

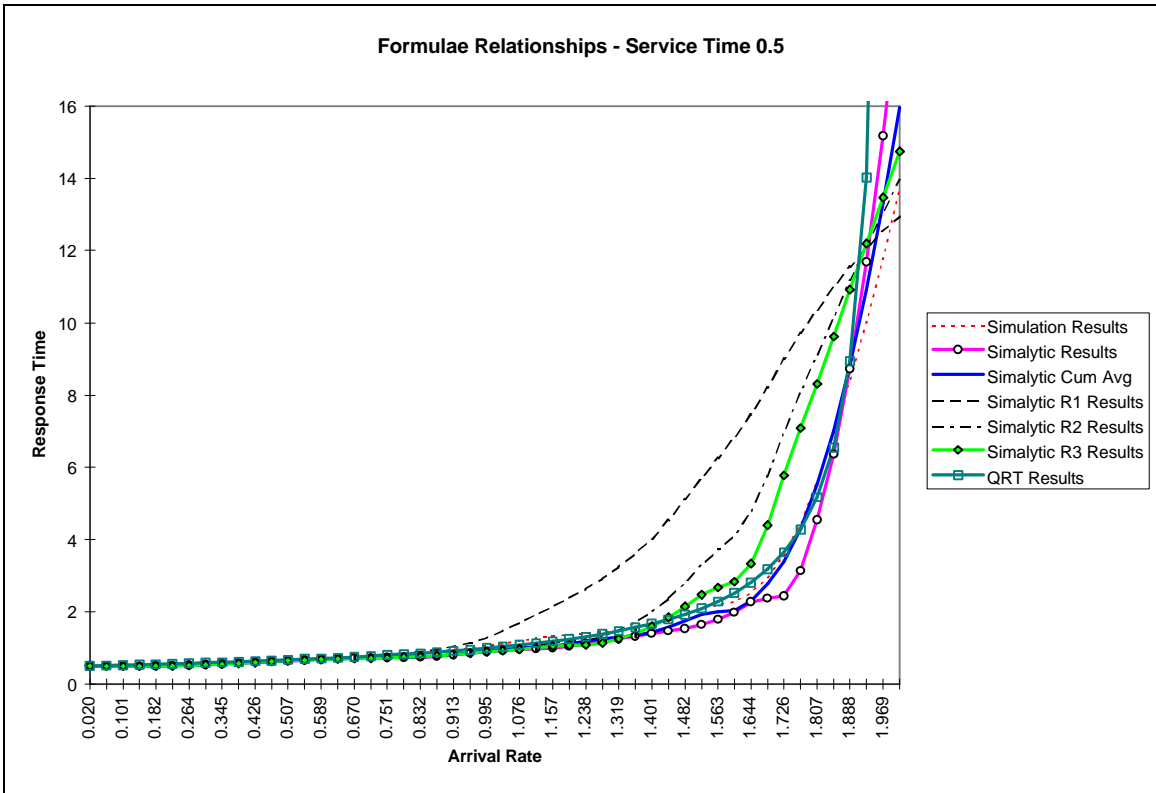
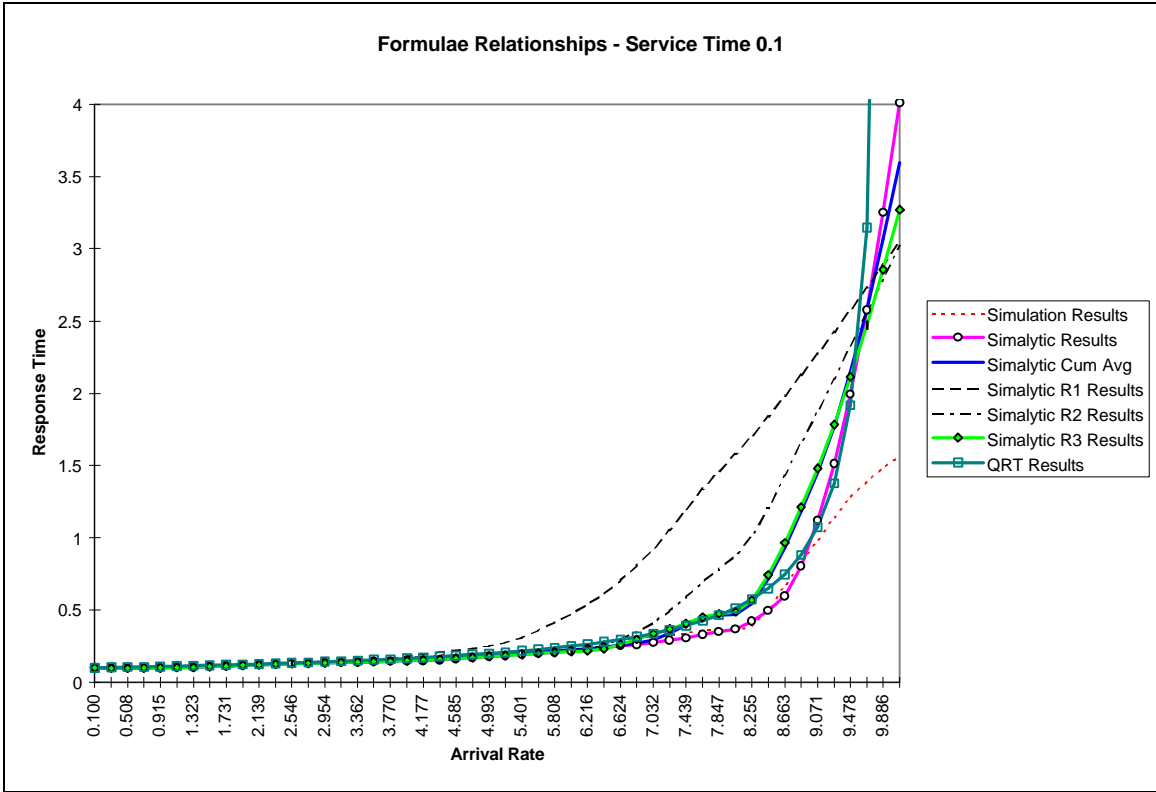
7.5 Appendix E: MathCAD Formulae Results Charts

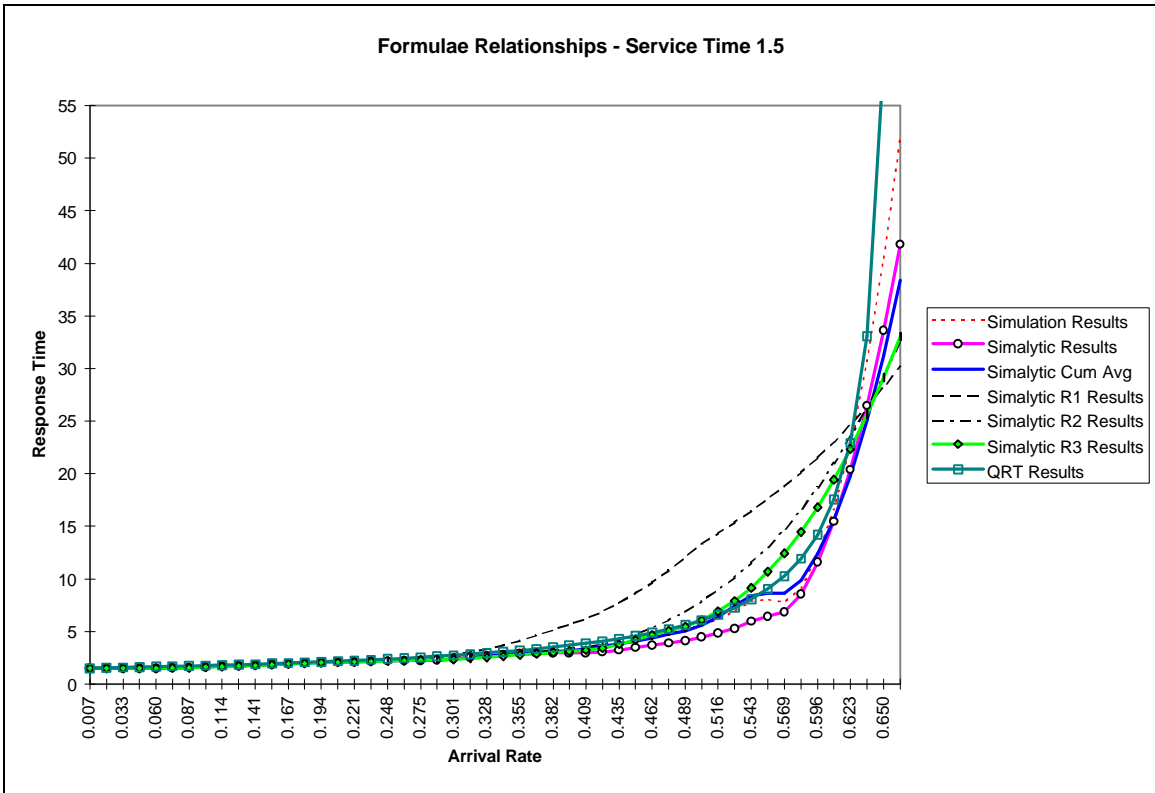
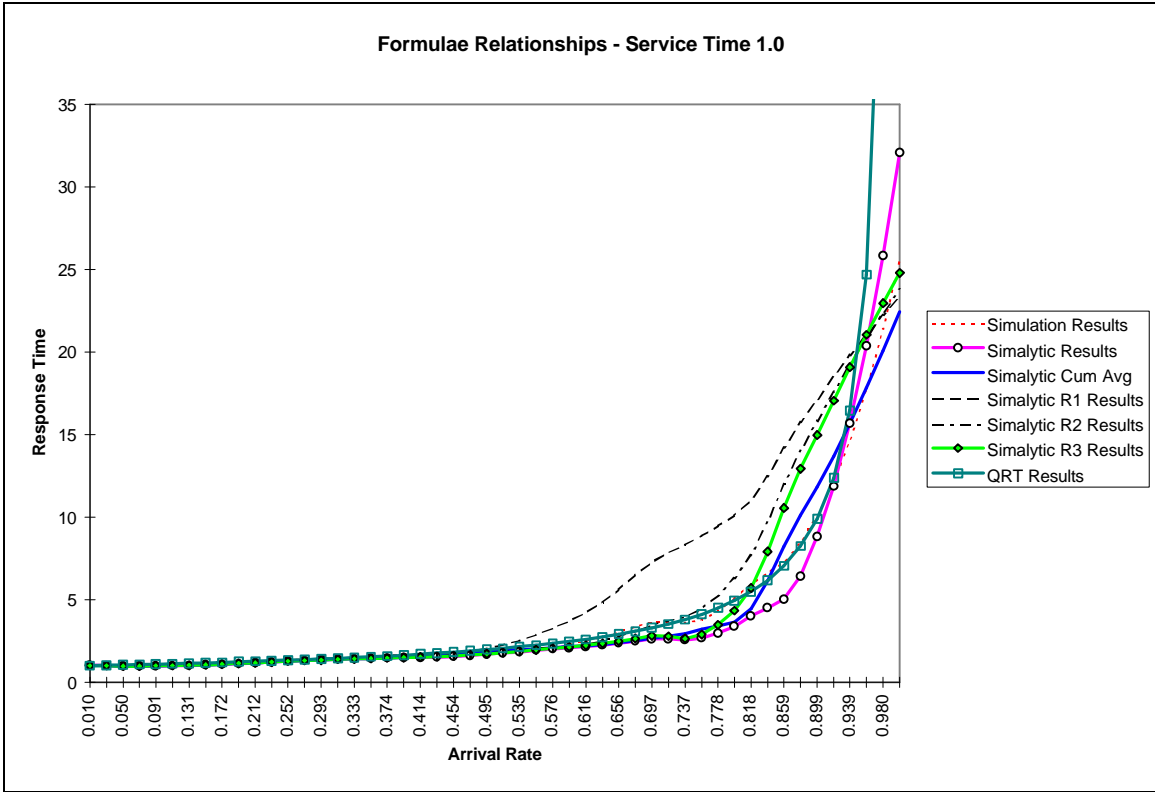
The following Microsoft Excel charts show the detailed results of the mathematical formulae as implemented using MathCAD. The MathCAD worksheet used to generate the results is listed in *Figure 60 MathCAD Worksheet Listing* on page 199. The data was collected into a spreadsheet from MathCAD output files. Each set of data is copied to a worksheet and then charted.

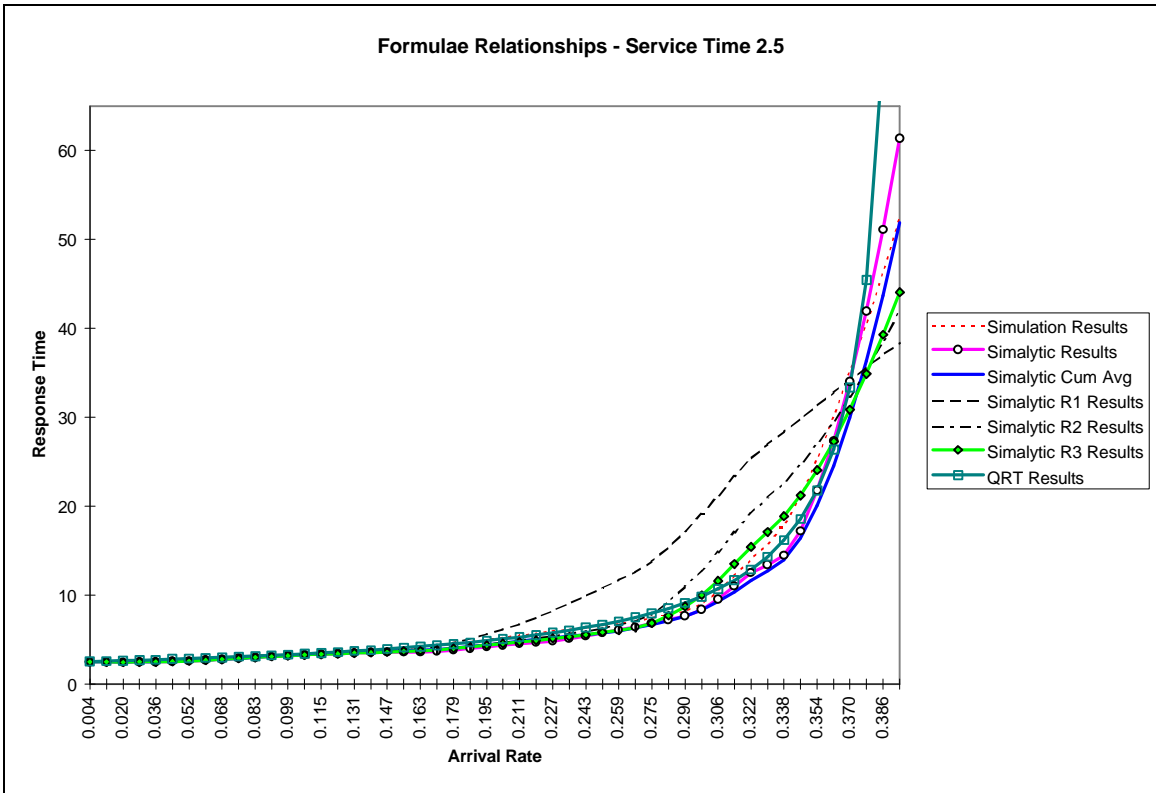
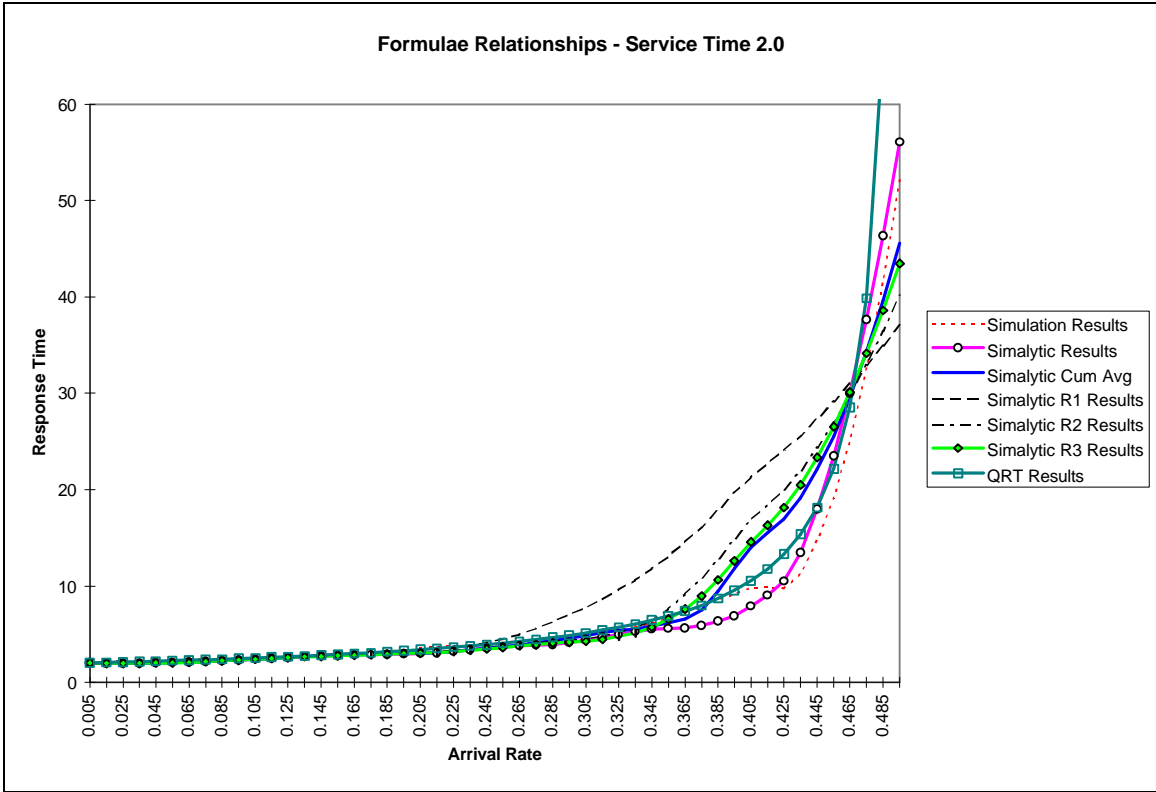
7.5.1 Formulae Relationships

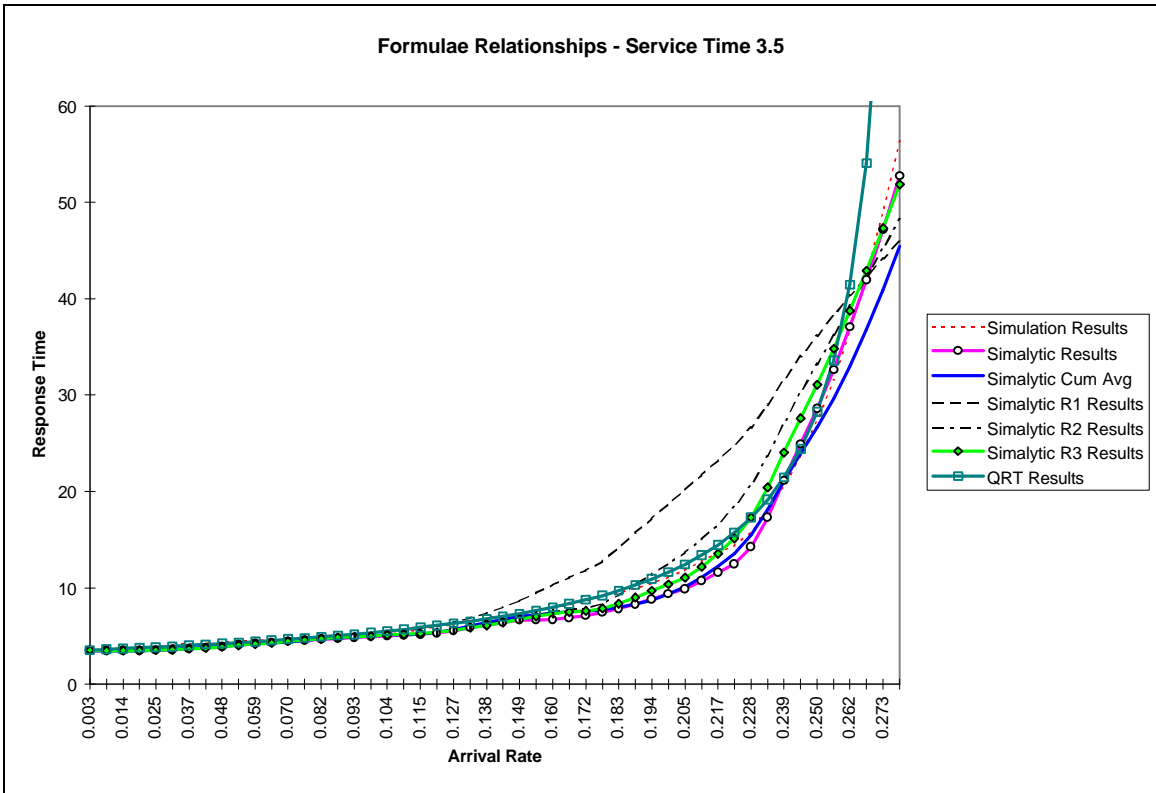
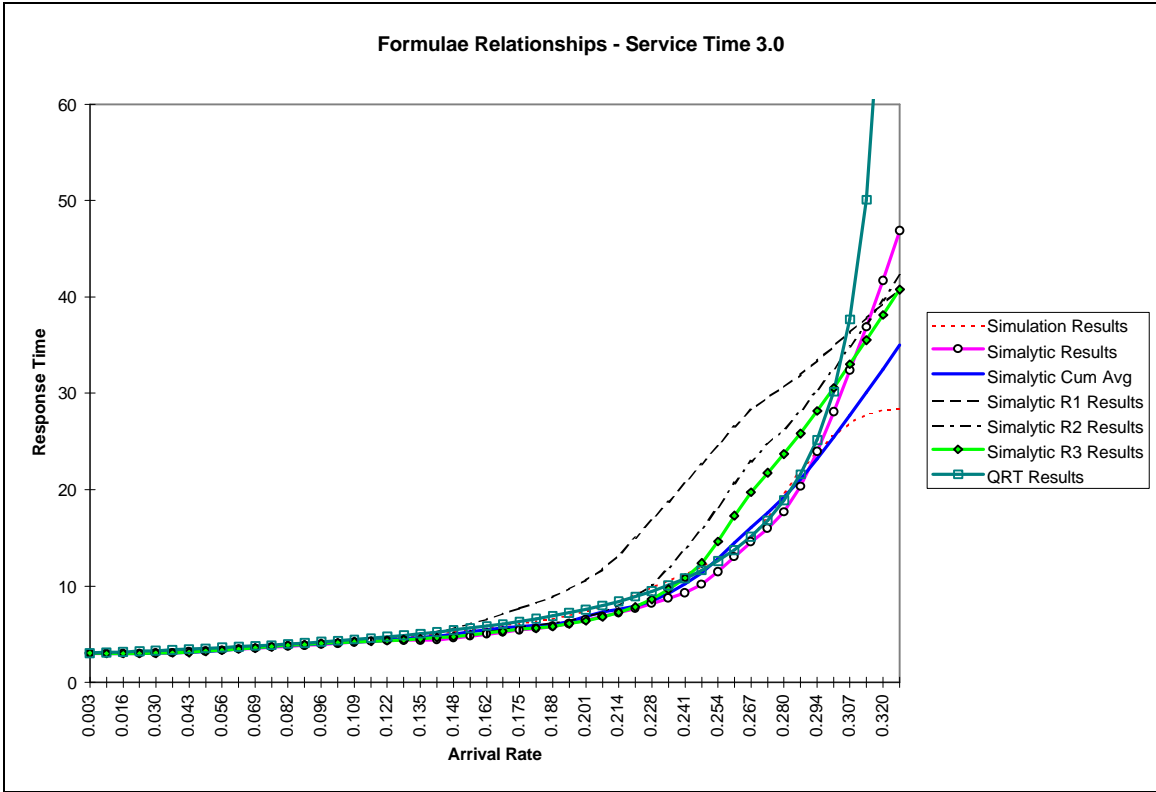
The first group of charts, labeled **Formulae Relationships - Service Time**, shows the relationships between the three formulae (*Equation 1 Analytic Response Time Formula* on page 32, *Equation 2 Simulation Response Time Formula* on page 35 and *Equation 3 Simalytic Response Time Formula* on page 61) implemented in MathCAD at a series of arrival rates as described in section 3.5.2 *Verification/Validation Approach* on page 81. The line labeled **Simulation Results** shows the response times calculated by a single execution of the simplistic simulation technique in MathCAD. Some anomalies are evident for this line in some of the charts because of the effect the arrival distribution variation has on a single model run. These anomalies, seen as curve fall-off at high arrival rates, disappear when the model is rerun using a different random number seed. The line labeled **QRT Results** shows the response times calculated by the MathCAD formula (QueuingRT) for queuing theory response time in *Figure 59 MathCAD Worksheet for Queuing Theory Surface Plots* on page 178. The lines labeled **Simalytic** show results calculated using different implementations of the Simalytic function. **Simalytic Results** shows the basic values in the Simalytic function response time table where the function is

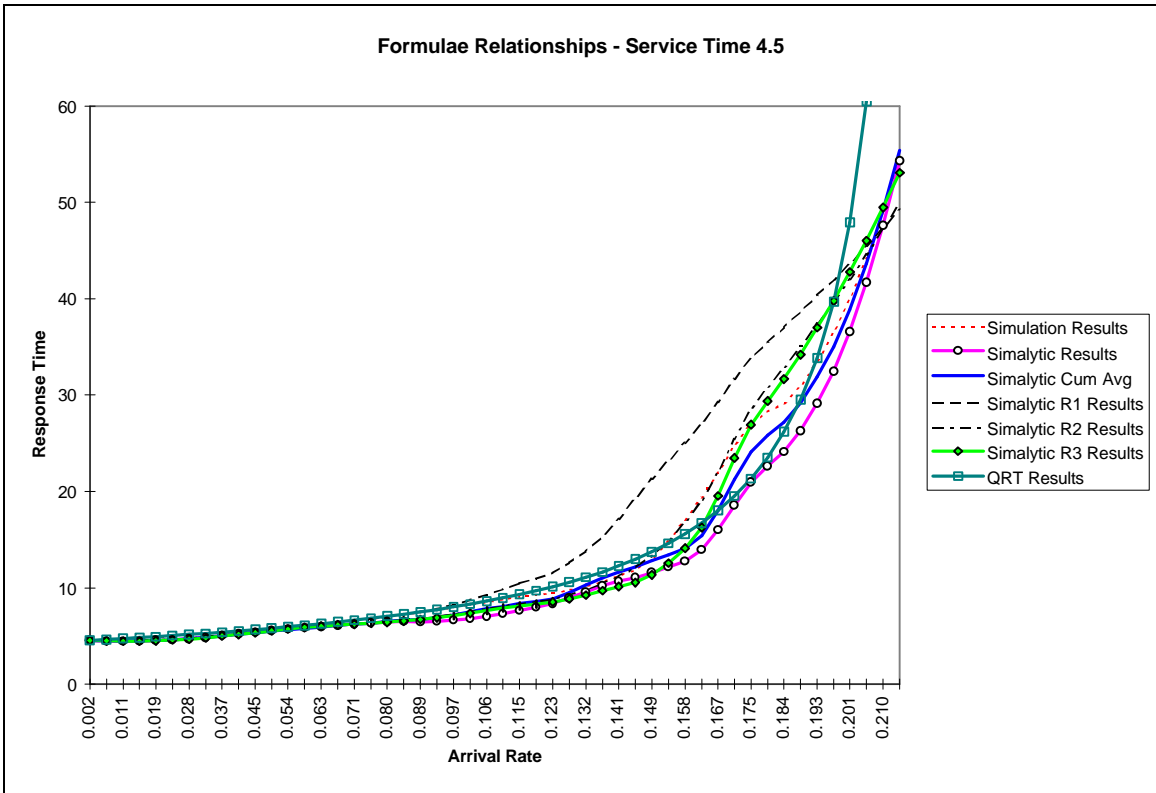
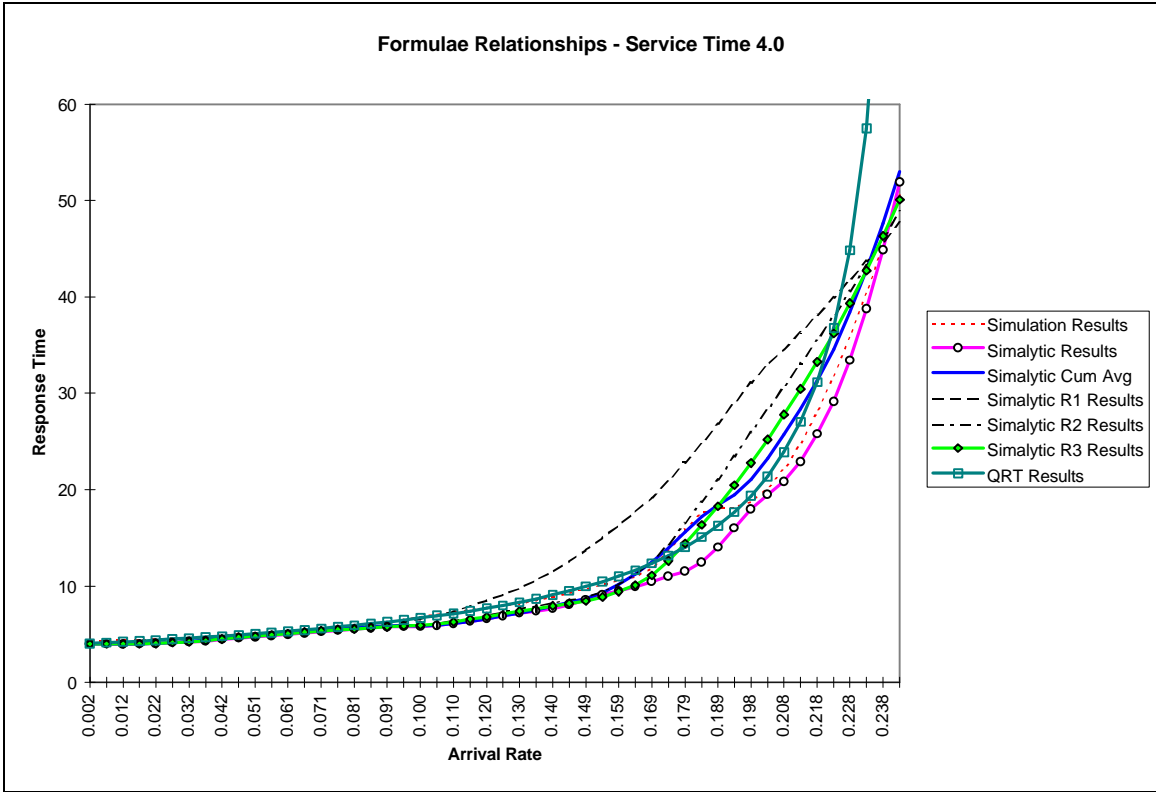
called once for each arrival rate with the average of all interarrival times. This is the best value that can be returned if the arrival rate is consistent across the sample period (i.e. a steady-state system). **Simalytic Cum Avg** shows the results when the function calculates the cumulative average of the interarrival times each time it is called and returns the response time from the table for that average. **Simalytic R1 Results**, **Simalytic R2 Results** and **Simalytic R3 Results** show the results when the function calculates a rolling average of the interarrival times each time it is called. R1 is a small rolling window (1.5% of the total events), R2 is a medium rolling window (5% of the total events) and R3 is a large rolling window (10% of the total events). In each case, the function sums the last x events (where x represents the rolling window size), divides by x and returns the response time from the table for that average. Each of these implementations is a compromise between sensitivity to short-term dynamics and consistent long-term results. The R1 results provide the greatest sensitivity to dynamic changes but experimentation has shown that the R1 results, and to a lesser degree the R2 results, are overly sensitive to distribution event order and look-up table step values. The R3 results were used in the comparisons to the simulation and queuing theory results because they provide the closest approximation to the steady-state system results but are calculated from parameters to the function and still remain sensitive to dynamic changes in the arrival rate patterns.

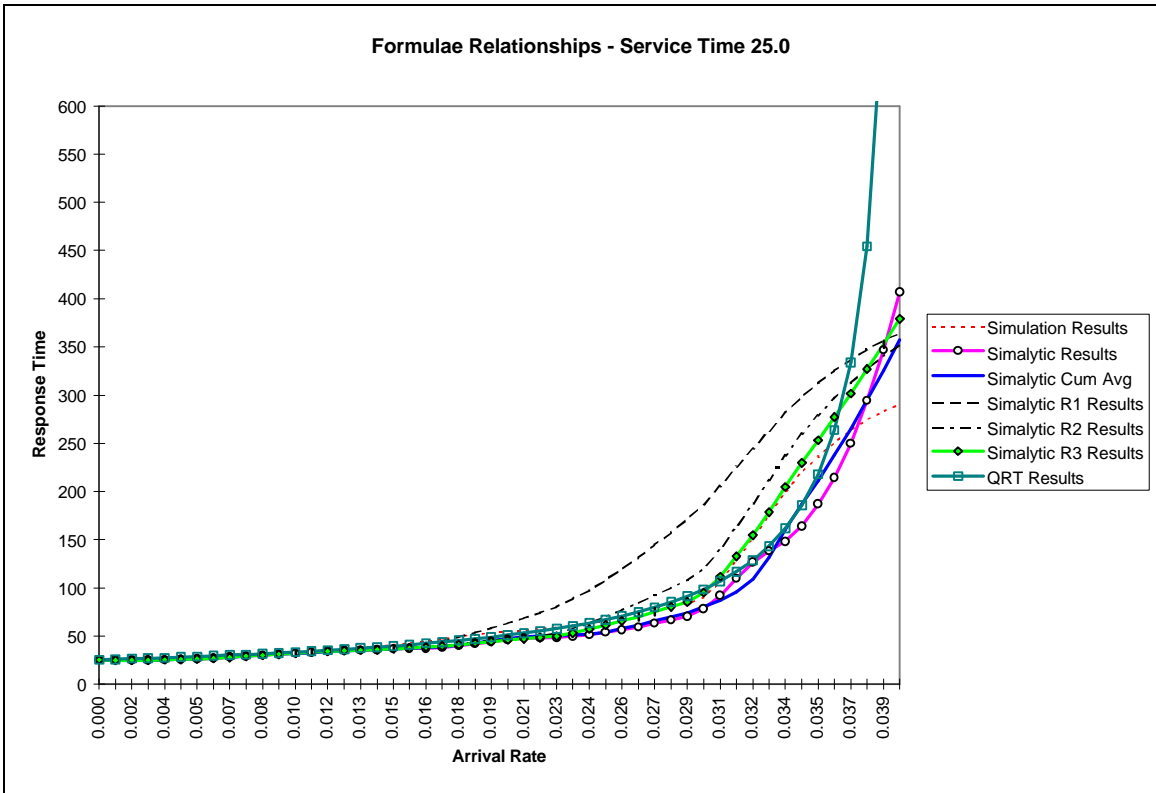
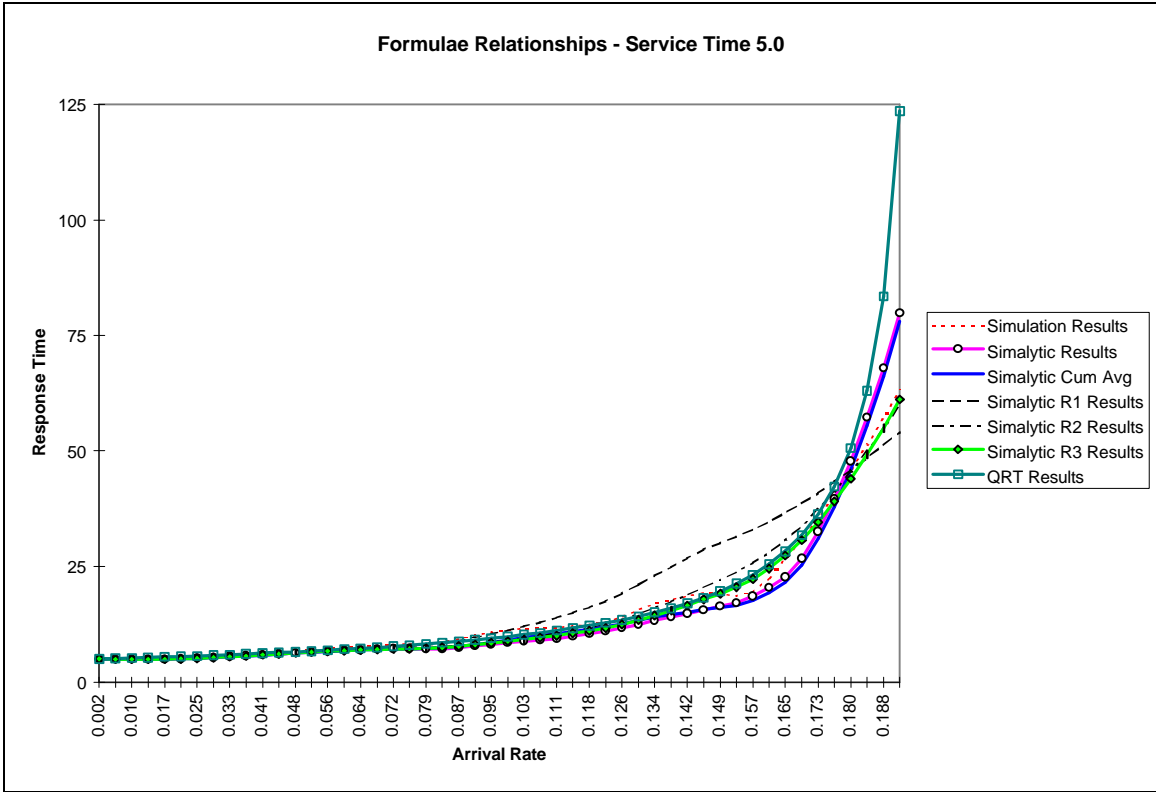


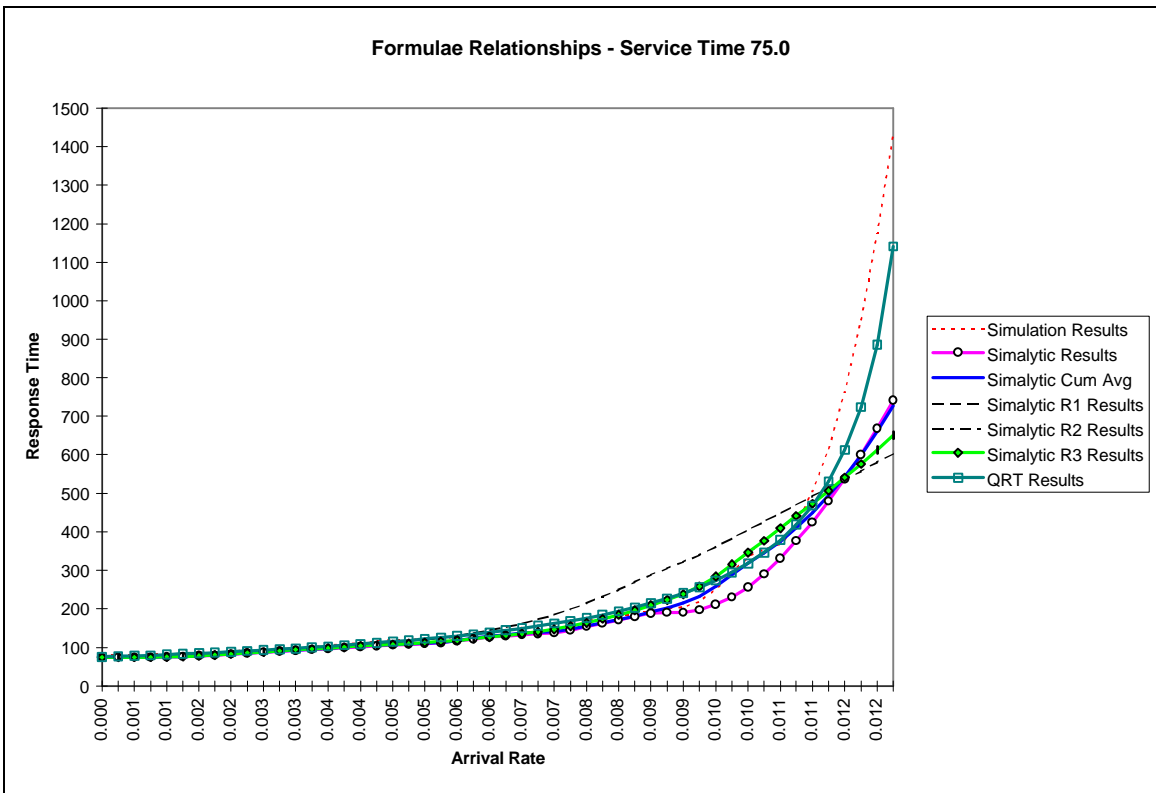
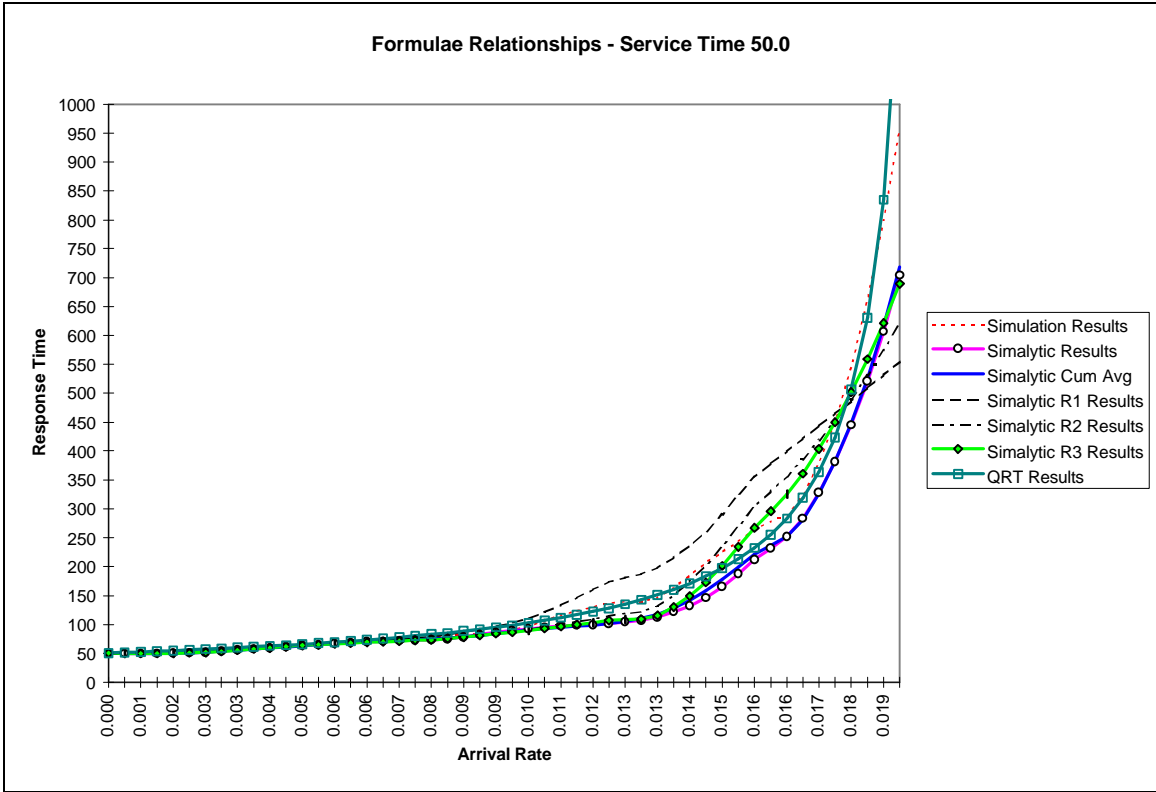


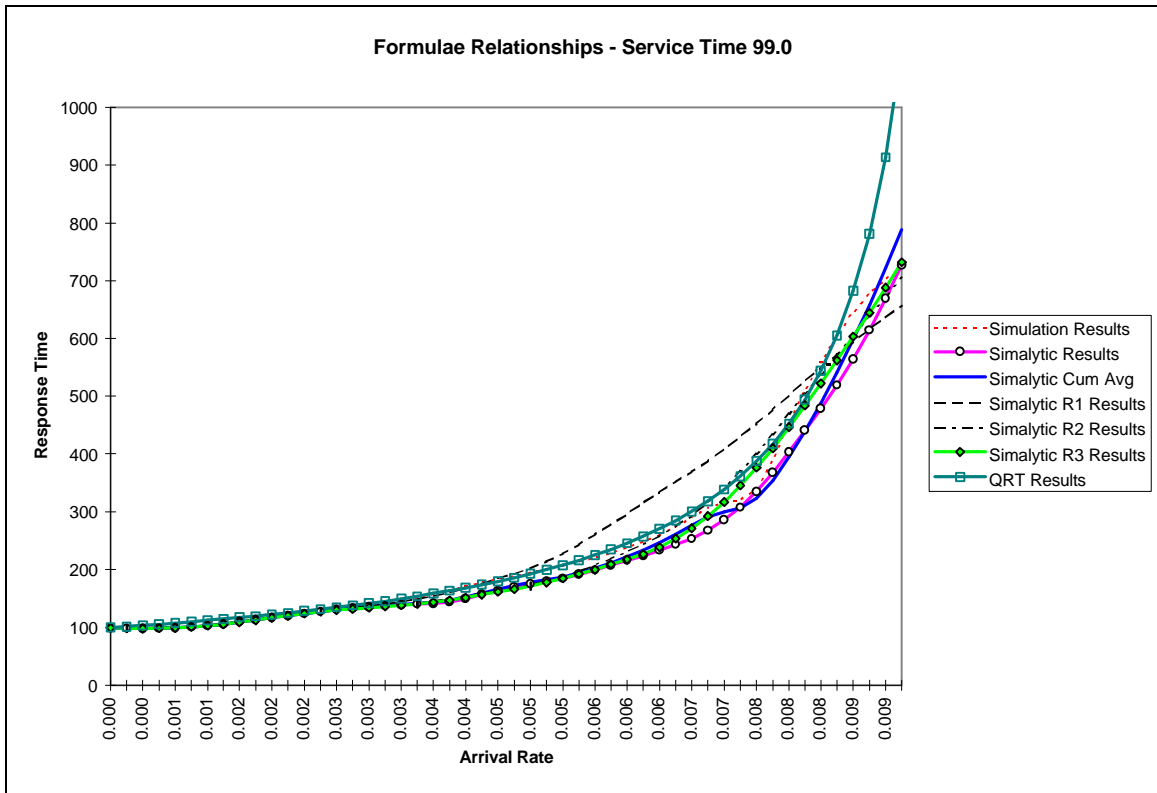












7.5.2 Relative Response Times

The next group of charts, labeled **Relative Response Times**, shows the resulting response times when servers are combined as discussed in sections 3.4.1.1 through 3.4.1.7. These charts show 32 scenarios (*Table 1 Mathematical Formulae Results* on page 93 lists the scenarios), including both series and routing systems. Each chart, which contains either three or four related scenarios (on the same chart to reduce the total number of charts and pages required), shows how the Simalytic function results track the simulation results.

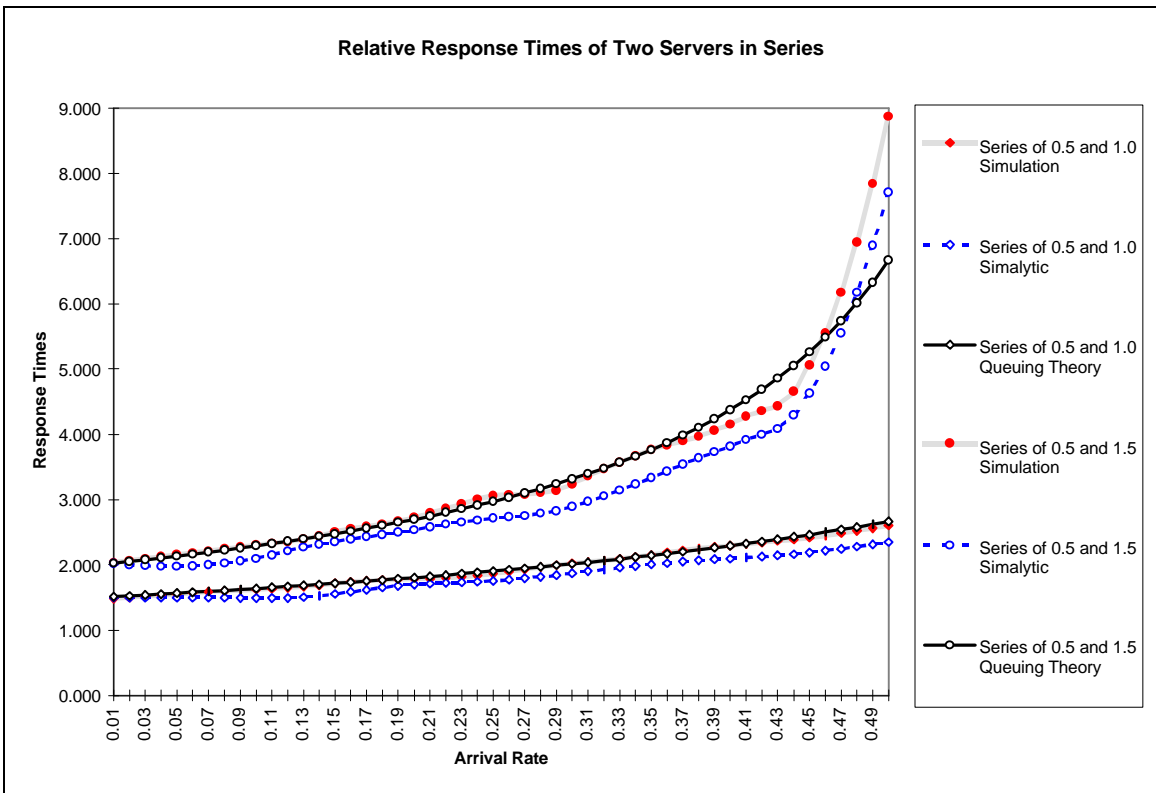
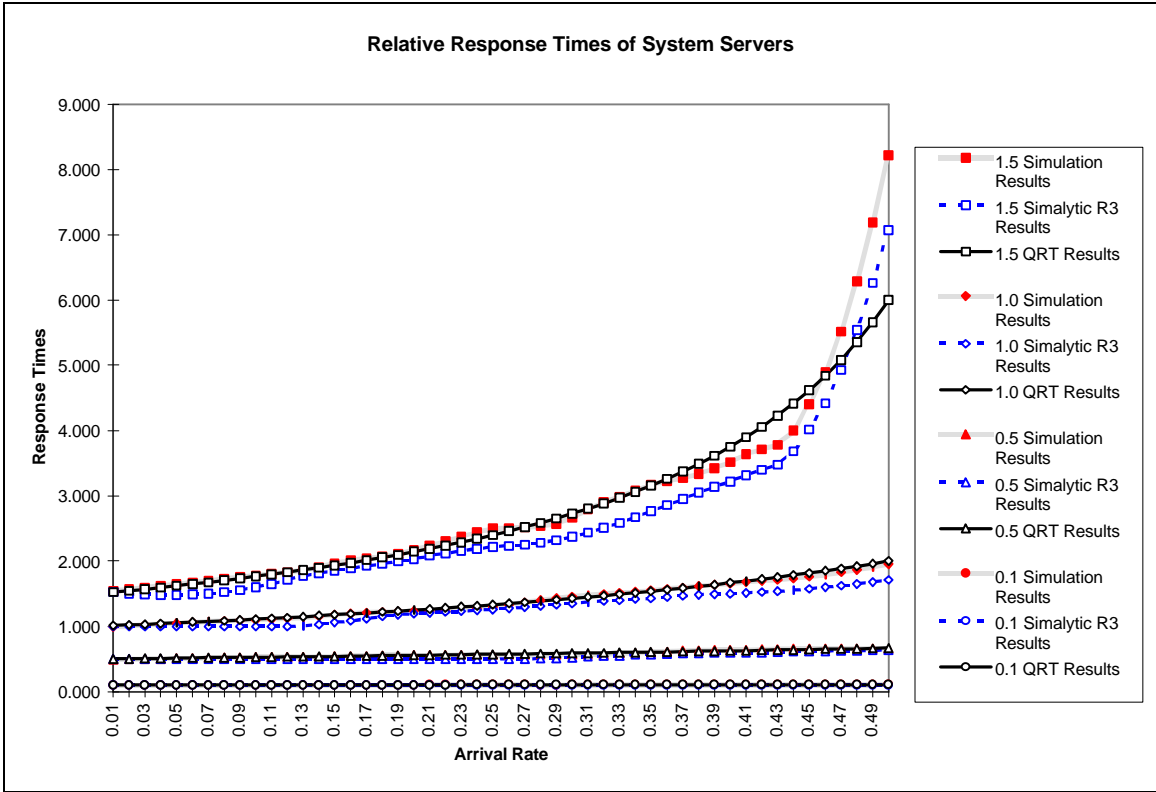
The first chart, **Relative Response Times of System Servers**, simply reproduces the results from the last charts in the same format, and using the same arrival rate scale, to provide easy reference between the single server results and the multi-server results. The

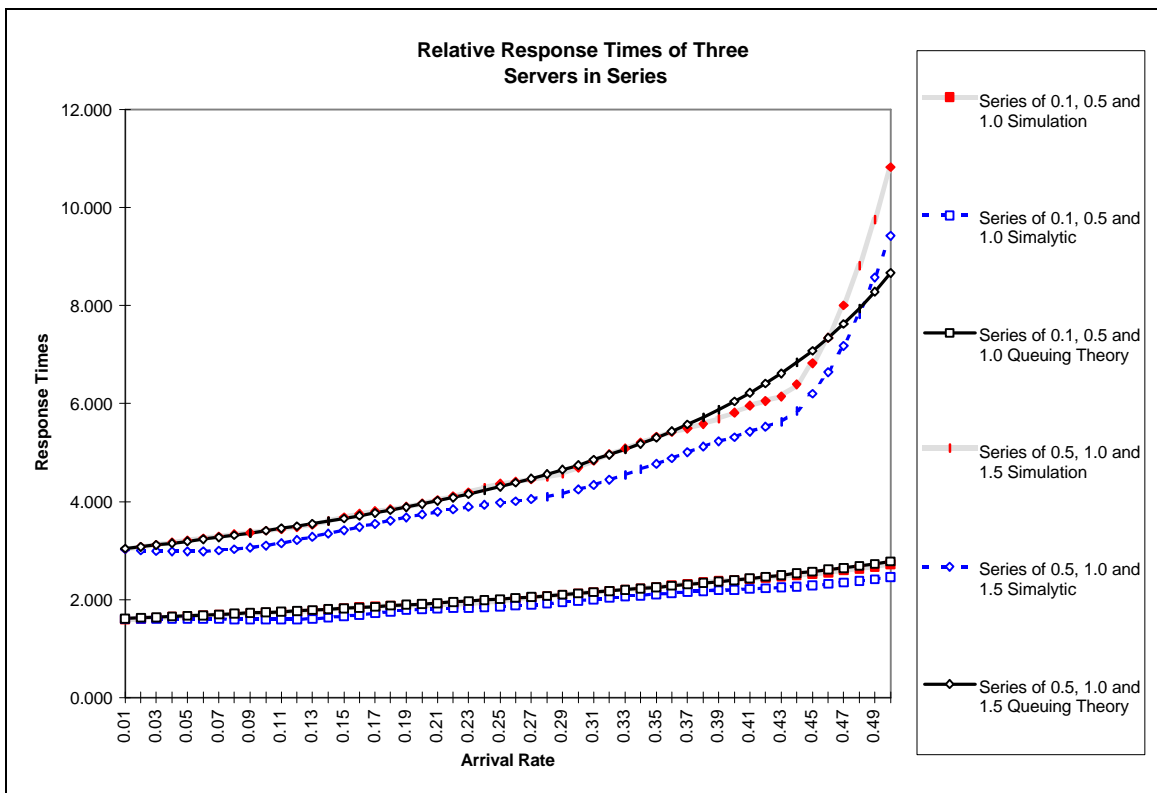
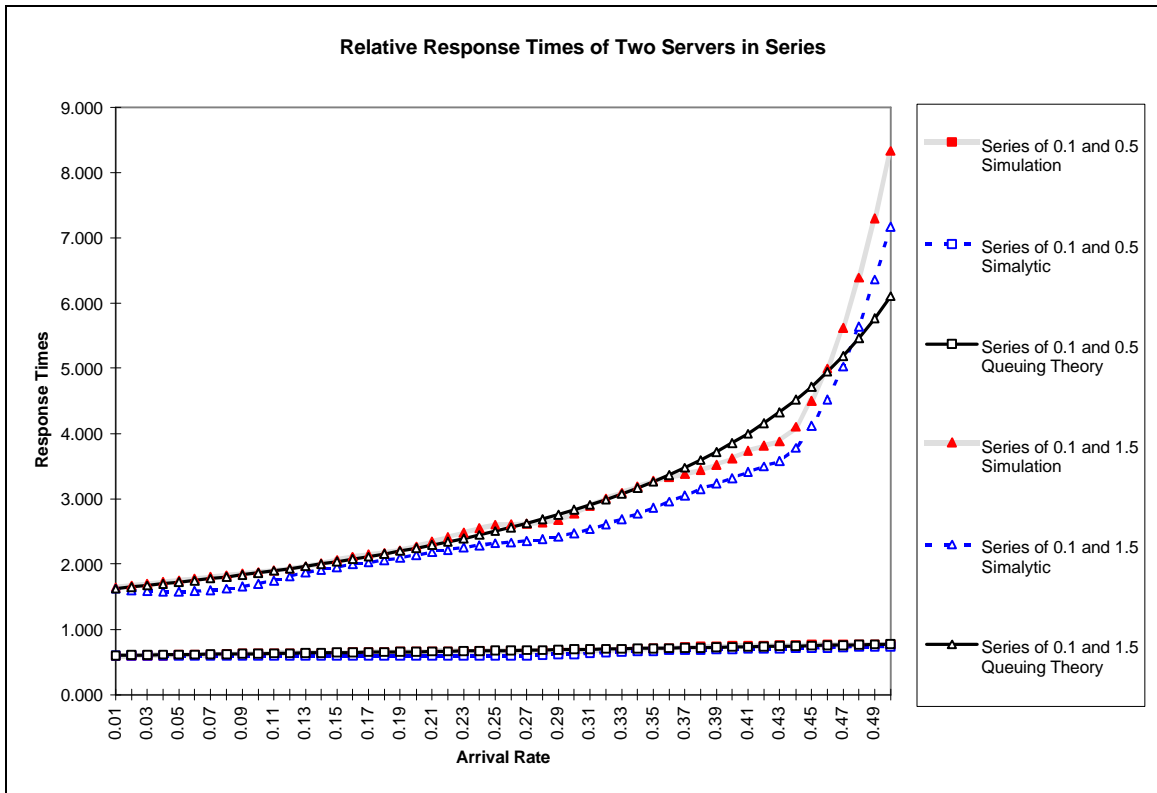
lines are labeled with the server service time and the MathCAD result name (**Simulation**, **Simalytic R3** (as discussed above) and the queuing theory function **QRT**). For example, the bottom line in the legend is labeled **0.1 QRT Results** and represents the response time for a server with a service time of 0.1 calculated by the MathCAD formula (QueuingRT) for queuing theory response time in *Figure 59 MathCAD Worksheet for Queuing Theory Surface Plots* on page 178.

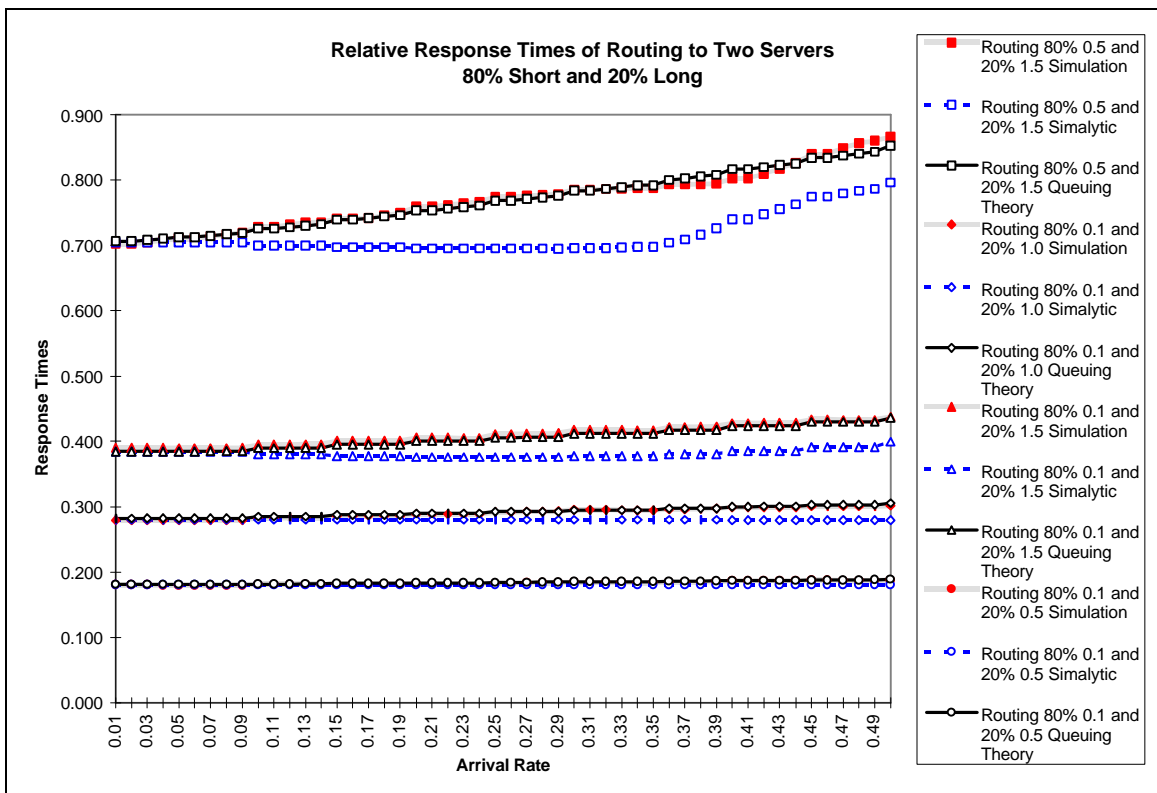
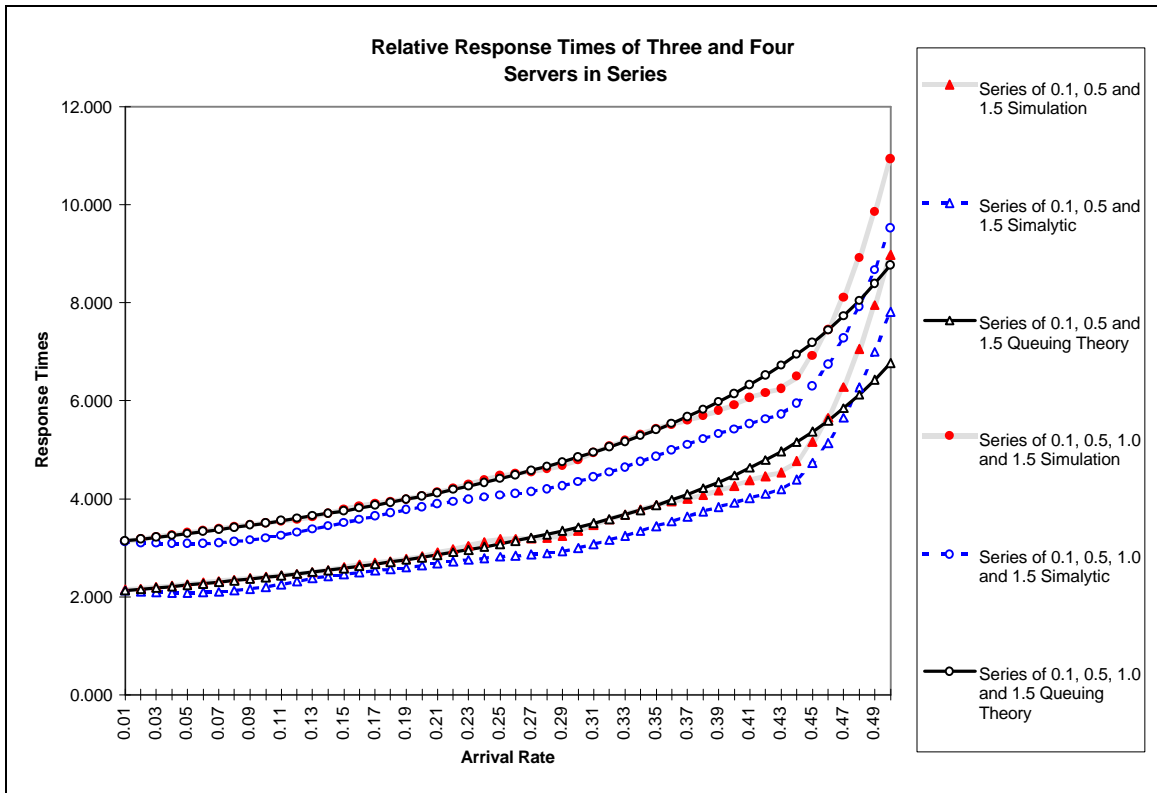
The rest of the charts show the resulting response times when servers are combined as discussed in sections 3.4.1.1 through 3.4.1.7. Each line is labeled with the name of the scenario, which is intended to be descriptive of the number of servers, the relationship between servers (series or routing) and the service times of each of the servers. For example, in the chart titled **Relative Response Times of Two Servers in Series**, the top line in the legend is labeled, **Series of 0.1 and 0.5 Simulation**, and shows the results of combining two servers in series where one has a service time of 0.1 and the other has a service time of 0.5, and the results were created using the MathCAD simulation technique. The bottom line in the same legend, **Series of 0.5 and 1.5 Queuing Theory**, shows the results of combining two servers in series where one has a service time of 0.5 and the other has a service time of 1.5, and the results were created using the MathCAD queuing theory technique. Another example is shown in the chart titled **Relative Response Times of Routing to Two Servers 80% Short and 20% Long**, where the top line in the legend is labeled **Routing 80% 0.5 and 20% 1.5 Simulation**, and shows the results of combining two servers where 80% of the transactions are short (routed to the 0.5 service time server) and 20%

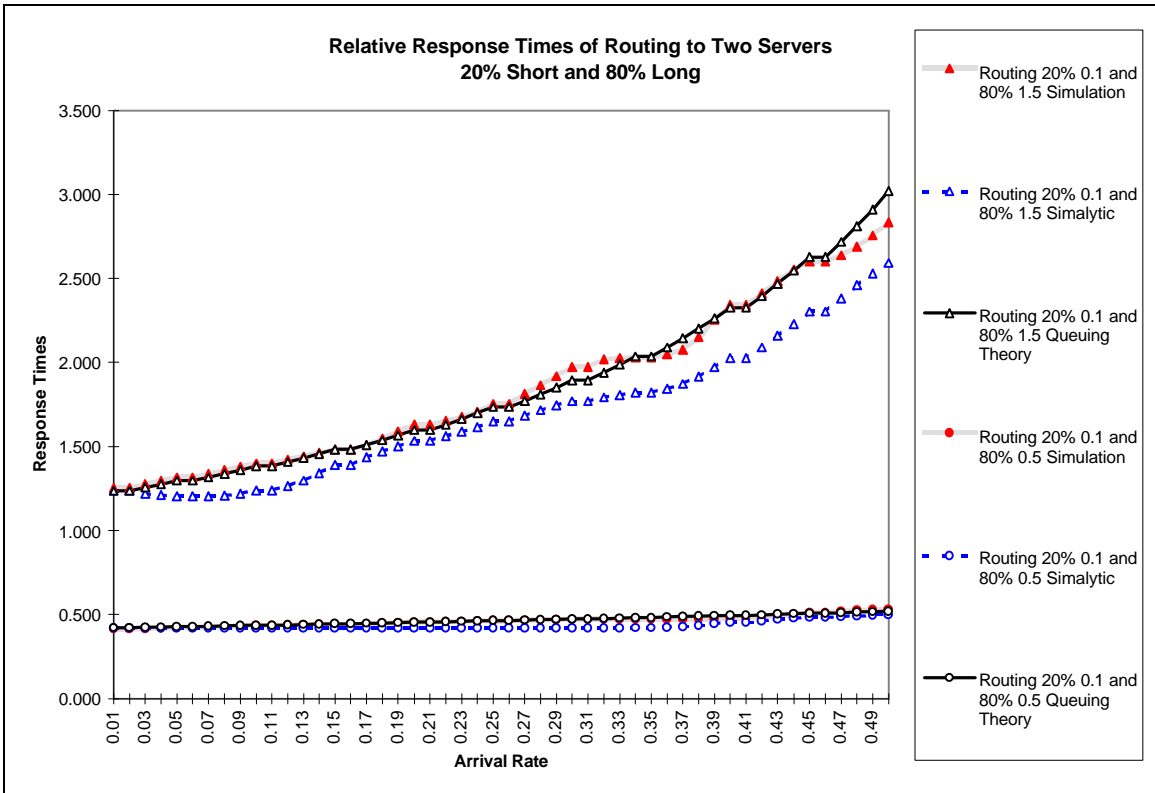
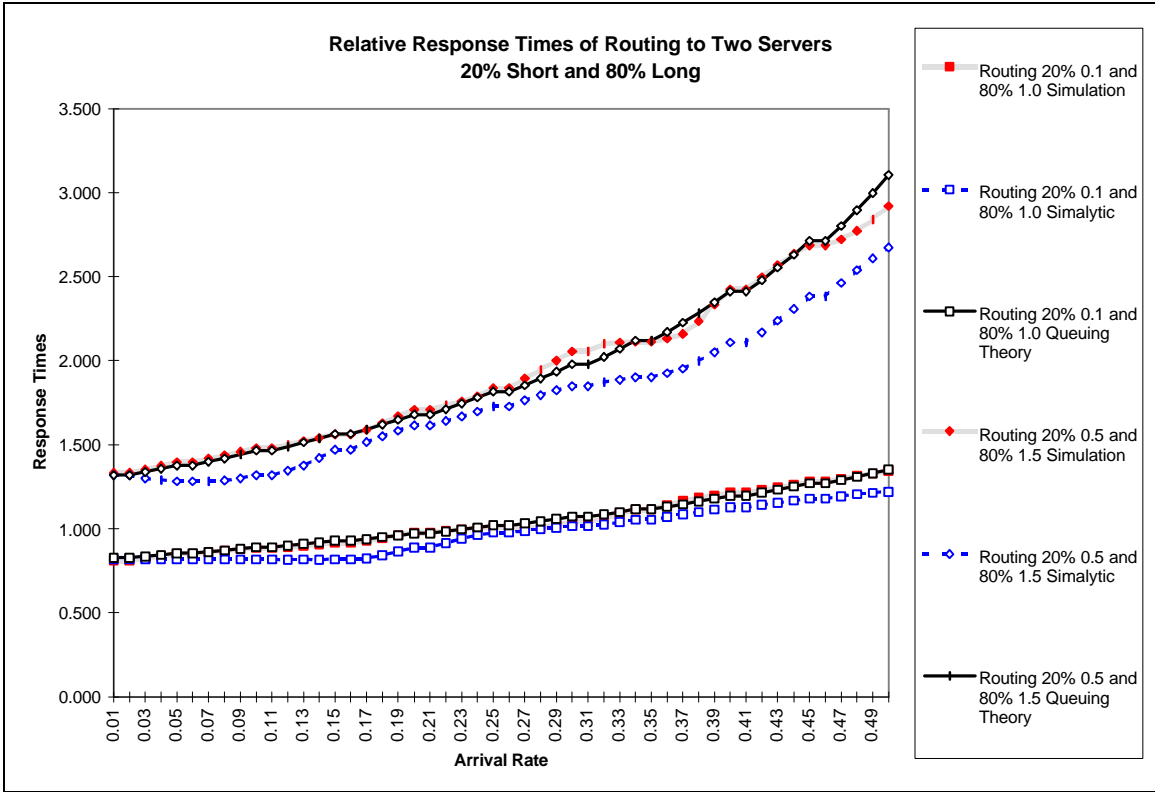
are long (routed to the 1.5 service time server) and the results were created using the MathCAD simulation technique.

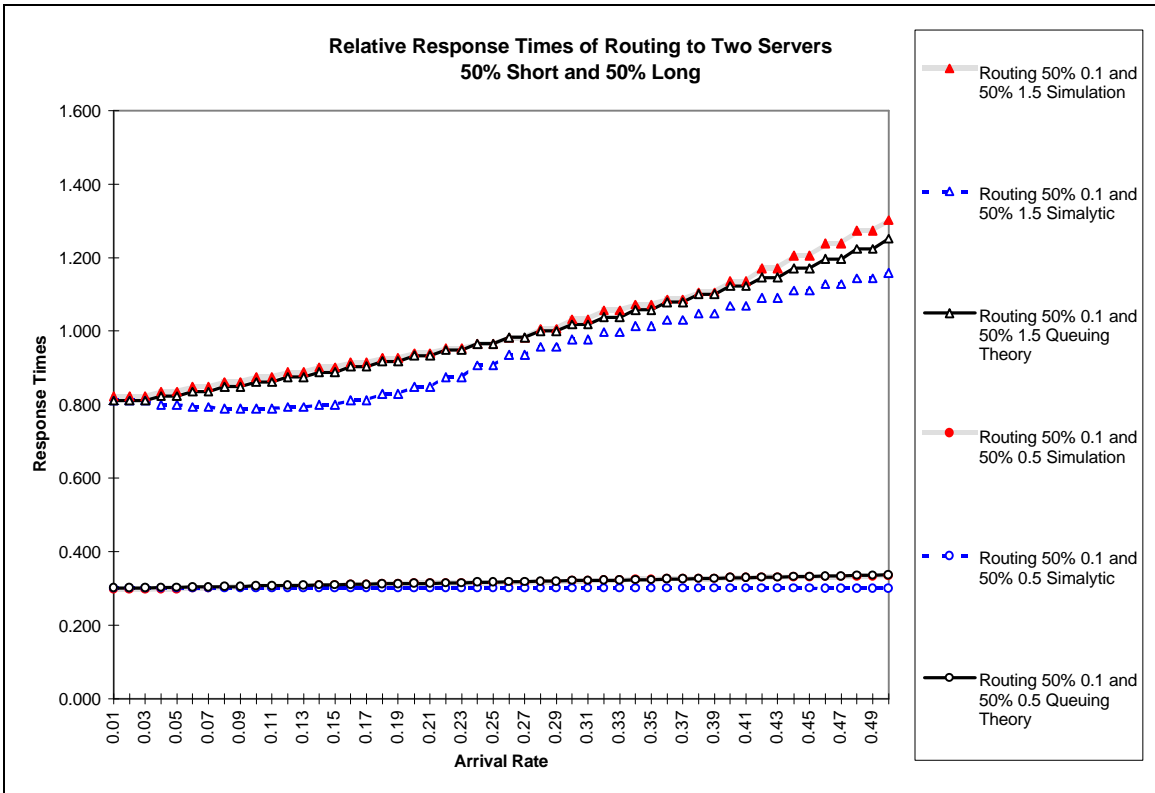
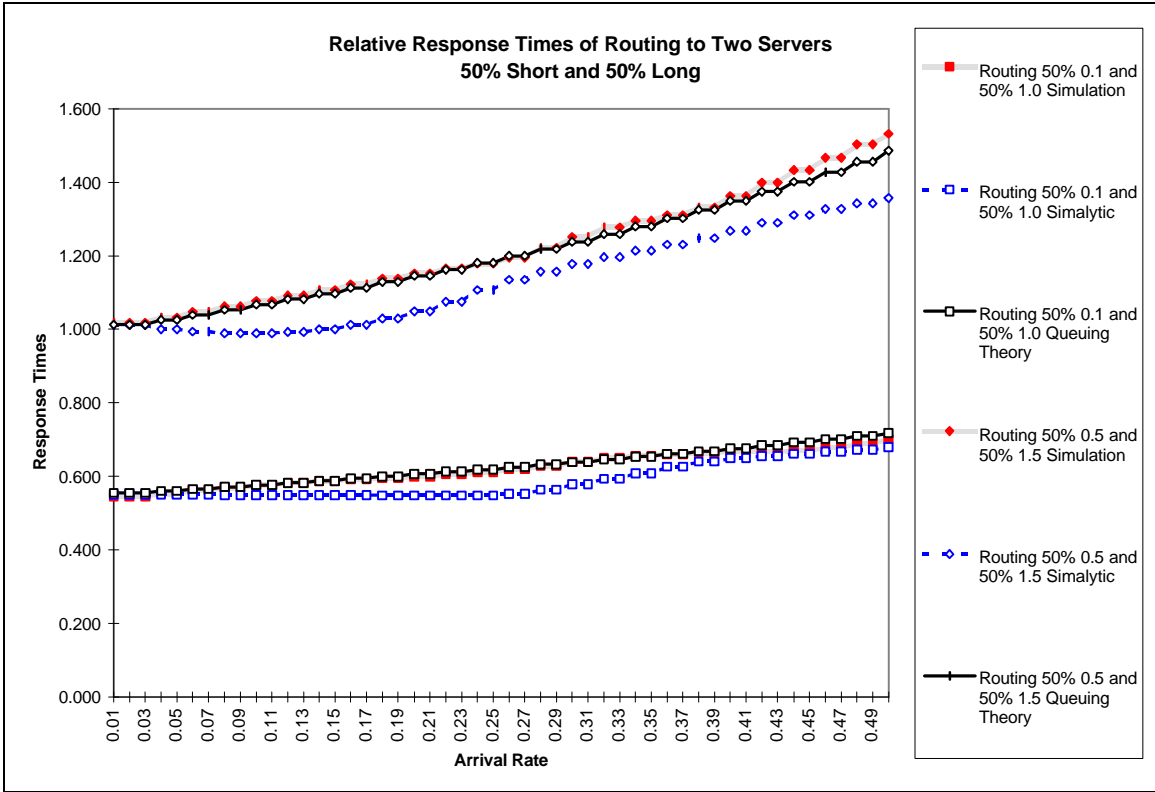
In theory, the lines with the same scenario name, except for the technique used, should produce the same curve (these have the same shaped symbol on each chart). For example, the lines labeled **Routing 80% 0.5 and 20% 1.5 Simulation**, **Routing 80% 0.5 and 20% 1.5 Simalytic** and **Routing 80% 0.5 and 20% 1.5 Queuing Theory**, (shown with a square symbol) should be indistinguishable. The small differences seen in the lines for each group show how well the results from the three techniques match. The purpose of these charts is to show that, in all cases, the differences are relatively small. Additional discussion and detailed analysis is in section *3.5.3 Validation of the Mathematical Foundation* on page 86.

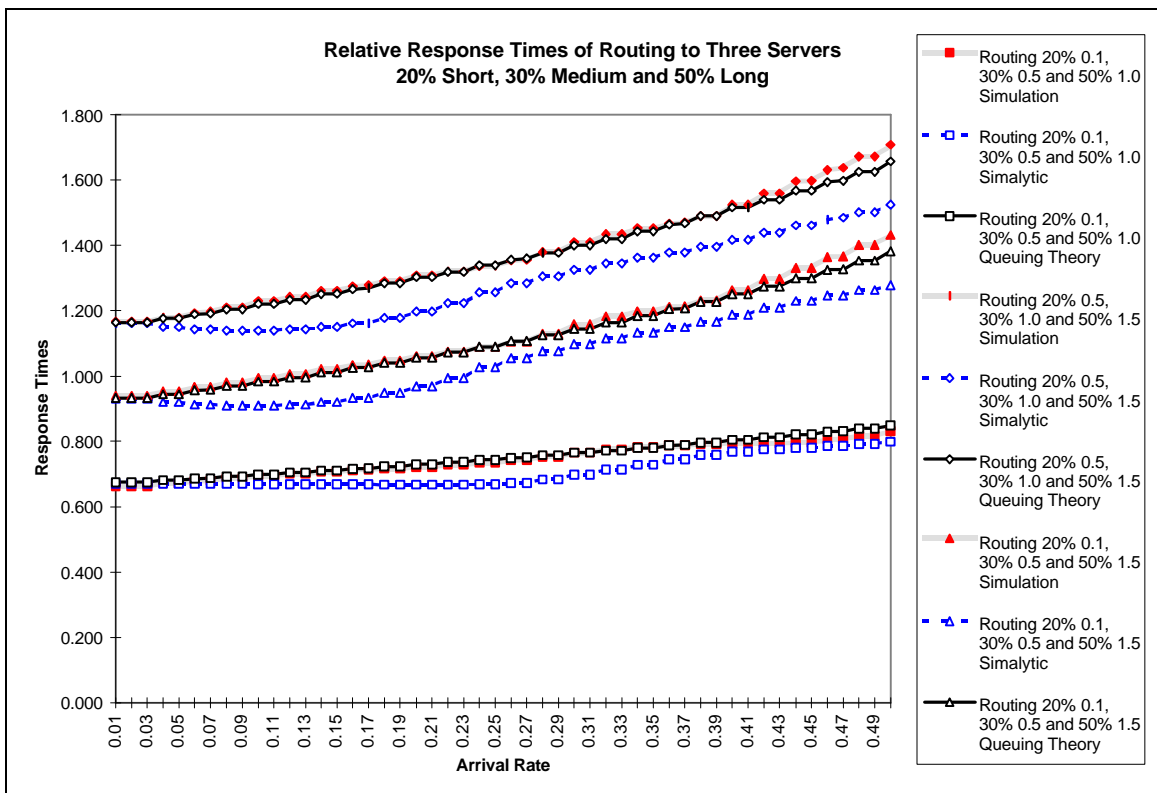
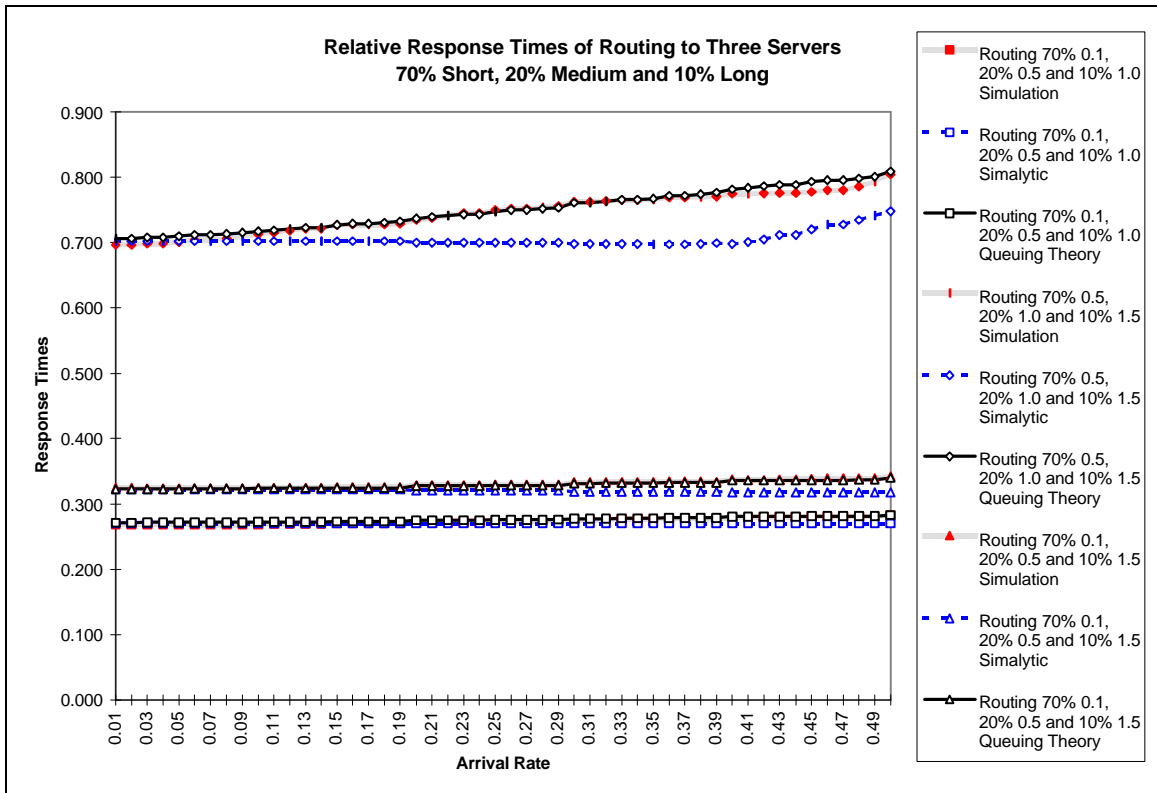


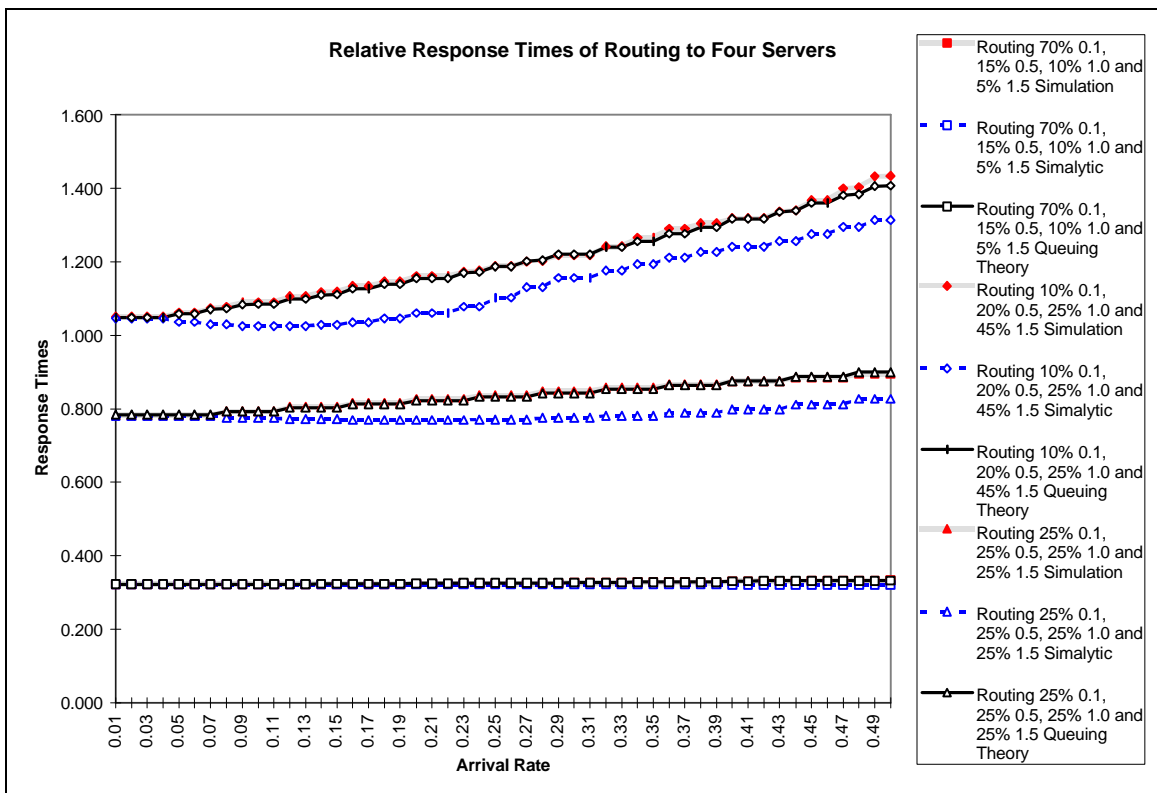
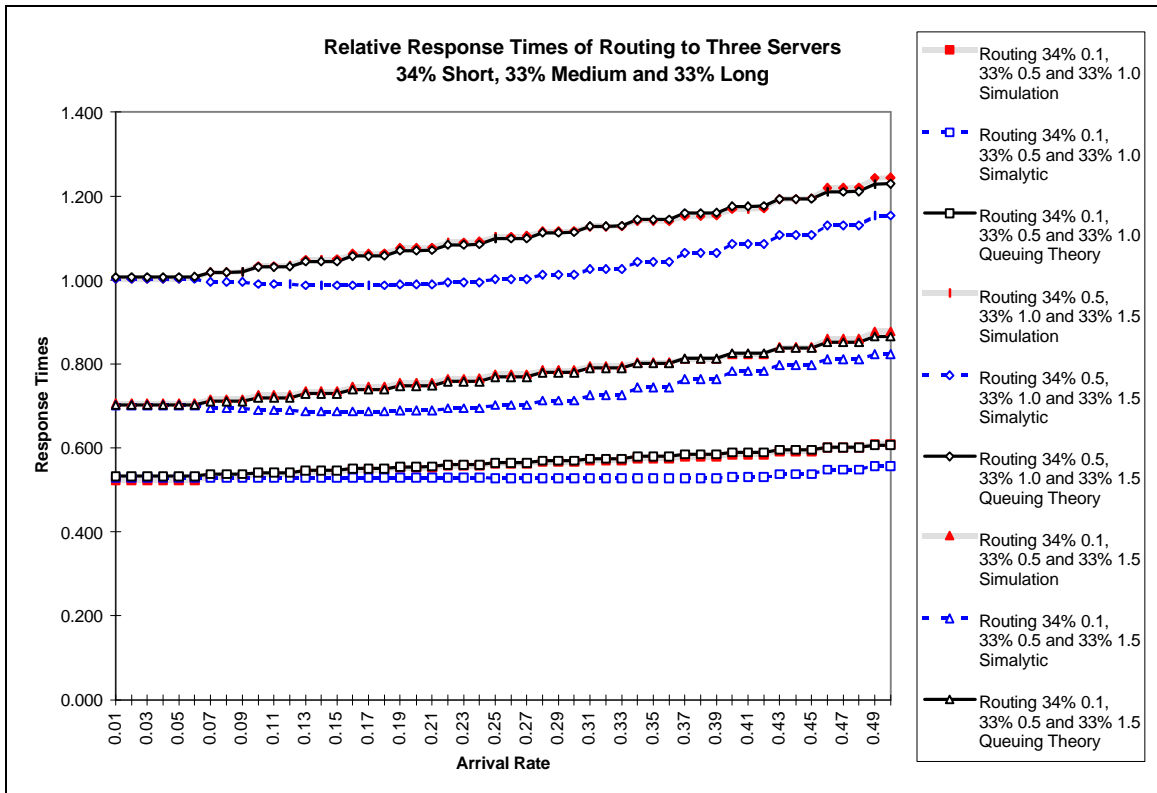












```

vsize, msize number of transactions    <== set by user
ServiceTime average server service time    <== set by user

IRT          an array containing the interarrival times between each
              transaction
Arv          an array containing the arrival time for each transaction
Dep          an array containing the departure time for each transaction
Srv          an array containing the service time for each transaction
              (exponential distribution)
SimulationResp  an array containing the response time for each transaction
              (service + wait time)
SRTimes      an array containing the response times for each transaction from
              the Simalytic function

rexp(x,y)    Built-in exponential function that returns an array of x elements
              with rate y.
QRT(r)       Queuing theory function returns response time for rate r and
              ServiceTime.
SRT(r)       Simulation function returns response time for rate r and
              ServiceTime.
Simalytic(ir) Transform function that returns response time for
              interarrival rate ir and ServiceTime.

Matrix indices are: i, n, r, m, m_i, m_n

Each of the major arrays (above) is a single column. The same names are used
prefixed with a lower case 'a' to show a multi-column version of the array
(i.e. IRT and aIRT).

Variables used to calculate the steps in the Simalytic function SrvRTime:
I1, I2, I3, I4, I5, I6, I7, I8 and I9 calculate the interarrival times for
the steps.
S1, S2, S3, S4, S5, S6, S7, S8 and S9 calculate the response times for the
steps.

The arrays vs1, vx and MaResp are used with the built-in functions loess and
interp to fit a curve to the results of the simulation.

msize is used for the simulation matrix to get a larger number to sample (same
usage a vsize).

ServiceTime := 5.0    Service rate for this analysis run.
vsize := 1000    Numbers of transactions for this analysis run.

rate is calculated as 1/ServiceTime to set the upper bound at the saturation
point.

rate :=  $\frac{1}{\text{ServiceTime}}$  - .01

Define a queuing theory function to calculate the response time when given an
arrival rate. The service time is a constant for this analysis run. (The
function only returns positive response times. The calculated time goes
negative when the server is saturated and the function returns infinity to
show this. Because of MathCAD rounding of intermediate values a small value
is added to avoid a divide by zero error.)

QRT(q_rate) := 
$$\left| \begin{array}{l} \text{rt} \leftarrow \frac{\text{ServiceTime}}{(1 - \text{q\_rate} \cdot \text{ServiceTime}) + .00000001} \\ \text{rt if } \text{rt} > 0 \\ \infty \text{ if } \text{rt} \leq 0 \\ \infty \text{ otherwise} \end{array} \right.$$


```

Figure 60 MathCAD Worksheet Listing

```

Set matrix indices (MathCAD Matrices are zero based)
i := 0..vsize - 1
n := 1..vsize - 1 (i for all elements, n for all but first)
r := .001,.002..rate (r sets the arrival rates to use for plots)
Create arrays of interarrival times (IRT) and service times (Srv)
IRT := rexp(vsize, rate) Srv := rexp(vsize, rate)
Create a series of step points for the transform function (interarrival times
for step points)
I1 :=  $\frac{1}{\text{rate}} \cdot 6$  I2 :=  $\frac{1}{\text{rate}} \cdot 4$  I3 :=  $\frac{1}{\text{rate}} \cdot 3$  I4 :=  $\frac{1}{\text{rate}} \cdot 2$  I5 :=  $\frac{1}{\text{rate}} \cdot 1.5$ 
I6 :=  $\frac{1}{\text{rate}} \cdot 1.25$  I7 :=  $\frac{1}{\text{rate}} \cdot 1.15$  I8 :=  $\frac{1}{\text{rate}} \cdot 1.1$  I9 :=  $\frac{1}{\text{rate}} \cdot 1.02$ 

(response times for step points)
S1 :=  $\text{QRT}\left(\frac{1}{I1}\right)$  S2 :=  $\text{QRT}\left(\frac{1}{I2}\right)$  S3 :=  $\text{QRT}\left(\frac{1}{I3}\right)$  S4 :=  $\text{QRT}\left(\frac{1}{I4}\right)$  S5 :=  $\text{QRT}\left(\frac{1}{I5}\right)$ 
S6 :=  $\text{QRT}\left(\frac{1}{I6}\right)$  S7 :=  $\text{QRT}\left(\frac{1}{I7}\right)$  S8 :=  $\text{QRT}\left(\frac{1}{I8}\right)$  S9 :=  $\text{QRT}\left(\frac{1}{I9}\right)$ 

(cumulative average index, ca_i=initital part, ca=main part )
Cavg := mean(submatrix(IRT, 0, i, 0, 0)) (cumulative average of all events)
(initialize the first part of the array to low ir value to represent warm-up
interval)
Cavgca_i := I1·1.1e

ra1 := 5          rai1 := ra1..vsize - 1   rai1_i := 0..ra1 - 1
ra2 := 10         rai2 := ra2..vsize - 1   rai2_i := 0..ra2 - 1
ra3 := 25         rai3 := ra3..vsize - 1   rai3_i := 0..ra3 - 1
Ravg1rai1 := mean(submatrix(IRT, rai1 - ra1, rai1, 0, 0))
Ravg1rai1_i := Ravg1rai1 (rolling average of ra1 events)
Ravg2rai2 := mean(submatrix(IRT, rai2 - ra2, rai2, 0, 0))
Ravg2rai2_i := Ravg2rai2 (rolling average of ra2 events)
Ravg3rai3 := mean(submatrix(IRT, rai3 - ra3, rai3, 0, 0))
Ravg3rai3_i := Ravg3rai3 (rolling average of ra3 events)

Define the transform function using the step point rates and service times
Simalytic(ir) := | ServiceTime if I1 < ir
                  | S1 if I2 < ir ≤ I1
                  | S2 if I3 < ir ≤ I2
                  | S3 if I4 < ir ≤ I3
                  | S4 if I5 < ir ≤ I4
                  | S5 if I6 < ir ≤ I5
                  | S6 if I7 < ir ≤ I6
                  | S7 if I8 < ir ≤ I7
                  | S8 if I9 < ir ≤ I8
                  | S9 otherwise

```

Figure 60 continued from previous page

Create arrays of transaction arrival times (Arv), transaction departure times (Dep) and transaction response times (SimulationResp) to simulate the system. This technique is taken from the MathCAD Statistics Electronic Book, Simulation section.

$Arv_0 := 0$ (the first arrival time is zero)

(each following time adds an interarrival time to the prior arrival time)

$Arv_n := Arv_{n-1} + IRT_{n-1}$

$Dep_0 := Arv_0 + Srv_0$ (the first departure time is the first arrival time plus the service time)

$Dep_n := \text{if}(Arv_n > Dep_{n-1}, Arv_n + Srv_n, Dep_{n-1} + Srv_n)$

(each following time is either the arrival time plus the service time (no queue) or the prior departure time plus the service time. the response times are the differences between transaction arrival times and departure times)

$SimulationResp := Dep - Arv$

$MaxATime := Arv_{vsize - 1}$

$MaxDTime := Dep_{vsize - 1}$

$ARate := \frac{vsize}{MaxATime}$

Create an array containing the response times for each transaction from the Simalytic function.

$SRTimes_n := \text{Simalytic}(IRT_n)$

$QSRTimes_n := \text{QRT}\left(\frac{1}{IRT_n}\right)$

$SR_Ravg1_n := \text{Simalytic}(Ravg1_n)$

$SR_Cavg_n := \text{Simalytic}(Cavg_n)$

$SR_Ravg2_n := \text{Simalytic}(Ravg2_n)$

$SR_Ravg3_n := \text{Simalytic}(Ravg3_n)$

Results are generated from a series of simulations at different arrival rates. The simulation builds columns of the array of response times at m different arrival rates between .1 and rate+.1. The average for each column is then plotted and a curve fit to the points. That curve is then laid on top of the queuing theory and Simalytic function curves from above.

(current values)

$m_{size} := 1000$ (number of columns in the matrix - similar to vsize above)

$m_j := 0..m_{size} - 1$

$m_n := 1..m_{size} - 1$ (indices for matrix)

(create an array of arrival rates)

$max_m := 49$ (number of columns on the matrix)

$m := 0..max_m$ (columns index for matrix)

$mrate_m := \text{rate}\left(\frac{m}{max_m} + .01\right)$

$z_m := 0$ (set a temporary vector to zeros)

Figure 60 continued from previous page

Set each column to be an array of exponential values

$aIRT^{<m>} := \text{rexp}(msize, mrate_m)$ (array of interarrival times at different means)
 $MaIRT_m := \text{mean}(aIRT^{<m>})$
 $aSrv^{<m>} := \text{rexp}(msize, \frac{1}{ServiceTime})$ (array of service times at the same mean)
 $MaSrv_m := \text{mean}(aSrv^{<m>})$
 $aArv^{<0>} := z$
 $aArv_{m,n,m} := aArv_{m,n-1,m} + aIRT_{m,n-1,m}$ (set the arrival times for each column)
 $aDep_{0,m} := aArv_{0,m} + aSrv_{0,m}$ (set the departure times for each column)
 $aDep_{m,n,m} := \text{if}(aArv_{m,n,m} > aDep_{m,n-1,m}, aArv_{m,n,m} + aSrv_{m,n,m}, aDep_{m,n-1,m} + aSrv_{m,n,m})$
 $aSimulationResp := aDep - aArv$ (set the response times for each column)
 $MaResp_m := \text{mean}(aSimulationResp^{<m>})$ (create a vector of the average response time of each column)

$mra1 := 15$
 $mrai1 := mra1..msize - 1$
 $mrai1_i := 0..mrai1 - 1$
 $mra2 := 50$
 $mrai2 := mra2..msize - 1$
 $mrai2_j := 0..mrai2 - 1$
 $mra3 := 100$
 $mrai3 := mra3..msize - 1$
 $mrai3_i := 0..mrai3 - 1$

$aCavg_{m_i,m} := \text{mean}(\text{submatrix}(aIRT^{<m>}, 0, m_i, 0, 0))$
 $aRavg1_{mrai1,m} := \text{mean}(\text{submatrix}(aIRT^{<m>}, mrai1 - mra1, mrai1, 0, 0))$
 $aRavg1_{mrai1_j,m} := aRavg1_{mrai1,m}$
 $aRavg2_{mrai2,m} := \text{mean}(\text{submatrix}(aIRT^{<m>}, mrai2 - mra2, mrai2, 0, 0))$
 $aRavg2_{mrai2_j,m} := aRavg2_{mrai2,m}$
 $aRavg3_{mrai3,m} := \text{mean}(\text{submatrix}(aIRT^{<m>}, mrai3 - mra3, mrai3, 0, 0))$
 $aRavg3_{mrai3_j,m} := aRavg3_{mrai3,m}$

$aSRTimesC_{m_i,m} := \text{Simalytic}(aCavg_{m_i,m})$
 $aSRTimesR1_{m_i,m} := \text{Simalytic}(aRavg1_{m_i,m})$
 $aSRTimesR2_{m_i,m} := \text{Simalytic}(aRavg2_{m_i,m})$
 $aSRTimesR3_{m_i,m} := \text{Simalytic}(aRavg3_{m_i,m})$
 $MaSimalyticResp_m := \text{Simalytic}(MaIRT_m)$
 $MaSimalyticRespC_m := \text{mean}(aSRTimesC^{<m>})$
 $MaSimalyticRespR1_m := \text{mean}(aSRTimesR1^{<m>})$
 $MaSimalyticRespR2_m := \text{mean}(aSRTimesR2^{<m>})$
 $MaSimalyticRespR3_m := \text{mean}(aSRTimesR3^{<m>})$

Figure 60 continued from previous page

The MathCAD built-in functions (loess and interp) are used to fit a line to the average response times of each column.

```

vx_m := m
vs1 := loess ( vx, MaResp, .35)
vs2 := loess ( vx, MaSimalyticResp, .35)
SimuResults ( x) := interp (vs1, vx, MaResp, vx_m)
SimalyResults ( x) := interp (vs2, vx, MaSimalyticResp, vx_m)
SimulationResults_m := SimuResults ( 1)
SimalyticResults_m := SimalyResults ( 1)
vs2C := loess ( vx, MaSimalyticRespC, .35)
vs2R1 := loess ( vx, MaSimalyticRespR1, .35)
SimalyResultsC ( x) := interp (vs2C, vx, MaSimalyticRespC, vx_m)
SimalyResultsR1 ( x) := interp (vs2R1, vx, MaSimalyticRespR1, vx_m)
SimalyticResultsC_m := SimalyResultsC ( 1)
SimalyticResultsR1_m := SimalyResultsR1 ( 1)
vs2R3 := loess ( vx, MaSimalyticRespR3, .35)
vs2R2 := loess ( vx, MaSimalyticRespR2, .35)
SimalyResultsR3 ( x) := interp (vs2R3, vx, MaSimalyticRespR3, vx_m)
SimalyResultsR2 ( x) := interp (vs2R2, vx, MaSimalyticRespR2, vx_m)
SimalyticResultsR3_m := SimalyResultsR3 ( 1)
SimalyticResultsR2_m := SimalyResultsR2 ( 1)
Write the results to file to be imported into Excel.
WRITE( McadData) := ServiceTime
APPEND( McadData) := ArrivalRate
WRITEPRN( McadD_U) := SimulationResults
WRITEPRN( McadD_Y) := SimalyticResults
WRITEPRN( McadD_YC) := SimalyticResultsC
WRITEPRN( McadD_YR1) := SimalyticResultsR1
WRITEPRN( McadD_YR2) := SimalyticResultsR2
WRITEPRN( McadD_YR3) := SimalyticResultsR3
WRITEPRN( McadD_mrate) := mrate
WRITEPRN( McadD_QT) := QRT( mrate_m)

```

Figure 60 continued from previous page

7.6 Appendix F: MathCAD/Simul8 Comparison Results

The charts in this section show the relationship between the Simul8 results and the MathCAD results, comparing the two modeling implementations. The charts show results for the scenarios in *Table 11 MathCAD/Simul8 Results Scenarios*. Each scenario is identified by the number of servers and the service time for each. The routing scenarios also indicate the percent of the transactions routed to each server. Each scenario was modeled at arrival rates of 0.1, 0.2,

0.3, 0.4, 0.45 and 0.5 transactions per second.

The charts beginning on page 207 show the relationship between the MathCAD models and the Simul8 models. A blank space separates each group of results for a given number of servers (one, two, three, or four).

Each group of three vertical bars represents the results of one modeling scenario implemented with Simul8.

There are three bars because

Series Scenarios	Routing Scenarios
Single server	Two servers
0.1	80% 0.1 & 20% 1.0
0.5	80% 0.5 & 20% 1.5
1.0	80% 0.1 & 20% 1.5
1.5	80% 0.1 & 20% 0.5
	20% 0.1 & 80% 1.0
Two servers	20% 0.5 & 80% 1.5
0.1 & 0.5	20% 0.1 & 80% 1.5
0.5 & 1.0	20% 0.1 & 80% 0.5
0.1 & 1.5	50% 0.1 & 50% 1.0
0.5 & 1.5	50% 0.5 & 50% 1.5
	50% 0.1 & 50% 1.5
Three servers	50% 0.1 & 50% 0.5
0.1, 0.5 & 1.0	
0.1, 0.5 & 1.5	Three servers
0.5, 1.0 & 1.5	70% 0.1, 20% 0.5 & 10% 1.0
	70% 0.5, 20% 1.0 & 10% 1.5
Four servers	70% 0.1, 20% 0.5 & 10% 1.5
0.1, 0.5, 1.0 & 1.5	20% 0.1, 30% 0.5 & 50% 1.0
	20% 0.5, 30% 1.0 & 50% 1.5
High service time	20% 0.1, 30% 0.5 & 50% 1.5
25.0	34% 0.1, 33% 0.5 & 33% 1.0
50.0	34% 0.5, 33% 1.0 & 33% 1.5
75.0	34% 0.1, 33% 0.5 & 33% 1.5
99.0	
	Four servers
	70% 0.1, 15% 0.5, 10% 1.0 & 5% 1.5
	10% 0.1, 20% 0.5, 25% 1.0 & 45% 1.5
	25% 0.1, 25% 0.5, 25% 1.0 & 25% 1.5
Table 11 MathCAD/Simul8 Results Scenarios	

Simul8 produces response time results for the $\pm 95\%$ confidence limits in addition to an average. The line on the chart represents the results of the MathCAD implementation from *Table 1 Mathematical Formulae Results* on page 93. The format of the charts is somewhat non-conventional in that the line does not represent a continuous series of related values. A line was used to connect the independent values for several reasons: to allow the reader to easily find each data point, to visually distinguish the MathCAD results from the Simul8 results, to provide an overall impression of the shape of the MathCAD results and to allow easy comparison between the two different groups of data points.

The last chart in this section is a similar comparison for the high service time servers shown in *Table 12 High Service Time Scenarios*, which also shows the service times

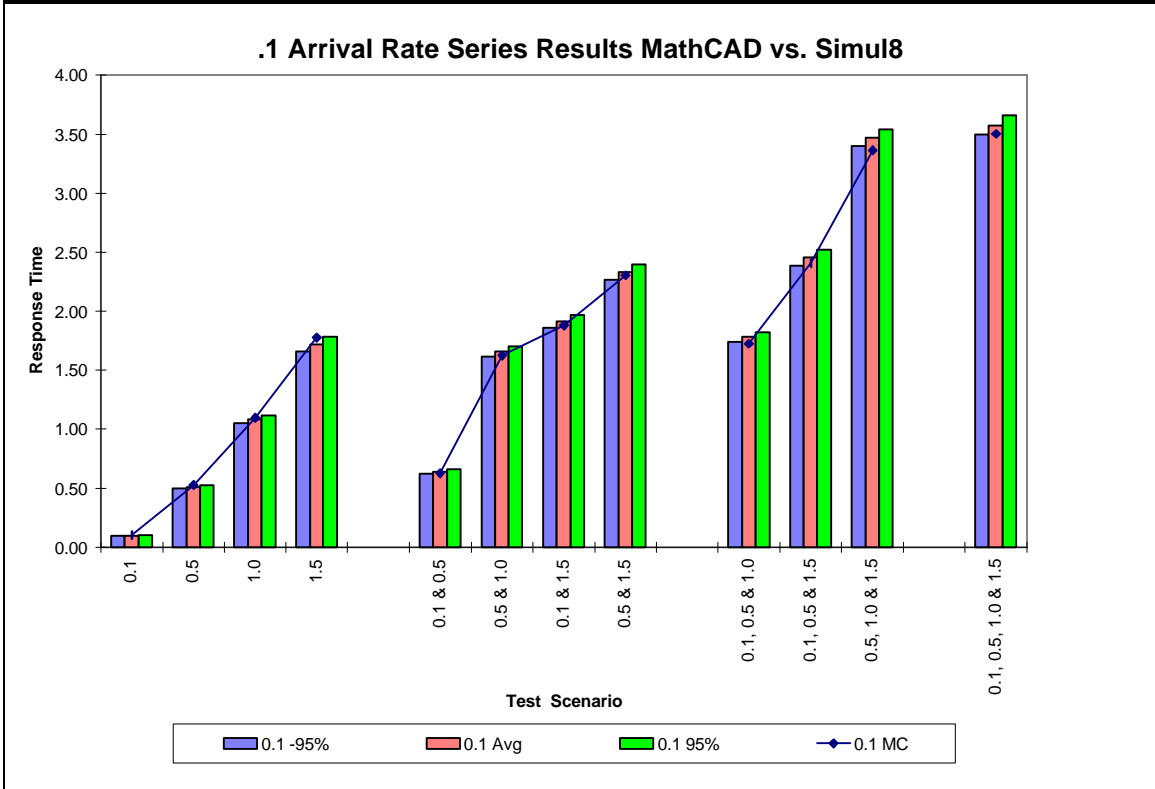
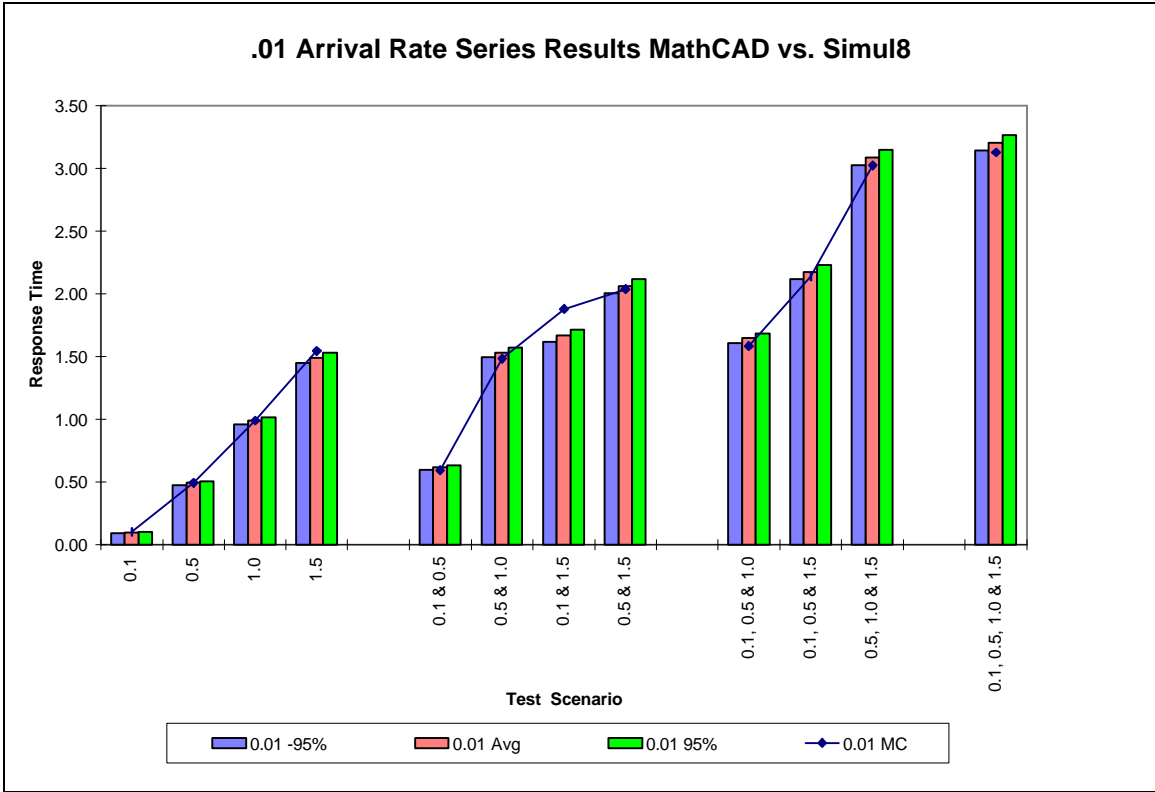
Arrival Rates	Service Times			
	25.0	50.0	75.0	99.0
0.001	✓	✓	✓	✓
0.005	✓	✓	✓	✓
0.009	✓	✓	✓	✓
0.0125	✓	✓	✓	
0.015	✓	✓		
0.02	✓	✓		
0.03	✓			
0.04	✓			

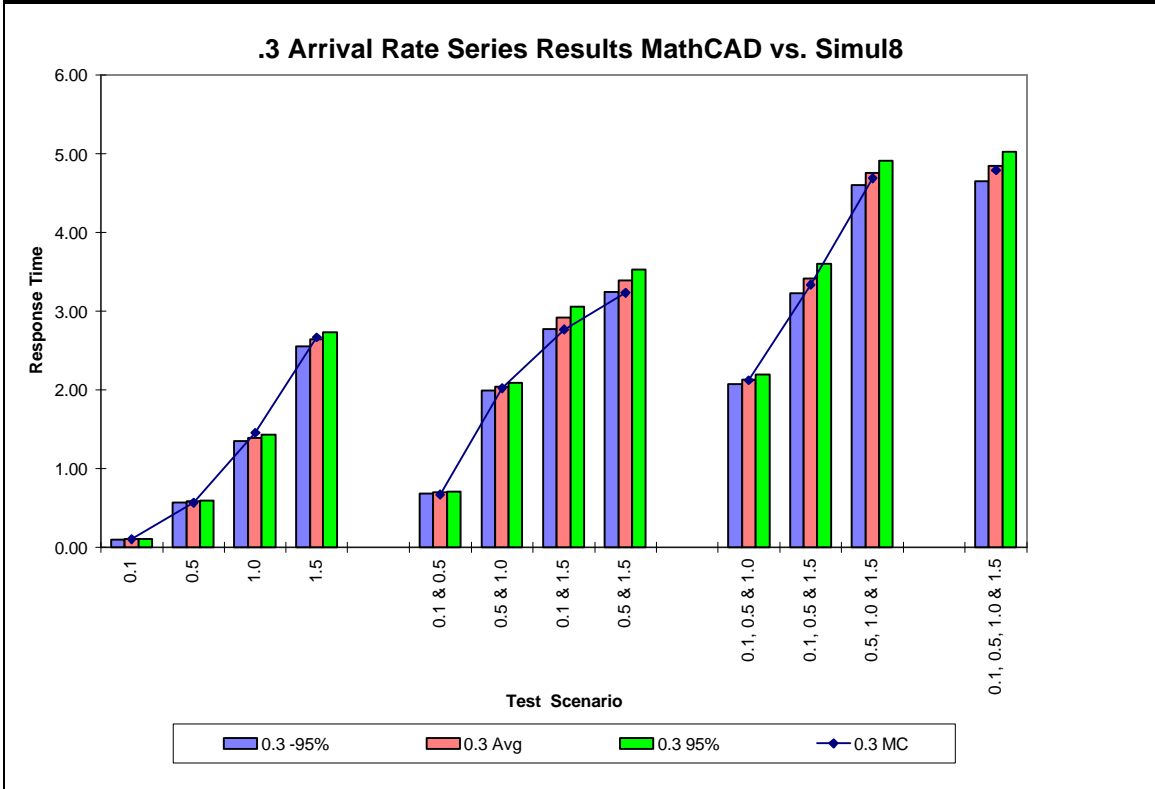
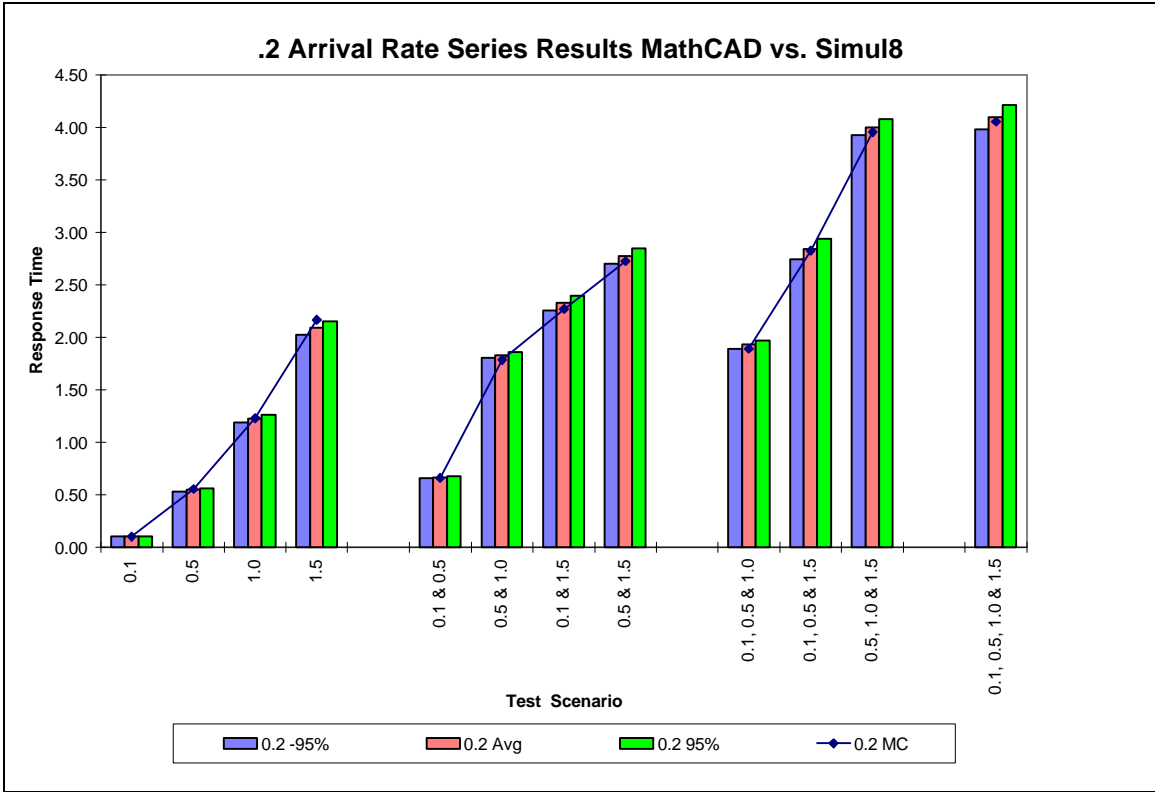
Table 12 High Service Time Scenarios

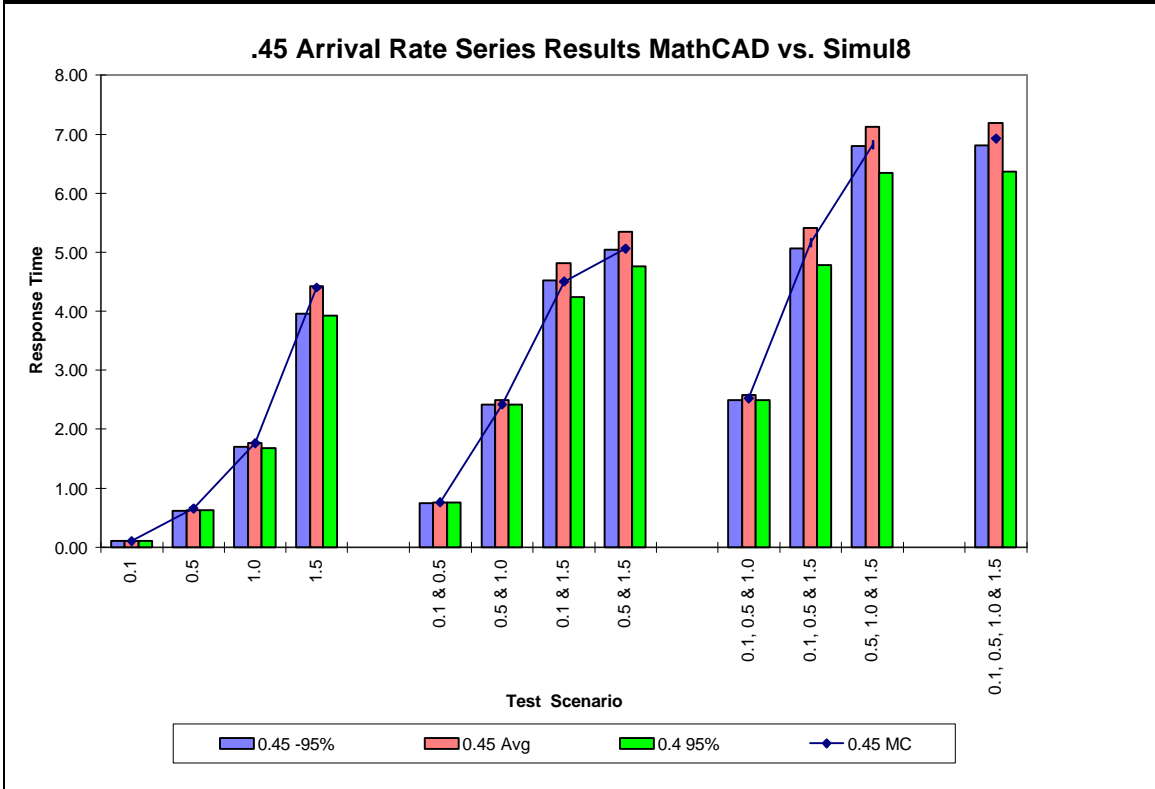
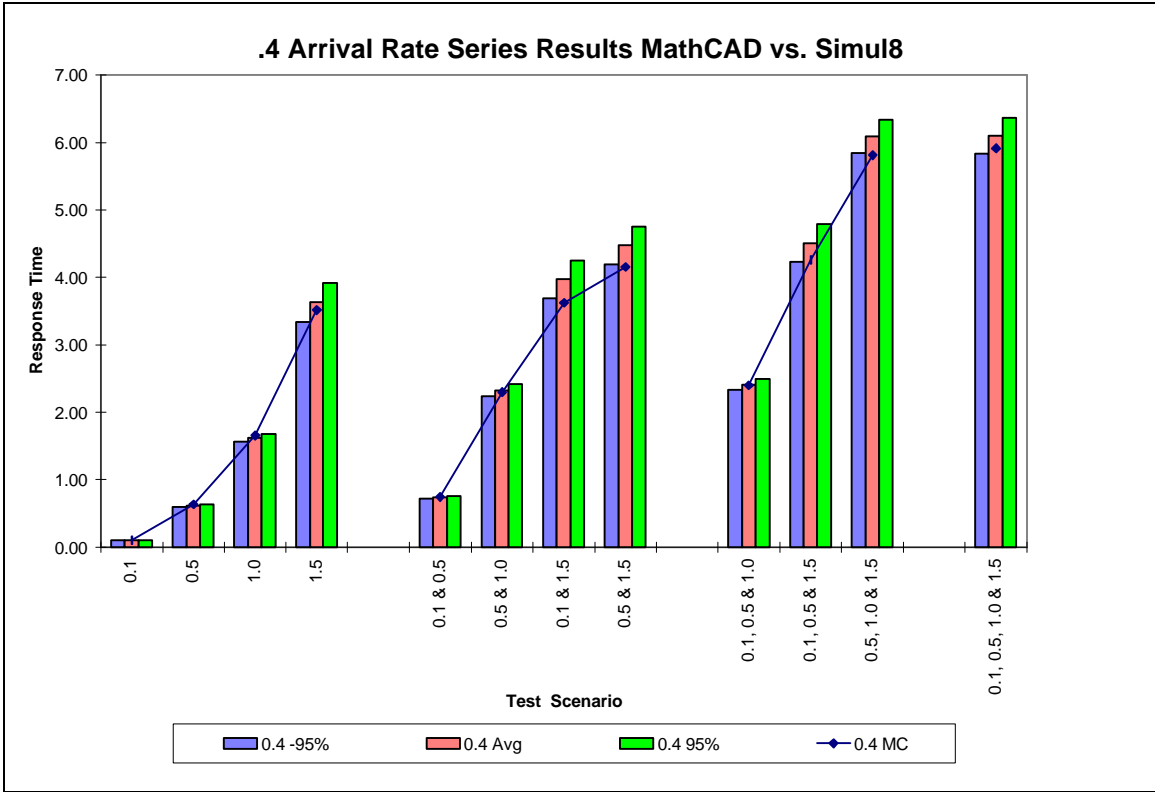
for these scenarios. A check mark (✓) indicates which arrival rates were modeled for each service time. A blank indicates that the model saturated and did not produce usable results.

In most cases the MathCAD result is very close to the average response time and between the $\pm 95\%$ confidence limits. The few cases where the MathCAD results were significantly different were assumed to be caused by distribution variations in the MathCAD results because they were generated by single run trials. To determine if this was a reasonable assumption, a number of scenarios showing the difference were modeled with

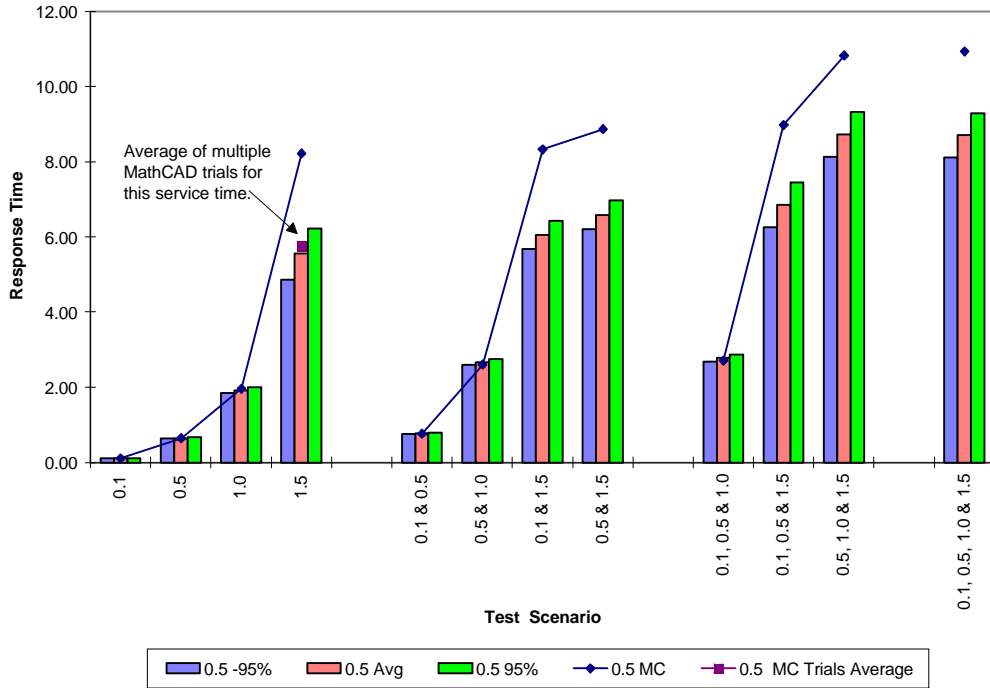
MathCAD using multiple run trials. These selected scenarios were: a single server with a service time of 1.5 and an arrival rate of 0.5, all of the scenarios for a single server with a service time of 25.0 and all of the scenarios for a single server with a service time of 75.0. The first was selected because most of the differences appeared to occur in a scenario that included this server. The last two were selected because they showed differences both significantly higher than and lower than the Simul8 results. Additional MathCAD model results were generated using seven model runs per trial, each with a different random number seed. These results were very consistent with the Simul8 model results and strongly indicate that the assumption was correct. Each of the multi-run MathCAD trials require significant effort and resources and it was determined that there was no additional benefit to conducting additional trials. These selected multiple run trials are indicated on the *.5 Arrival Rate Series Results* and the *High Service Time Results* charts as the series labeled *MC Trials Average*.



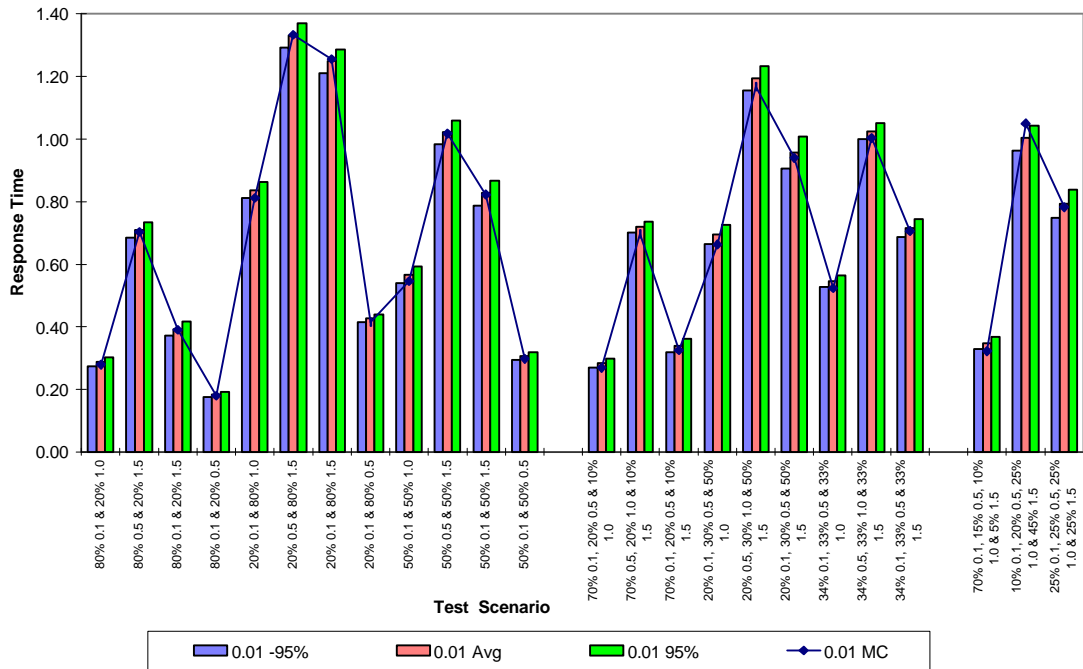




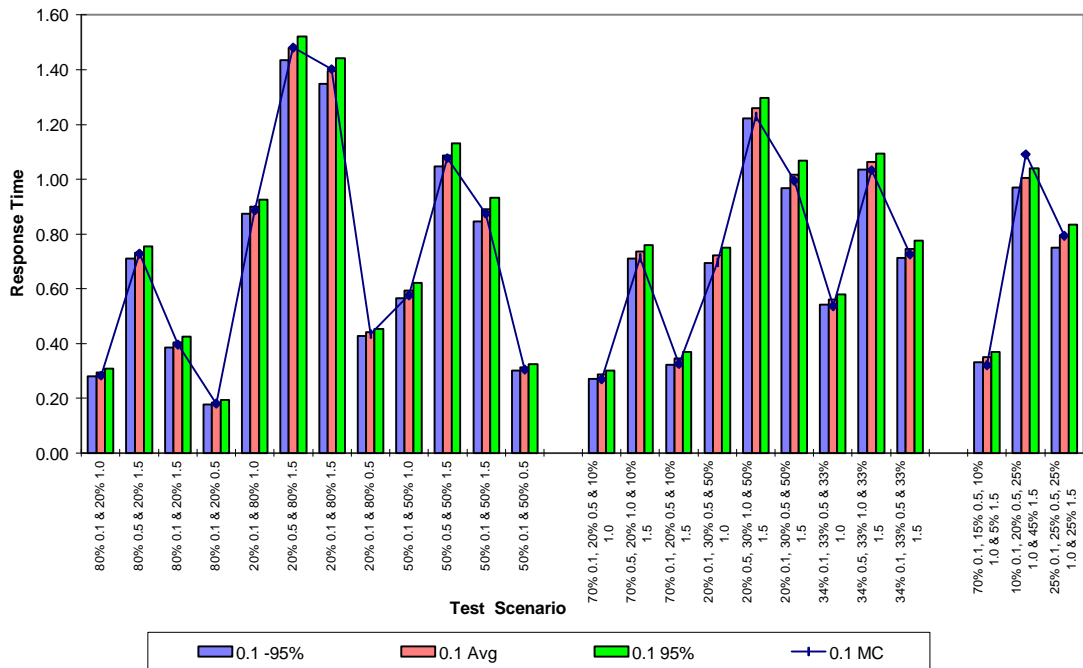
.5 Arrival Rate Series Results MathCAD vs. Simul8



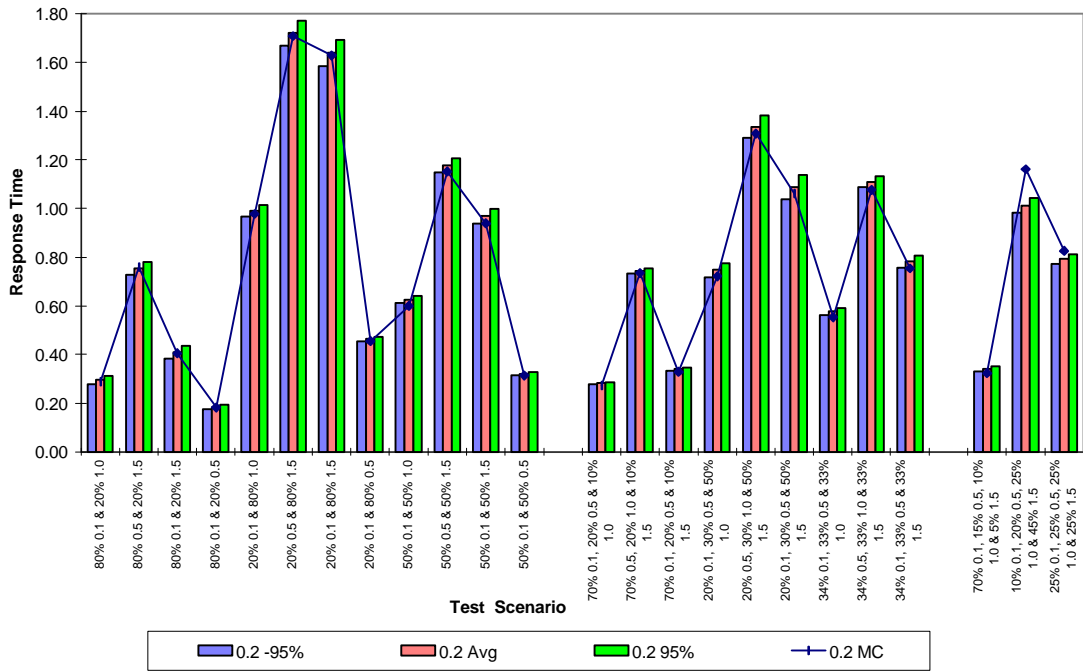
.01 Arrival Rate Routing Results MathCAD vs. Simul8



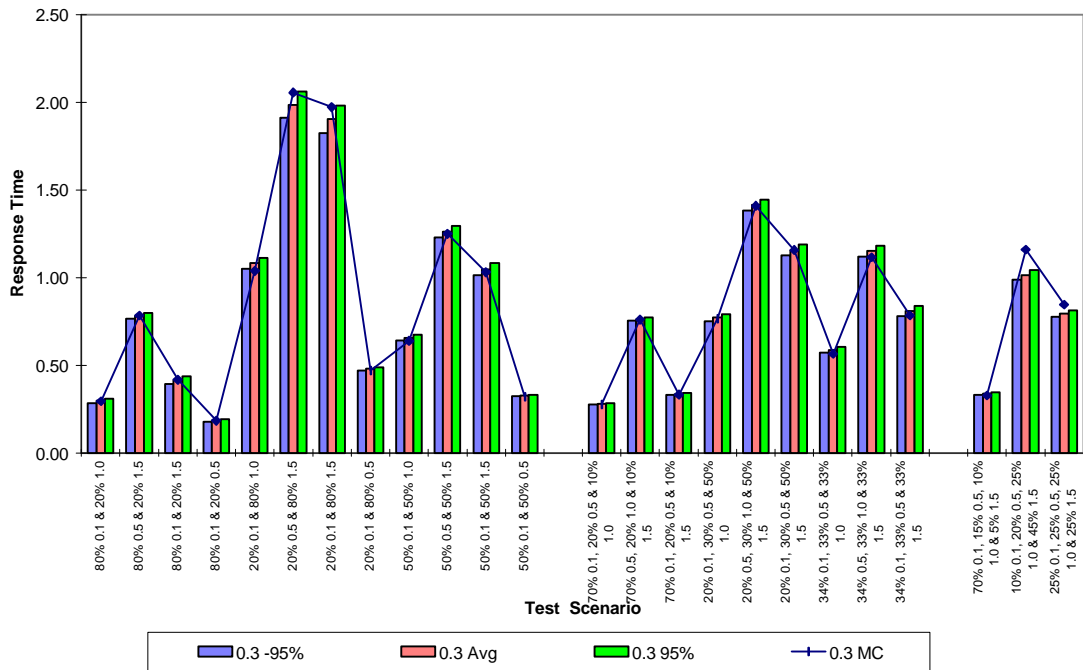
.1 Arrival Rate Routing Results MathCAD vs. Simul8



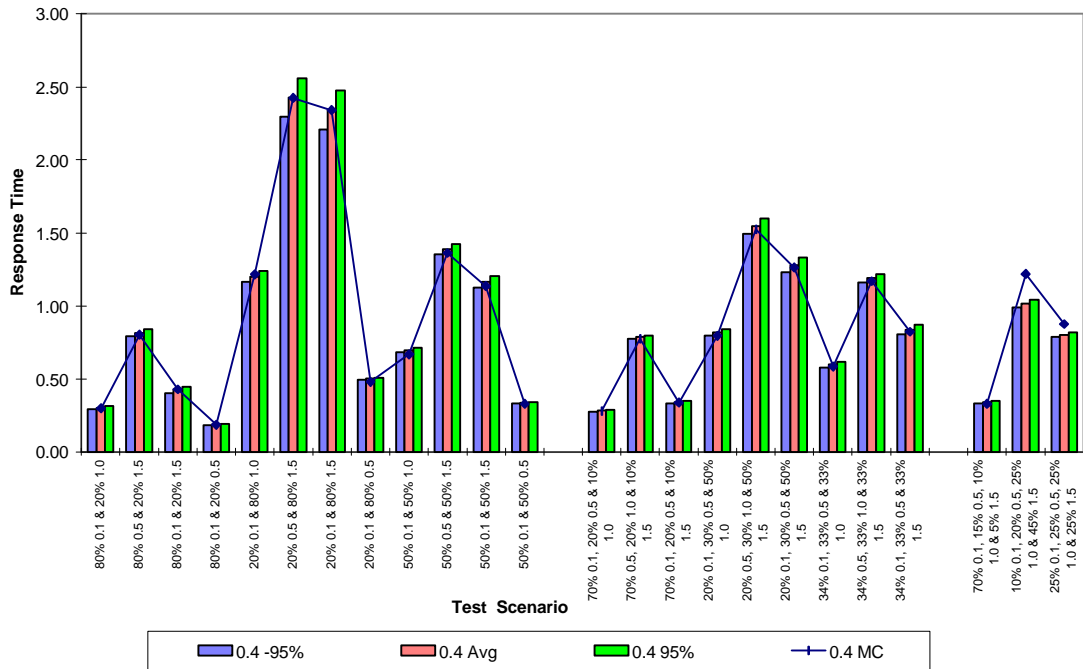
.2 Arrival Rate Routing Results MathCAD vs. Simul8



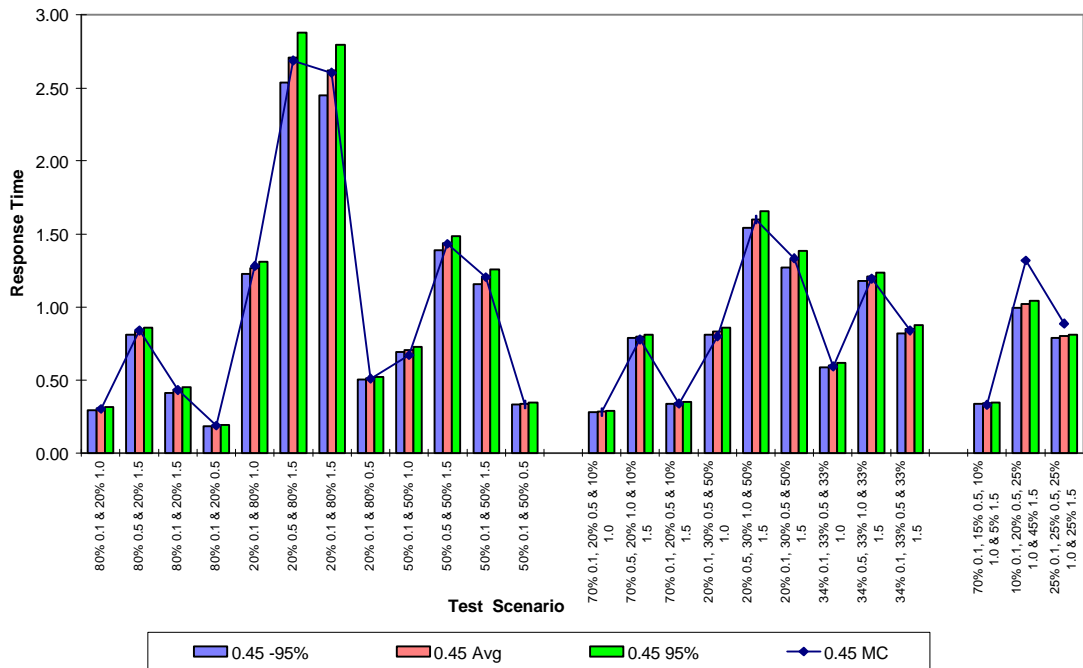
.3 Arrival Rate Routing Results MathCAD vs. Simul8



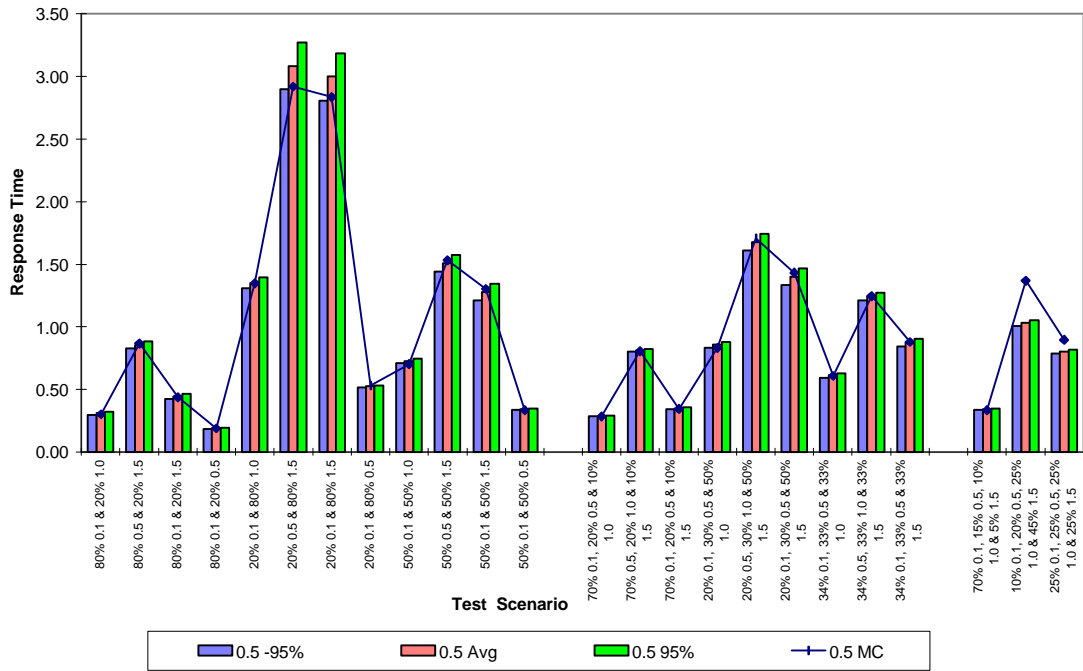
.4 Arrival Rate Routing Results MathCAD vs. Simul8



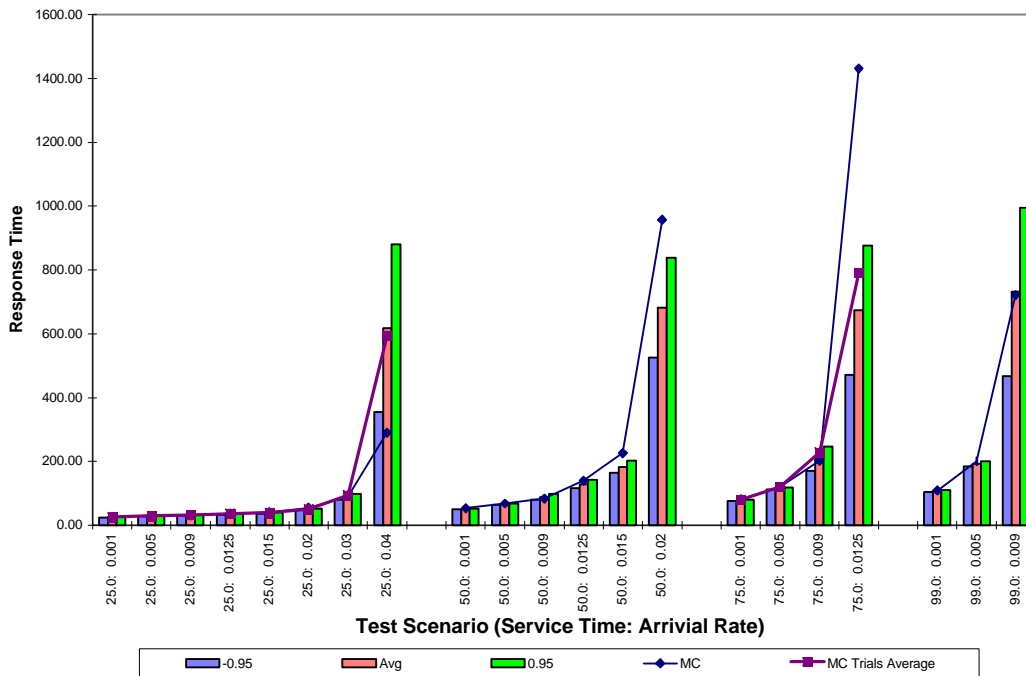
.45 Arrival Rate Routing Results MathCAD vs. Simul8



.5 Arrival Rate Routing Results MathCAD vs. Simul8



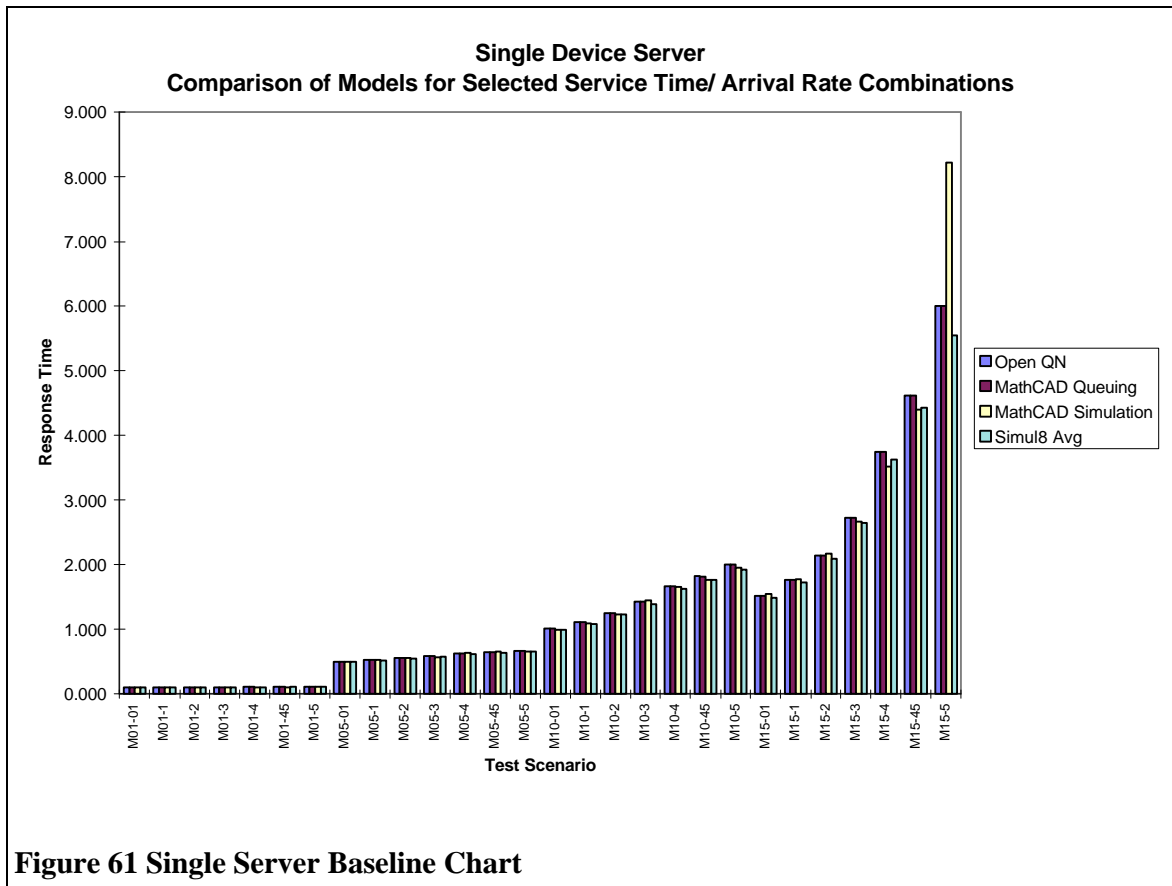
High Service Time Results MathCAD vs. Simul8



7.7 Appendix G: Tool Baseline Comparison Results

This final set of comparisons establishes the validity of the Simul8 tool to provide baseline simulation results by comparing those results to both the MathCAD results (simulation and queuing theory) and the OpenQN results. These comparisons are shown on two charts.

Figure 61 Single Server Baseline Chart compares the Simul8 results for single server model test scenarios to all three sets of the other results for the same environment, MathCAD simulation, MathCAD queuing theory and OpenQN. This chart shows a very close correlation between all four of the groups of results. The **Test Scenario** names began with **M** (for model) and reflect the arrival rate and service time for each scenario in the



format **Mss-aa**, where **ss** is the service time with an assumed decimal between the digits (**01**=0.1) and **aa** is the arrival rate with an assumed decimal before the digits (**45**=.45).

The large difference between the MathCAD simulation result and the other results for the far right test scenario (**M15-5**) was due to the single run trial for this model as discussed in section 7.5 *Appendix E: MathCAD Formulae Results Charts* on page 180. The anomaly disappeared when the MathCAD model was rerun using a multi-run trial. That result is not reflected in this chart because it only shows the single run trial for consistency across the chart.

Figure 62 Multi-device Server Baseline Chart compares the Simul8 results for system (multi-device server) test scenarios to the same environment modeled with

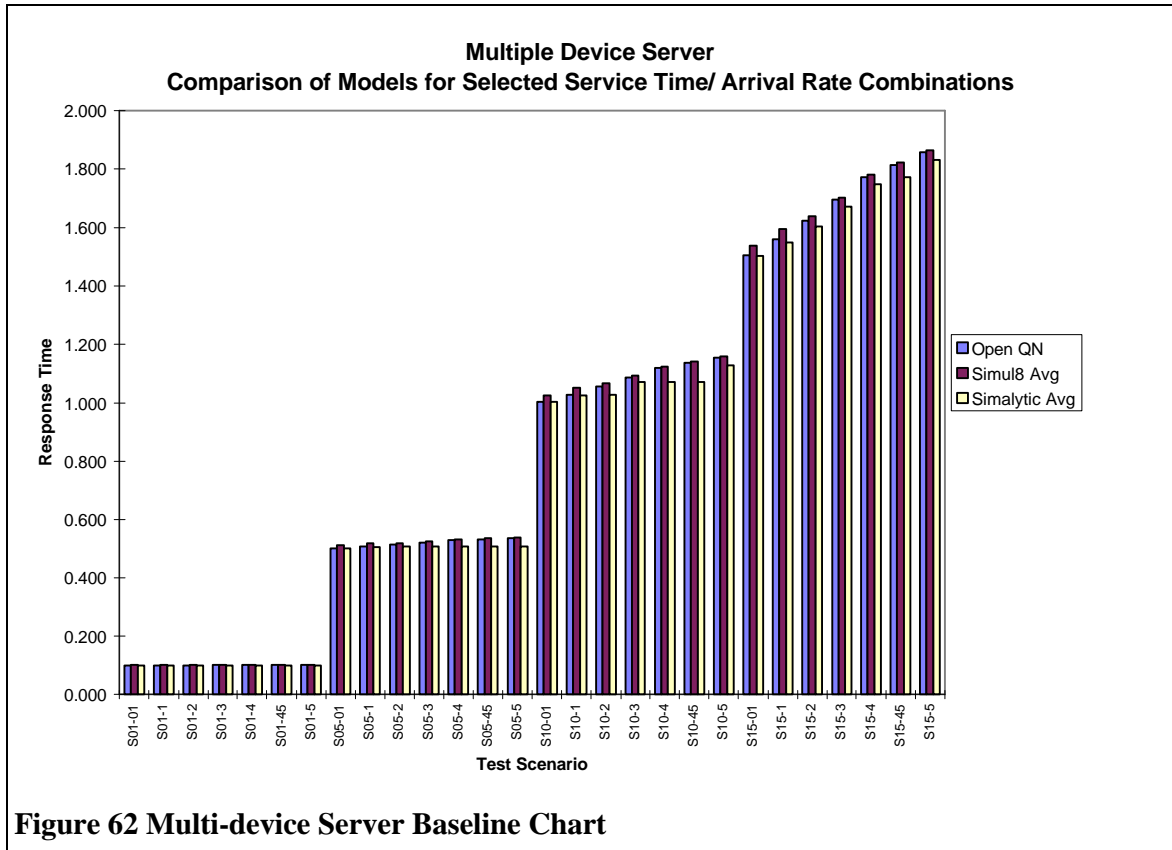


Figure 62 Multi-device Server Baseline Chart

OpenQN and with Simul8 using a Simalytic Function (these models were not implemented in MathCAD). The **Test Scenario** names began with **S** (for server) and reflect the arrival rate and service time for each scenario in the format **Sss-aa**, where **ss** is the service time with an assumed decimal between the digits (**01**=0.1) and **aa** is the arrival rate with an assumed decimal before the digits (**45**=.45). This chart shows the near identical results for all three of the modeling techniques, thus establishing the Simul8 modeling tools as valid to use for the baseline for comparisons with Simalytic Models and that Simalytic Models produce equivalent results for individual nodes.

7.8 Appendix H: Glossary

CICS	Customer Information Control System. TP system developed for IBM System/360 mainframe computer systems. Currently CICS is available from IBM for a number of operating systems (and the respective hardware platforms), including AIX, MVS, Windows NT and OS/2.
DB2	Data Base 2. A relational data base system developed for IBM System/360 mainframe computer systems. Currently DB2 is available from IBM for a number of operating systems (and the respective hardware platforms), including AIX, MVS, Windows NT and OS/2
GUI	Graphical User Interface. The interface between the user and a program that uses the computer's graphics capabilities to make the application easier to use, generally by representing program functions with icons the user clicks on with a mouse instead of typing application commands.
Informix	Informix is an object-relational database management system that is available on a large number of Unix systems from Informix Systems, Inc.
MPP	Massively Parallel Processing. Processors constructed of large numbers of CPU's and connected using some type of communications paths (generally the CPU's do not have access to the same memory and disks, i.e. "shared nothing").

NFS	Network File System. Protocol used to allow one system to offer disk files to be shared by other systems over a network.
node	A computer system (including CPU, disk and network communications) used as a server in a client/server environment. In the context of a model, the term is used to mean a combination of server and queue or a submodel and queue that represent some relatively independent component of an application or provide service to an application.
OLTP	On-Line Transaction Processing. Generic term for developing an application design implemented with end-user terminals or workstations and a TP system.
responsiveness	The quality of something to respond to a request from a particular viewpoint. The responsiveness of a device can be its service time from the hardware viewpoint or the response time from the application viewpoint. The responsiveness of an application or transaction is the user's perception of response time. Responsiveness is good based on what is expected or required by that viewpoint.
SPE	Software Performance Engineering. The technique developed by Dr. Connie Smith as an early lifecycle method for designing and constructing software systems to meet the required performance objectives by identifying performance problems in the design phase (Smith 1990).
TP	Transaction Processor. A system component, generally considered an extension of the operating system, that provides a programming API, a

	<p>user interface and resource management functions to schedule and execute transactions. TP systems generally provide services to control terminals or workstations so the transaction program can be developed independent of the required communications protocols and screen formatting commands.</p>
TPC	<p>Transaction Processing Performance Council. A series of benchmarks developed to provide objective comparisons between processors from different manufacturers. The individual benchmarks are designated by a suffix letter, such as TPC-A, TPC-B, TPC-C and TPC-D.</p>
Tuxedo	<p>A transactions processing system developed for Unix systems. Also referred to as middleware because it generally is implemented in the middle layer of a three tiered architecture. Currently Tuxedo is available from BEA Systems, Inc. for both Microsoft NT and a large number of Unix operating systems (and the respective hardware platforms). Client software is also available on MS-DOS, Macintosh OS , Windows 95 and OS/2.</p>
transaction	<p>A relatively small independent unit of work entered into the system by an end-user to receive some information as a response in near real-time. Transactions include entering an order at a terminal (business transaction), an SQL command (database transaction), some key-strokes followed by a carriage-return (interactive transaction)—whatever is meaningful from the end-user’s point-of-view.</p>

	<p>A database transaction is considered to be self-contained, or atomic, and is guaranteed to complete successfully or not at all. If an error prevents a partially-performed transaction from proceeding to completion, it must be "backed out" to prevent the database being left in an inconsistent state.</p>
trial	<p>A series of model executions, or runs, used to generate a single result, usually averaged. Trials are used to reduce the influence of variations in arrival and service distributions in simulation models.</p>

8. Bibliography

- Abate, Joseph and Ward Whitt. 1994. Transient behavior of the M/G/1 workload process. Operations Research 42 (4): 750-764.
- ACR. CAPS . Applied Computer Research, Phoenix, AZ (800) 234-2227.
- Ahn, M. S. and T. G. Kim. 1994. A Framework for Hybrid Modeling/Simulation of Discrete Event Systems. In Proceedings. The Fifth Annual Conference on AI, Simulation, and Planning in High Autonomy Systems.
<http://www.computer.org/conferen/ais94/ais94toc.html>: University of Florida, The Advanced Research Projects Agency (ARPA), and The Army Research Office (ARO).
- Allen, Arnold O. 1978. Probability, Statistics and Queueing Theory With Computer Science Applications. New York, NY: Academic Press.
- Allen, Arnold O. 1990. Probability, Statistics and Queueing Theory With Computer Science Applications. San Diego, CA: Academic Press.
- Anonymous. 1990. Estimating Workloads. Capacity Management Review 18 (7): 3-4.
- Anonymous. 1992. Capacity Planning: More than Sizing CPUs. Capacity Management Review 20 (1): 7-8.
- Anonymous. 1992,1994. The American Heritage Electronic Dictionary: Houghton Mifflin Company & Softkey International Inc.
- Anonymous. 1996. Simulation. HTML: <http://jaguar.is.unimt.mt/staff/jskl/simul.html>.
- Artis, H. Pat. 1994a. DASD Subsystems: Evaluating the Performance Envelope. In Transactions. Computer Measurement Group: CMG, Inc.
- Artis, H. Pat. 1994b. A Graphical Paradigm for the Design and Analysis of Distributed Systems. In Proceedings. Computer Measurement Group: CMG, Inc.
- AST. QASE . Advanced System Technologies, 12200 E. Briarwood Ave., Suite 260, Englewood, CO 80112.
- B&B. CMF/MODEL . Boole & Babbage, Inc., 510 Oakmead Parkway, Sunnyvale, CA 94086.
- Baker, M. and W. Shih. 1992. Performance prototyping: a simulation methodology for software performance engineering. In Proceedings. CompEuro Proceedings, IEEE Computer Systems and Software Engineering.,: 624-9: IEEE.

- Barroso, Luiz André and Michael Dubois. 1995. Performance Evaluation of the Slotted Ring Multiprocessor. IEEE Transactions on Computers 44 (7): 878-890.
- Baylor, S. Johnson, C. Benveniste, and L. J. Boelhouwer. 1994. A methodology for evaluating parallel I/O performance for massively parallel processors. In Proceedings. 27th Annual Simulation Symposium Proceedings,: 31-40. La Jolla, CA: IEEE Computer Society Press.
- BGS. Best/1 and Crystal . BGS Systems, Inc., 128 Technology Center, Waltham, MA 02254.
- BGS. 1996. UNIX Client Server Sizing and Performance: BGS Systems, Inc. White Paper.
- Biswas, Prabuddha and Christine Pietras. 1988. Performance Analysis and Modeling of an Office Server in a Distributed PC Environment. In Proceedings. Computer Measurement Group: CMG, Inc.
- Bleker, Julius. 1994. Unix Capacity Planning Hints and Tips. In Proceedings. Computer Measurement Group: CMG, Inc.
- Brunetto, Anthony F. 1984. Benchmarking Decisions - A Management Tutorial. In Proceedings. Computer Measurement Group,: 755-761: CMG, Inc.
- Buzen, Jeffrey P. 1984. A Simple Model of Transaction Processing. In Proceedings. Computer Measurement Group,: 835-839: CMG, Inc.
- Buzen, J.P. 1995. Performance Assurance Tools for LAN Based Client/Server Environments. In British Proceedings. Computer Measurement Group: CMG, Inc.
- Buzen, J.P. and D. Potier. 1977. Accuracy Of Exponential Assumptions In Closed Queueing Models. In Proceedings. Computer Measurement Group: CMG, Inc.
- Buzen, J.P. and A.W. Shum. 1994. Architecture and Performance of S/390 Parallel Transaction Servers. In Proceedings. Computer Measurement Group: CMG, Inc.
- CA-Legent. OptiModel . Legent Corporation, PO Box 9345, Framington, MA 01701.
- Carrasquilla, Uriel. 1991. Benchmarking a Co-operative Application. In Proceedings. Computer Measurement Group,: 333-342: CMG, Inc.
- Carroll, John R. and William S. Cool. 1994. Surviving My First Client/Server, Distributed Systems Capacity Study. In Proceedings. Computer Measurement Group: CMG, Inc.

- Chandy, K. and E. Sauer. 1978. Approximate methods for analyzing queuing network models of computing systems. ACM Computing Surveys 10 (3).
- Confrey, Tom. 1996. Perspectives on modeling and capacity planning: a conversation with Dr. Jeff Buzen. Capacity Management Review 24 (1): 1-8.
- Cortois, P. 1975. Decomposability, instabilities and saturation in multiprogramming systems. Communications of the ACM 18 (7).
- Domanski, Bernard. 1983. System Performance Management and Capacity Planning. In Proceedings. Computer Measurement Group: CMG, Inc.
- Domanski, Bernard. 1995. Capacity Management for Client-Server Architectures. Enterprise Systems Journal 10 (11): p78(6).
- Dwyer, Jim. 1984. Modeling Large Ims Systems Using Commercial Analytic Queueing Model Software Products. In Proceedings. Computer Measurement Group: CMG, Inc.
- Eick, Stephen G., William A. Massey, and Ward Whitt. 1993. The physics of the Mt/G/infinity queue. Operations Research 41 (4): 731-742.
- Elmqvist, Hilding, Francois E. Cellier, and Martin Otter. 1996. Object-Oriented Modeling of Hybrid systems. Postscript:
<http://www.ece.arizona.edu:8001/~cellier/delft.html>.
- Ferry, Dennis L. and Fred M. Richards. 1989. What Makes a Good Benchmark? In Proceedings. Computer Measurement Group: 820-830: CMG, Inc.
- Fishwick, Paul A. 1995. Computer Simulation: The Art and Science of Digital World Construction. <http://www.cis.ufl.edu:80/~fishwick/tr/tr95-017.html>: University of Florida. TR95-017.
- Fishwick, Paul A., Gyooseok Kim, and Jin Joo Lee. 1996. Improved Decision Making through Simulation Based Planning. <http://www.cis.ufl.edu:80/~fishwick/tr/tr96-014.html>: University of Florida. TR96-014.
- Ghosh, Deb, Ishwar Murthy, and Allen Moffett. 1992. File Allocation Problem: Comparison of Models with Worst Case and Average Communication Delays. Operations Research 40 (6): 1074-1085.
- Gilmore, Martha R. 1980. Capacity Planning Methodologies. In Proceedings. Computer Measurement Group: CMG, Inc.

- Goettge, Robert T. 1990. PEDAS: An Expert System for Performance Engineering of Time-Critical Software. In Proceedings. Computer Measurement Group: CMG, Inc.
- Gunther, Neil J. 1995. Performance Analysis and Capacity Planning for Datacenter Parallelism. In Proceedings. Computer Measurement Group: CMG, Inc.
- Hanna, Carolyn. 1988. A Production Control Model Of On-Line Systems A Capacity Planning Overview. In Proceedings. Computer Measurement Group,: 592-598: CMG, Inc.
- Harbitter, Alan. 1990. Hybrid Techniques in Benchmarking: A Scalable Approach. In Proceedings. Computer Measurement Group,: 387-394: CMG, Inc.
- Harrison, Ken. 1990. Forecasting from Business Volumes. Capacity Management Review 18 (7): 1-2.
- Hatheson, Amanda. 1995. Two Unix Client/Server Capacity Planning Case Studies. In British Proceedings. Computer Measurement Group: CMG, Inc.
- Hersh, William. 1978. Capacity Planning - Scape. In Proceedings. Computer Measurement Group: CMG, Inc.
- ISM. ISM/CP . The Information Systems Manager, Inc., 301 Broadway, Bethlehem, PA 18015.
- Kobayashi, Hisashi. 1981. Modeling and Analysis: An Introduction to System Performance Evaluation Methodology. The Systems Programming Series. Reading, MA: Addison-Wesley Publishing Company.
- Lehmann, Axel. 1984. A Multi-stage Hierarchical Performance Evaluation Concept - Design, Application and Experiences. In Proceedings. Computer Measurement Group: CMG, Inc.
- Linthicum, David S. 1997. Performance Anxiety. DBMS Magazine 10 (8): reprint.
- Lipovich, G. Jay. 1995. Performance assurance; a new paradigm for performance and capacity planning. Enterprise Systems Journal 9 (11): 48-55.
- Liu, Hong, Jun-Cheol Park, and Raymond E. Miller. 1996. On Hybrid Synthesis for Hierarchical Structured Petri Nets. College Park, MD: Univ. of Maryland. CS 3628.
- MathSoft. Mathcad 6.0. MathSoft Inc., Cambridge MA.

- Menascé, D., V. Almeida, and L. Dowdy. 1994. Capacity Planning and Performance Modeling: from mainframes to client-server systems. Englewood Cliffs, New Jersey: Prentice Hall.
- Menascé, Daniel A. and Duane Dregits. 1995. On the design and implementation of a capacity management tool for lan environments. In Proceedings. Computer Measurement Group: CMG, Inc.
- Menascé, Daniel A. , Duane Dregits, Robert Rossin, and Don Gantz. 1995. A federation-oriented capacity management methodology for LAN environments. In Proceedings. Computer Measurement Group: CMG, Inc.
- Miller, Victor Todd and Paul A. Fishwick. 1992. Hybrid Heterogeneous Hierarchical Models for Knowledge-Based Autonomous Systems. Gainesville, Florida: University of Florida. TR92-038.
- Mills, Terry L. 1991. Capacity Planning For Managers: an Implementation Architecture. In Transactions. Computer Measurement Group: CMG, Inc.
- Mohr, Jeffrey M. 1979. The Time Varying Nature of Computer Workloads. In Proceedings. Computer Measurement Group,: 305-319: CMG, Inc.
- MS. CSIM . Mesquite Systems, Inc., 8920 Business Park Sr., Austin, TX 78759.
- MSI. Pro-Sim and ATHENE . Metron Systems Inc., 352 Hungerford Dr., Suite 400, Rockville, MD 20850.
- Mullen, J. William. 1988. MVS/XA Performance and Resource Measurement Facility(RMF) Data Analysis. In Proceedings. Computer Measurement Group,: 983-1002: CMG, Inc.
- Norton, Tim R. 1995. Modeling Distributed Transaction Response Times As Impacted By In-Storage Buffer Accesses. In Proceedings. Computer Measurement Group,: 1-8: CMG, Inc.
- Norton, Tim R. 1996. Simalytic Enterprise Modeling: The Best of Both Worlds. In Proceedings. Computer Measurement Group,: 1-12. San Diego, CA: CMG, Inc.
- Norton, Tim R. 1997a. Simalytic Hybrid Modeling: Planning the Capacity of Client/Server Applications. In Proceedings. 15th IMACS World Congress,: 1-6. Berlin, Germany: International Association for Mathematics and Computers in Simulation.
- Norton, Tim R. 1997b. Simalytic Modeling: A Hybrid Technique for Client/Server Capacity Planning. In Proceedings. Summer Computer Simulation Conference,: 1-6. Arlington, Virginia: Society for Computer Simulation.

- PES. SPE-ED . Performance Engineering Services, P.O. box 2640, Santa Fe, NM 87504.
- Pfeiff, Bill. 1993. Basics Of Queueing Theory For Capacity Planning. In Proceedings. Computer Measurement Group: CMG, Inc.
- Place, Jerry P. 1986. A Capacity Planning Model Relating the Degree of Multiprogramming, Page Fault Rate and CPU Utilization in a Large Computing System. In Proceedings. Computer Measurement Group: CMG, Inc.
- Pooch, Udo W. and James A. Wall. 1993. Discrete Event Simulation: A Practical Approach. Edited by Udo W. Pooch. Computer Engineering Series. Boca Raton: CRC Press.
- Pooley, Rob. 1995. Performance Analysis Tools in Europe. Informationstechnik und Technische Informatik 37 : 10-16.
- Pooley, R. J. 1992. The Integrated Modelling Support Environment a New Generation of Performance Modelling Tools. In Computer Performance Evaluation, ed. G. Balbo and G. Serazzi:1--15: Elsevier Science Publishers.
- ProModel. ProModel . ProModel Corp., Orem, UT (801) 223-4600.
- Rosenberg, Jerry L. 1986. Is There Capacity Planning After the Model? In Proceedings. Computer Measurement Group: CMG, Inc.
- Rosenberg, Jerry L. 1988. Great Myths of Capacity Planning. In Proceedings. Computer Measurement Group: CMG, Inc.
- Rosenberg, Jerry L. and Ellen M. Friedman. 1984. Capacity Planning in a Decentralized Environment. In Proceedings. Computer Measurement Group,: 422-424: CMG, Inc.
- Salsburg, Michael. 1994. Capacity Planning in the Interoperable Enterprise. In Proceedings. Computer Measurement Group: CMG, Inc.
- SAS. QNAP II and Modline . Simulog AS, France.
- Schaming, Bernie. 1996. HW/SW Co-Design Results to Date. postscript.
- Schriber, Thomas J. and Daniel T. Brunner. 1996. Inside Simulation Software: How it works and why it matters. In Proceedings. Winter Simulation Conference,: 1-18: <http://www-personal.umich.edu/~schriber/papers/tjs-dtb-96.html>.
- Schwartz, Jeffry A. 1995. Modeling Considerations When Merging Computers. In Proceedings. Computer Measurement Group: CMG, Inc.

- Schwetman, Herb. 1978. Hybrid Simulation Models of Computer Systems. Communications of the ACM 21 (9): 718-723.
- Schwetman, Herb. 1996. Private email. July/August 1996
- Senson, Geoff. 1985. How to Assess the Results of Your CPU Benchmark. In Proceedings. Computer Measurement Group,: 658-667: CMG, Inc.
- SES. SES/workbench and SES/strategizer . Scientific and Engineering Software, Inc, 4301 West Bank Dr. Blgg A, Austin, TX 78746.
- Smith, Connie U. 1990. Performance Engineering of Software Systems. The SEI Series in Software Engineering. Reading, MA: Addison-Wesley Publishing Co.
- Smith, Connie U. 1995. The Evolution of Performance Analysis Tools. Informationstechnik und Technische Informatik 37 : 17-20.
- Smith, Connie U. and Lloyd G. Williams. 1995. A Performance Model Interchange Format. Santa Fe, NM: Performance Engineering Services, PO Box 2640, 87504-2640. <http://www.rt66.com/~cusmith/cspubs.htm#Conf/pmifpaper.pdf>.
- Sobh, Tarek M. and Beno Benhabib. 1996. Describe Event Dynamic System Technical Committee. HTML: <http://sgi01.acim.usl.edu/RAS/RASdedsTC.html>.
- Taylor, David J. 1994. Hybrid modeling: The next step. Computer-aided Engineering 13 (7): 32-35.
- Thomasian, Alexander and Kameshwar Gargeya. 1984. Speeding Up Computer System Simulations Using Hierarchical Modeling. In Proceedings. Computer Measurement Group,: 845-850: CMG, Inc.
- Visual. Simul8 . Visual Thinking International Limited, 141 St James Rd., Glasgow, UK G4 0LT.
- Wallace, Rodney B. 1995. A Performance Monitoring and Capacity Planning Methodology for Transaction Processing Systems in a Distributed Unix Environment. In Proceedings. Computer Measurement Group: CMG, Inc.
- Whitt, Ward. 1993. Large fluctuations in a deterministic multiclass network of queues. Management Science 39 (8): 1020-1028.
- Wicks, Ray. 1989. Balanced Systems. In Transactions. Computer Measurement Group: CMG, Inc.
- Wilson, Gregory L. 1994. Capacity planning in a high-growth organization. In Proceedings. Computer Measurement Group. Orlando, FL: CMG, Inc.

Zibitsker, Boris E. 1994. Massively Parallel Processing: Performance Management Challenge. In Proceedings. Computer Measurement Group: CMG, Inc.

Zimmer, Harry. 1987. Rules of Thumb. In Proceedings. Computer Measurement Group,: 981-987: CMG, Inc.