of a PBH evaluated at N is given by $f(i, 1, N)$, where

$$f(j, p, N) = L^{(p)} + \frac{f(j-1, p+1, N-1)}{Z + f(j-1, 1, N-1)} (N-1), \tag{22}$$

and the boundary conditions are

$$f(0, p, N) = NL_b^{p-1} \tag{23}$$

for the pessimistic hierarchy, and

$$f(0, p, N) = \frac{L^{(p-1)}}{K} \max[NL_b - Z, 1] \tag{24}$$

for the optimistic hierarchy. These equations can be verified using eqs. (5), (6), (7), and (13), noting that $f(j, p, N) = \sum_{k=1}^{K} L_k^{p-1} R_k^{(j)}(N)$.

The evaluation of $f(i, 1, N)$ requires approximately $i^2/2$ function evaluations, each involving about four operations. (Thus, the computational cost of this formulation is related to that of both the solution algorithm based on Polya's theory of enumeration [7] and CCNC [3].) Whenever i is chosen to be not much larger than K, this formulation is more efficient than the straightforward method based on MVA-like iterations. The results of the previous section and experimental results suggest that, in practice, i would rarely be chosen to be significantly larger than K.

4. CONCLUSIONS

The *performance bound hierarchy* technique presented in this paper provides links among Mean Value Analysis, approximate solution algorithms based on Mean Value Analysis, and prior bounding techniques. The level 0 and 1 PBH members correspond, respectively, to ABA and BJB bounds. For a system with N customers, the exact solution is attained by the level N PBH member. With comparable computational cost, PBHs provide guaranteed bounds on performance and convergence toward the exact solution (although each bound requires a separate computation), while the approximate MVA-based approaches produce results without any bound on the maximum possible error, and converge, in general, to an answer other than the exact solution. In systems with a very large number of devices and customers, the computational advantage of using an estimate based on PBH bounds rather than an exact MVA solution can be substantial.

Performance bound hierarchies permit a smooth trade-off between accuracy and computational cost. For a stated constraint on accuracy, the bound level can be chosen to minimize the computational cost. Alternatively, for a given computational budget, the bound level can be chosen such that the most accurate possible answer is obtained.

In the context of multiple customer classes, the likelihood is greater that the computational cost (in both space and time) of an exact solution may be unacceptably high. We are currently developing PBHs for the case of multiple classes.

APPENDIX A

For simplicity of notation, $\bar{n}_k(N)$ and $\bar{n}_{k\,\mathrm{opt}}^{(i)}(N)$ will be denoted as n_k and $n_k^{(i)}$, respectively, in the following proof that each bound in the optimistic PBH is indeed an optimistic bound. Also, the fixed-rate centers are assumed to be indexed such that $L_k \geq L_{k+1}$ for $1 \leq k \leq K - 1$. It is first necessary to prove the following lemma.

LEMMA. *Suppose that for some $i \geq 0$,*

$$\frac{n_j^{(i)}}{\sum_{k=1}^{j} n_k^{(i)}} \geq \frac{n_j}{\sum_{k=1}^{j} n_k} \qquad 2 \leq j \leq K+1. \tag{A-1}$$

Then

$$\sum_{k=1}^{j} n_k \geq \sum_{k=1}^{j} n_k^{(i)} \qquad 2 \leq j \leq K+1, \tag{A-2}$$

and

$$\frac{\sum_{k=1}^{j} L_k n_k}{\sum_{k=1}^{j} n_k} \geq \frac{\sum_{k=1}^{j} L_k n_k^{(i)}}{\sum_{k=1}^{j} n_k^{(i)}} \qquad 2 \leq j \leq K. \tag{A-3}$$

PROOF OF RELATION A-2. The proof is by downwards induction on j. For $j = K+1$, $\sum_{k=1}^{K+1} n_k = N = \sum_{k=1}^{K+1} n_k^{(i)}$, establishing the induction basis.

Assume that relation A-2 holds for $j+1$, where $2 \leq j \leq K$, and consider the relation for j. Relation A-1 applied for $j+1$ yields

$$\frac{n_{j+1}^{(i)}}{\sum_{k=1}^{j+1} n_k^{(i)}} \geq \frac{n_{j+1}}{\sum_{k=1}^{j+1} n_k} \Rightarrow 1 - \frac{n_{j+1}}{\sum_{k=1}^{j+1} n_k} \geq 1 - \frac{n_{j+1}^{(i)}}{\sum_{k=1}^{j+1} n_k^{(i)}} \Rightarrow \frac{\sum_{k=1}^{j} n_k}{\sum_{k=1}^{j+1} n_k} \geq \frac{\sum_{k=1}^{j} n_k^{(i)}}{\sum_{k=1}^{j+1} n_k^{(i)}}.$$

Since $\sum_{k=1}^{j+1} n_k \geq \sum_{k=1}^{j+1} n_k^{(i)}$ by the inductive hypothesis, it must be the case that $\sum_{k=1}^{j} n_k \geq \sum_{k=1}^{j} n_k^{(i)}$; therefore, relation A-2 is established for j. By induction, relation A-2 is established for $2 \leq j \leq K+1$. □

PROOF OF RELATION A-3. The proof is by induction on j. For $j = 2$, relation A-3 becomes

$$\frac{L_1 n_1 + L_2 n_2}{n_1 + n_2} \geq \frac{L_1 n_1^{(i)} + L_2 n_2^{(i)}}{n_1^{(i)} + n_2^{(i)}}.$$

Since $L_1 \geq L_2$ and

$$\frac{n_2^{(i)}}{n_1^{(i)} + n_2^{(i)}} \geq \frac{n_2}{n_1 + n_2}$$

by relation A-1, this relation will hold if

$$\frac{L_1 n_1 + L_1 n_2}{n_1 + n_2} \geq \frac{L_1 n_1^{(i)} + L_1 n_2^{(i)}}{n_1^{(i)} + n_2^{(i)}}.$$

As this expression is an equality, the induction basis is shown.

Assume that relation A-3 holds for j, where $2 \leq j < K$, and consider the relation for $j+1$. First, note that

$$\frac{\sum_{k=1}^{j+1} L_k n_k^{(i)}}{\sum_{k=1}^{j+1} n_k^{(i)}}$$

is maximized for minimized $n_{j+1}^{(i)}$, since for $n_{j+1}^{(i)} \geq \varepsilon \geq 0$,

$$\frac{\sum_{k=1}^{j} L_k n_k^{(i)} + L_{j+1}(n_{j+1}^{(i)} - \varepsilon)}{\sum_{k=1}^{j+1} n_k^{(i)} - \varepsilon} < \frac{\sum_{k=1}^{j+1} L_k n_k^{(i)}}{\sum_{k=1}^{j+1} n_k^{(i)}} \Rightarrow L_{j+1} \sum_{k=1}^{j+1} n_k^{(i)} > \sum_{k=1}^{j+1} L_k n_k^{(i)}.$$

Given the assumed center indexing, this last relation is a contradiction. Therefore, relation A-3 need only be shown under the assumption that $n_{j+1}^{(i)}$ is minimized, which by relation A-1 corresponds to

$$\frac{n_{j+1}^{(i)}}{\sum_{k=1}^{j+1} n_k^{(i)}} = \frac{n_{j+1}}{\sum_{k=1}^{j+1} n_k}.$$

Starting with the inductive hypothesis, it follows that

$$\begin{aligned}\frac{\sum_{k=1}^{j} L_k n_k}{\sum_{k=1}^{j} n_k} \geq \frac{\sum_{k=1}^{j} L_k n_k^{(i)}}{\sum_{k=1}^{j} n_k^{(i)}} \Rightarrow \left(1 - \frac{n_{j+1}}{\sum_{k=1}^{j+1} n_k}\right) \frac{\sum_{k=1}^{j} L_k n_k}{\sum_{k=1}^{j} n_k} \\ \geq \left(1 - \frac{n_{j+1}^{(i)}}{\sum_{k=1}^{j+1} n_k^{(i)}}\right) \frac{\sum_{k=1}^{j} L_k n_k^{(i)}}{\sum_{k=1}^{j} n_k^{(i)}} \\ \Rightarrow \frac{\sum_{k=1}^{j} L_k n_k}{\sum_{k=1}^{j+1} n_k} \geq \frac{\sum_{k=1}^{j} L_k n_k^{(i)}}{\sum_{k=1}^{j+1} n_k^{(i)}}.\end{aligned}$$

Since

$$\frac{L_{j+1} n_{j+1}}{\sum_{k=1}^{j+1} n_k} = \frac{L_{j+1} n_{j+1}^{(i)}}{\sum_{k=1}^{j+1} n_k^{(i)}}$$

by assumption, this last relation establishes relation A-3 for $j + 1$. By induction, relation A-3 is established for $2 \leq j \leq K$, which establishes the lemma. □

The main result can now be shown.

THEOREM. *For all $i \geq 0$, $N \geq 1$,*

$$R(N) \geq R_{opt}^{(i)}(N), \tag{A-4}$$

and

$$\frac{n_j^{(i)}}{\sum_{k=1}^{j} n_k^{(i)}} \geq \frac{n_j}{\sum_{k=1}^{j} n_k} \quad \textit{for} \quad 2 \leq j \leq K + 1. \tag{A-5}$$

PROOF. The proof is by induction on i.

Consider $i = 0$. By the correctness of ABA,

$$R(N) \geq \max[NL_b - Z, 1] = R_{\text{opt}}^{(0)}(N).$$

Also, for $2 \leq j \leq K$, relation A-5 becomes

$$\frac{n_j^{(0)}}{\sum_{k=1}^{j} n_k^{(0)}} = \frac{\dfrac{a(N)}{Z + a(N)} \dfrac{1}{K} N}{j \dfrac{a(N)}{Z + a(N)} \dfrac{1}{K} N} = \frac{1}{j} \geq \frac{n_j}{\sum_{k=1}^{j} n_k}.$$

This last relation holds since $L_1 \geq L_2 \geq \cdots \geq L_K$ implies that $n_1 \geq n_2 \geq \cdots \geq n_K$. Finally, for $j = K + 1$, relation A-5 becomes

$$\frac{n_{K+1}^{(0)}}{\sum_{k=1}^{K+1} n_k^{(0)}} = \frac{\dfrac{Z}{Z + a(N)} N}{N} = \frac{Z}{Z + a(N)} \geq \frac{Z}{Z + R(N)} = \frac{n_{K+1}}{\sum_{k=1}^{K+1} n_k}.$$

The induction basis is now established.

Assume that the theorem holds for i, and consider the theorem for $i + 1$ ($i \geq 0$). For $N = 1$, relations A-4 and A-5 are equalities; therefore, only $N + 1$ ($N \geq 1$) need be considered. The preceding lemma applied with the inductive hypothesis yields

$$\frac{\sum_{k=1}^{K} L_k n_k}{\sum_{k=1}^{K} n_k} \geq \frac{\sum_{k=1}^{K} L_k n_k^{(i)}}{\sum_{k=1}^{K} n_k^{(i)}},$$

and

$$\sum_{k=1}^{K} n_k \geq \sum_{k=1}^{K} n_k^{(i)}.$$

Applying the latter relation to the former yields

$$\sum_{k=1}^{K} L_k n_k \geq \sum_{k=1}^{K} L_k n_k^{(i)},$$

implying that

$$R(N+1) = \sum_{k=1}^{K} L_k(1 + n_k) \geq \sum_{k=1}^{K} L_k(1 + n_k^{(i)}) = R_{\text{opt}}^{(i+1)}(N+1).$$

This establishes relation A-4 for $i + 1$.

Now, assume that relation A-5 does not hold for $i + 1$ for some j, where $2 \leq j \leq K$. The definitions of $\bar{n}_k(N+1)$ and $\bar{n}_{k\,\text{opt}}^{(i+1)}(N+1)$ then yield

$$\frac{\dfrac{L_j(1+n_j^{(i)})}{Z + R_{\text{opt}}^{(i+1)}(N+1)}(N+1)}{\sum_{k=1}^{j} \dfrac{L_k(1+n_k^{(i)})}{Z + R_{\text{opt}}^{(i+1)}(N+1)}(N+1)} < \frac{\dfrac{L_j(1+n_j)}{Z + R(N+1)}(N+1)}{\sum_{k=1}^{j} \dfrac{L_k(1+n_k)}{Z + R(N+1)}(N+1)}$$

$$\Rightarrow \frac{1 + n_j^{(i)}}{\sum_{k=1}^{j} L_k(1 + n_k^{(i)})} < \frac{1 + n_j}{\sum_{k=1}^{j} L_k(1 + n_k)}.$$

From the inductive hypothesis and the previous lemma, this yields

$$\frac{1 + n_j \dfrac{\sum_{k=1}^{j} n_k^{(i)}}{\sum_{k=1}^{j} n_k}}{\sum_{k=1}^{j} L_k + \dfrac{\sum_{k=1}^{j} n_k^{(i)}}{\sum_{k=1}^{j} n_k} \sum_{k=1}^{j} L_k n_k} < \frac{1 + n_j}{\sum_{k=1}^{j} L_k(1 + n_k)}.$$

Multiplying the numerator and denominator of the left-hand side by $\sum_{k=1}^{j} n_k$, cross-multiplying, and then simplifying produces

$$\sum_{k=1}^{j} L_k n_k \left(\sum_{k=1}^{j} n_k - \sum_{k=1}^{j} n_k^{(i)} \right) < n_j \sum_{k=1}^{j} L_k \left(\sum_{k=1}^{j} n_k - \sum_{k=1}^{j} n_k^{(i)} \right).$$

From the preceding lemma and the inductive hypothesis, $\sum_{k=1}^{j} n_k - \sum_{k=1}^{j} n_k^{(i)} \geq 0$; this fact applied to the preceding inequality yields

$$\sum_{k=1}^{j} L_k n_k < n_j \sum_{k=1}^{j} L_k.$$

Since $n_1 \geq n_2 \geq \cdots \geq n_K$, this last inequality is a contradiction. Therefore, relation A-5 is established for $2 \leq j \leq K$.

For $j = K + 1$, relation A-5 becomes

$$\frac{\dfrac{Z}{Z + R_{\mathrm{opt}}^{(i+1)}(N+1)}(N+1)}{N+1} \geq \frac{\dfrac{Z}{Z + R(N+1)}(N+1)}{N+1}$$

$$\Leftrightarrow \frac{Z}{Z + R_{\mathrm{opt}}^{(i+1)}(N+1)} \geq \frac{Z}{Z + R(N+1)}.$$

As it has been established that $R(N+1) \geq R_{\mathrm{opt}}^{(i+1)}(N+1)$, this last relation holds. Relation A-5 and the theorem are therefore established for $i + 1$. By induction, the theorem is established for $i \geq 0$. □

APPENDIX B

In the following proof that the pessimistic PBH forms a nested hierarchy of pessimistic bounds, $\bar{n}_{k\,\mathrm{pess}}^{(i)}(N)$ and $\bar{n}_{k\,\mathrm{pess}}^{(i+1)}(N)$ will be denoted by $n_k^{(i)}$ and $n_k^{(i+1)}$, respectively. Since $R_{\mathrm{pess}}^{(i)}(N)$ is exact for $i \geq N$, it need only be shown that $R_{\mathrm{pess}}^{(i)}(N) \geq R_{\mathrm{pess}}^{(i+1)}(N)$ for $i \geq 0$, $N \geq 1$.

THEOREM. *For $i \geq 0$, $N \geq 1$,*

$$R_{\mathrm{pess}}^{(i)}(N) \geq R_{\mathrm{pess}}^{(i+1)}(N) \tag{B-1}$$

$$n_b^{(i)} \geq n_b^{(i+1)}, \tag{B-2}$$

and

$$n_k^{(i+1)} \geq n_k^{(i)} \qquad 1 \leq k \leq K, \quad k \neq b. \tag{B-3}$$

PROOF. The proof will be by induction on i. Consider $i = 0$.

$$R_{\mathrm{pess}}^{(0)}(N) = N \geq 1 + L_b\left(\frac{N-1}{Z+N-1}\right)(N-1) = R_{\mathrm{pess}}^{(1)}(N),$$

establishing relation B-1. For $i = 0$, relation B-2 becomes

$$\frac{N}{Z+N}N \geq \frac{R_{b\,\mathrm{pess}}^{(1)}(N)}{Z + R_{\mathrm{pess}}^{(1)}(N)}N$$

$$\Leftrightarrow N(R_{\mathrm{pess}}^{(1)}(N) - R_{b\,\mathrm{pess}}^{(1)}(N)) + Z(N - R_{b\,\mathrm{pess}}^{(1)}(N)) \geq 0.$$

This last relation is easily seen to hold. Finally, relation B-3 holds for $i = 0$ since $n_k^{(0)} = 0$ for $k \neq b$. The induction basis is now established.

Assume that the theorem holds for $i \geq 0$ and consider $i + 1$. For $N = 1$, relations B-1, B-2, and B-3 are equalities; therefore, only $N + 1$ ($N \geq 1$) need be considered. Relation B-1 becomes

$$R_{\mathrm{pess}}^{(i+1)}(N+1) \geq R_{\mathrm{pess}}^{(i+2)}(N+1) \Leftrightarrow \sum_{k=1}^{K} L_k(1 + n_k^{(i)}) \geq \sum_{k=1}^{K} L_k(1 + n_k^{(i+1)})$$

$$\Leftrightarrow L_b(n_b^{(i)} - n_b^{(i+1)}) - \sum_{k \neq b} L_k(n_k^{(i+1)} - n_k^{(i)}) \geq 0.$$

From the inductive hypothesis, each term in the above expression is nonnegative. This implies that the above relation will be established if it can be shown that

$$L_b(n_b^{(i)} - n_b^{(i+1)}) - \sum_{k \neq b} L_b(n_k^{(i+1)} - n_k^{(i)}) \geq 0.$$

By dividing through by L_b, and noting that

$$\sum_{k=1}^{K} n_k^{(i)} = \frac{R_{\text{pess}}^{(i)}(N)}{Z + R_{\text{pess}}^{(i)}(N)} N$$

and

$$\sum_{k=1}^{K} n_k^{(i+1)} = \frac{R_{\text{pess}}^{(i+1)}(N)}{Z + R_{\text{pess}}^{(i+1)}(N)} N,$$

it can be seen that this is equivalent to

$$\frac{R_{\text{pess}}^{(i)}(N)}{Z + R_{\text{pess}}^{(i)}(N)} \geq \frac{R_{\text{pess}}^{(i+1)}(N)}{Z + R_{\text{pess}}^{(i+1)}(N)}.$$

Since $R_{\text{pess}}^{(i)}(N) \geq R_{\text{pess}}^{(i+1)}(N)$ by the inductive hypothesis, this last relation holds, and relation B-1 is established for $i + 1$.

Relation B-1 then yields

$$\frac{R_{\text{pess}}^{(i+1)}(N+1)}{Z + R_{\text{pess}}^{(i+1)}(N+1)} (N+1) \geq \frac{R_{\text{pess}}^{(i+2)}(N+1)}{Z + R_{\text{pess}}^{(i+2)}(N+1)} (N+1).$$

Expanding the numerators produces

$$\sum_{k=1}^{K} \frac{R_{k\,\text{pess}}^{(i+1)}(N+1)}{Z + R_{\text{pess}}^{(i+1)}(N+1)} (N+1) \geq \sum_{k=1}^{K} \frac{R_{k\,\text{pess}}^{(i+2)}(N+1)}{Z + R_{\text{pess}}^{(i+2)}(N+1)} (N+1).$$

From this last relation, it can be seen that relation B-2 or relation B-3 could be violated for $i + 1$ (at $N + 1$) only if there is a center $k \neq b$ such that

$$\frac{R_{k\,\text{pess}}^{(i+1)}(N+1)}{Z + R_{\text{pess}}^{(i+1)}(N+1)} (N+1) > \frac{R_{k\,\text{pess}}^{(i+2)}(N+1)}{Z + R_{\text{pess}}^{(i+2)}(N+1)} (N+1)$$

$$\Leftrightarrow \frac{L_k(1 + n_k^{(i)})}{Z + R_{\text{pess}}^{(i+1)}(N+1)} > \frac{L_k(1 + n_k^{(i+1)})}{Z + R_{\text{pess}}^{(i+2)}(N+1)}$$

$$\Leftrightarrow \frac{L_k(1 + n_k^{(i)})}{L_k(1 + n_k^{(i+1)})} > \frac{Z + R_{\text{pess}}^{(i+1)}(N+1)}{Z + R_{\text{pess}}^{(i+2)}(N+1)}.$$

However, the right-hand side has been proven to be greater than or equal to 1, and the left-hand side is less than or equal to one by the inductive hypothesis. As this is a contradiction, relations B-2 and B-3 and the theorem are established for $i + 1$. By induction, the theorem is established for $i \geq 0$. □

ACKNOWLEDGMENTS

We thank B. I. Galler, G. S. Graham, and J. Zahorjan, along with several anonymous referees, for their helpful and constructive comments on this work and its presentation.

After this paper was accepted for presentation at the 1982 ACM SIGMETRICS Conference on Measurement and Modeling of Computer Systems, and while it was being refereed for publication, we learned that Professor Rajan Suri of Harvard University had independently formulated a generalization of BJBs quite similar to PBH bounds, calling them *Generalized Quick Bounds.*

REFERENCES

1. BARD, Y. Some extensions to multiclass queueing network analysis. In *Performance of Computer Systems*, M. Arato, A. Butrimenko, and E. Gelenbe (Eds.), North Holland, Amsterdam, 1979.
2. CHANDY, K. M., AND NEUSE, D. Linearizer: A heuristic algorithm for queueing network models of computing systems. *Commun. ACM* 25, 2 (Feb. 1982), 126–134.
3. CHANDY, K. M., AND SAUER, C. H. Computational algorithms for product form queueing networks. *Commun. ACM* 23, 10 (Oct. 1980), 573–583.
4. DENNING, P. J., AND BUZEN, J. P. The operational analysis of queueing network models. *ACM Comput. Surv. 10*, 3 (Sept. 1978), 225–261.
5. GRAHAM, G. S. Asymptotic properties of mean device queue length. Unpublished report, Univ. of Toronto, Toronto, Canada.
6. KLEINROCK, L. *Queueing Systems.* Volume 2, *Computer Applications.* Wiley, New York, 1976.
7. KOBAYASHI, H. A computational algorithm for queue distributions via the Polya theory of enumeration. In *Performance of Computer Systems*, M. Aruto, A. Butrimenko, and E. Gelenbe (Eds.), North Holland, Amsterdam, 1979.
8. LITTLE, J. D. C. A proof of the queueing formula $L = \lambda W$. *Oper. Res. 9*, 3 (May 1961), 383–387.
9. MUNTZ, R. R., AND WONG, J. Asymptotic properties of closed queueing network models. In *Proc. 8th Annual Princeton Conference on Information Science and Systems.* Princeton Univ., March 1974.
10. REISER, M., AND LAVENBERG, S. S. Mean value analysis of closed multichain queueing networks. *J. ACM 27*, 2 (April 1980), 313–322.
11. REISER, M., AND SAUER, C. H. Queueing network models: Methods of solution and their program implementation. In *Current Trends in Programming Methodology.* Vol. III, *Software Modeling*, K. M. Chandy and R. T. Yeh (Eds.), Prentice-Hall, Englewood Cliffs, N.J., 1978.
12. SAUER, C. H., AND CHANDY, K. M. *Computer System Performance Modeling.* Prentice-Hall, Englewood Cliffs, N.J., 1981.
13. SCHWEITZER, P. Approximate analysis of multiclass closed networks of queues. In *Proc. Int. Conf. Stochastic Control and Optimization* (Amsterdam), 1979.
14. ZAHORJAN, J. The approximate solution of large queueing network models. Ph.D. dissertation (Tech. rep. CSRG-122), Univ. of Toronto, Toronto, Canada, 1980.
15. ZAHORJAN, J., SEVCIK, K. C., EAGER, D. L., AND GALLER, B.I. Balanced job bound analysis of queueing networks. *Commun. ACM 25*, 2 (Feb. 1982), 134–141.

Received February 1982; revised July 1982; accepted August 1982