Approximate Analytical Models for Networked Servers Subject to MMPP Arrival Process

Bruno Ciciani, Andrea Santoro, Paolo Romano

Computer Engineering Department, University of Rome "La Sapienza"

Main research project: SLA and penalty minimization

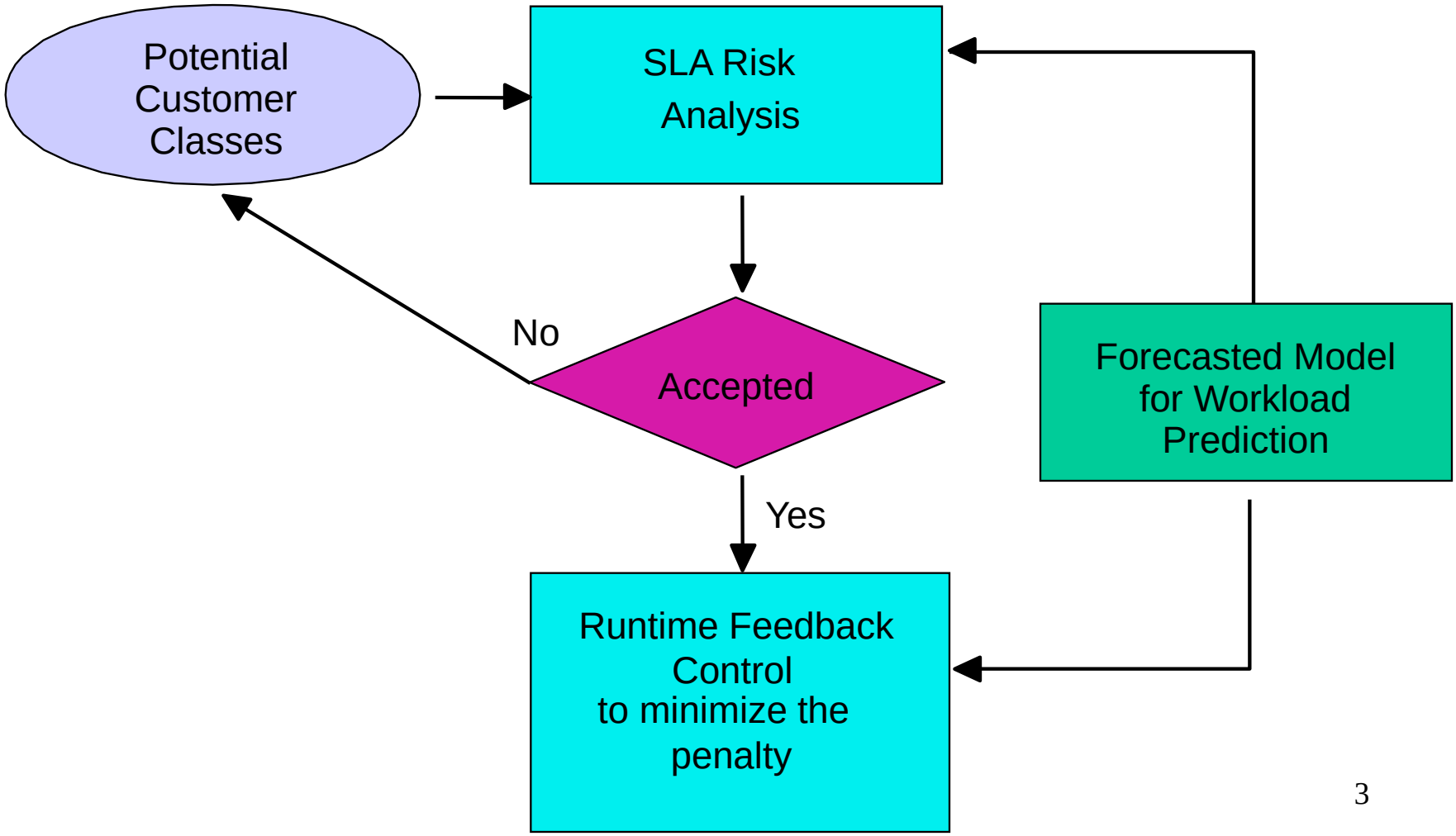
- Service provider economical risk analysis in planning phase

- Run-time minimization penalty control

Reference platform:

- WWW content hosting
- Grid platforms

Process Flow



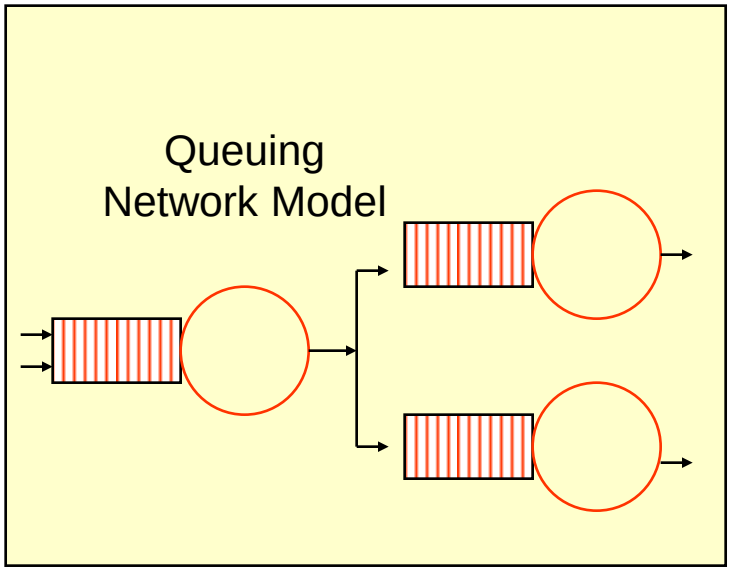
SLA Risk analysis (4 phases)

1. Definition of the parameters involved in the SLA.
2. Worload characterization and service time identification.
3. Platform and resource allocation policy modeling and evaluation.
4. Economical risk identification.

Platform and resource allocation policy modeling and evaluation

System Description

- System parameters
- Resources parameters
- Workload parameters
  - service demands
  - workload intensity



Performance Measures

- Response time
- Throughput
- Utilization
- Queue length

Characteristics of the incoming traffic for GRID and WWW content delivery platforms

- Heavy-tailed distributions in workload characteristics, that means a very large variability in the values of the workload parameters.
- Burstiness behavior – the arrivals are coming with different intensity during the time, in some of these they arrive in a burst way.
- Self-similarity - a self-similar process looks bursty across several time scales, i.e. incoming traffic looks the same when measured over scales ranging from millisecond to minutes and hours.

Markov Modulated Poisson Process captures last two characteristics

Power Laws: $y \propto x^{\alpha}$

- Heavy-tailed distribution

$$P[X > x] = kx^{-\alpha}L(x)$$

  - Great degree of variability, and a non negligible probability of high sample values
  - When α is less then 2, the variance is infinite, when α is less than 1, the mean is infinite.
  - Zipf's Law describes phenomena where large events are rare, but small ones are quite common
  - Popularity of static pages

Accounting for Heavy Tails: an example (1)

- The HTTP LOG of a Web server was analyzed during 1 hour. A total of 21,600 requests were successfully processed during the interval.
- Let us use a multiclass model to represent the server.
- There are 5 classes in the model, each corresponding to the 5 file size ranges.

Accounting for Heavy Tails: an example (2)

- File Size Distributions.

Class	File Size Range (KB)	Percent of Requests
1	Size < 5	25
2	5 ≤ size ≤ 50	40
3	50 ≤ size ≤ 100	20
4	100 ≤ size ≤ 500	10
5	size ≥ 500	5

Accounting for Heavy Tails: an example (3)

- The arrival rate for each class r is a fraction of the overall arrival rate $\lambda = 21{,}600/3{,}600 = 6$ requests/sec.
  - $\lambda_1 = 6 \times 0.25 = 1.5$ req./sec
  - $\lambda_2 = 6 \times 0.40 = 2.4$ req./sec
  - $\lambda_3 = 6 \times 0.20 = 1.2$ req./sec
  - $\lambda_4 = 6 \times 0.10 = 0.6$ req./sec
  - $\lambda_5 = 6 \times 0.05 = 0.3$ req./sec

Markov Modulated Poisson Process

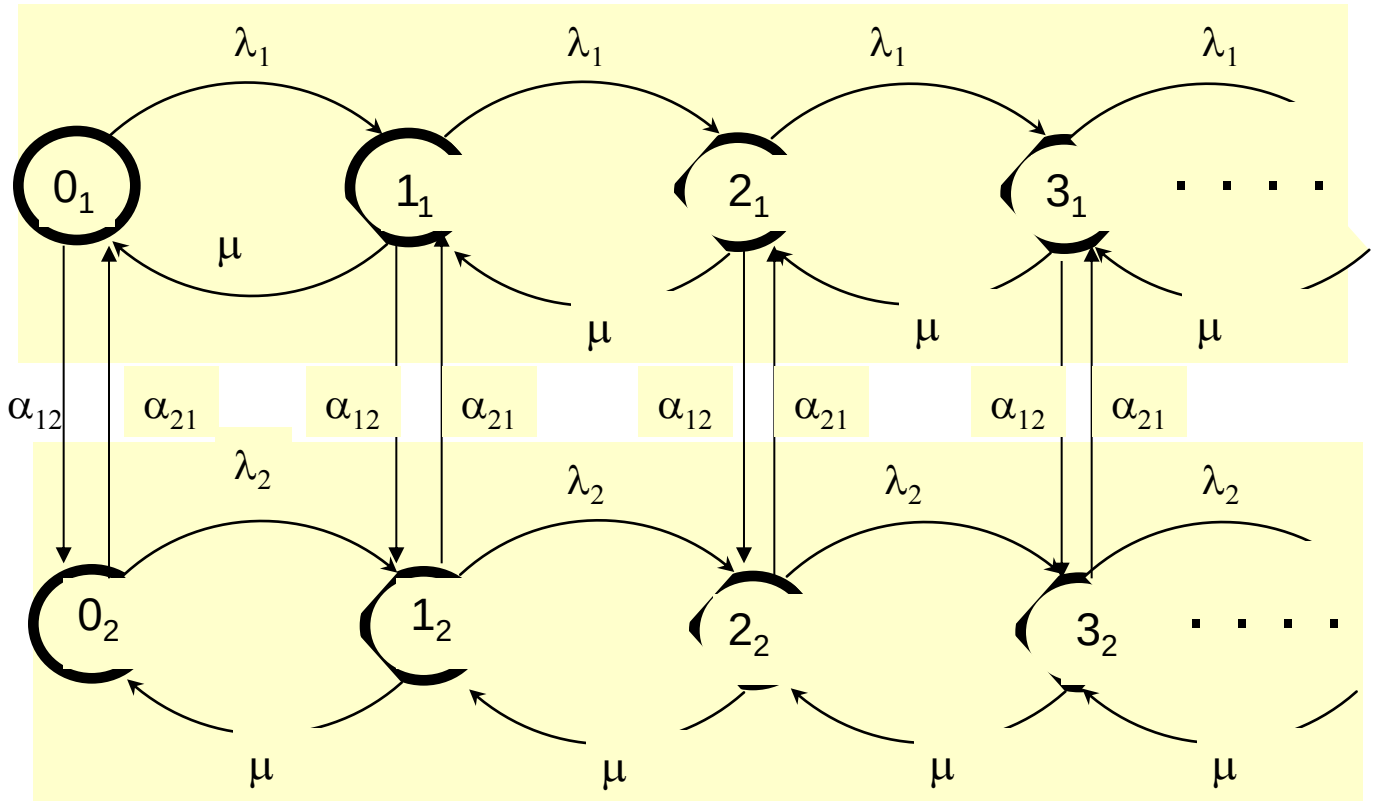
Goal of the contribution

Give a technique to evaluate a MMPP/M/1 with a computational complexity comparable to M/M/1 in a QoS modeling context

Outline of the presentation

- MMPP/M/1 modeling and its evaluation state of the art
- Main idea of our evaluation technique
- Unbiased approximation
- Lower bound approximation
- Upper bound approximation
- Validation (synthetic benchmarks)
- Real case study (Grid platform analysis)
- Conclusions and future work

MMPP/M/1 states representation



State of the art for MMPP/M/1 evaluation

- Basic exact solution techniques:
  - Matrix geometric methods
  - Generating function methods
  - Spectral expansion methods
- Combination of previous methods

Drawbacks of previous MMPP/M/1 evaluation techniques

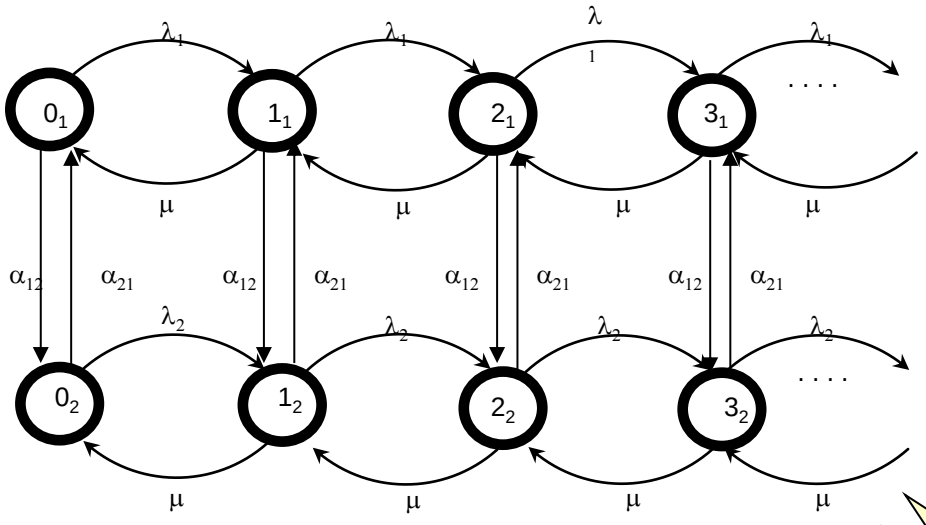
- They require iterative solutions or numerical methods (e.g. for matrix eingenvalues determination) whose computational cost is very high. Hence they are not useful for:
  - Real-time decision making aimed at server platform reconfiguration (e.g., via request redirection towards a different server instance in case of critical events) while still ensuring adequate service levels

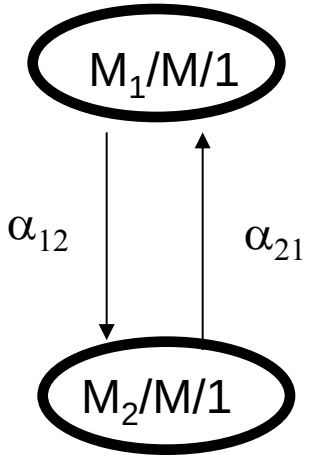
Basic idea (1/2)

- Model an MMPP/M/1 server as a combination of M/M/1 process
- The approximation must be used to evaluate platforms subject to SLA constrains based on percentile, i.e. the response time or the queue length must be less of a threshold T (e.g. 3 sec) for a given probability P (e.g. 95%)
- Denoting with $F_{MMPP/M/1}(t)$ and with $F_{approximation}(t)$ the cumulative distribution function of the response time of the original process and of the approximated model, respectively, we have that

$$F_{approximation}(T) < F_{MMPP/M/1}(T) < P$$

Basic idea (2/2)



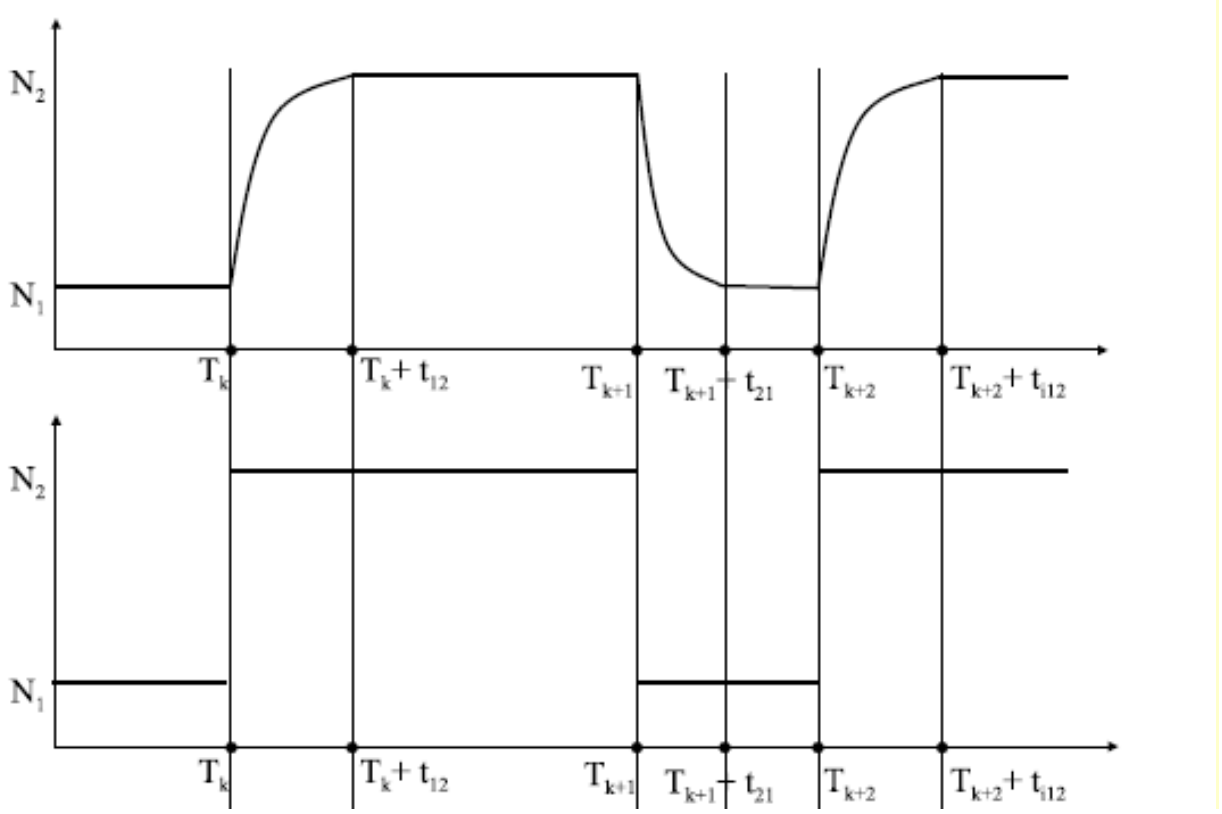


Approximation construction

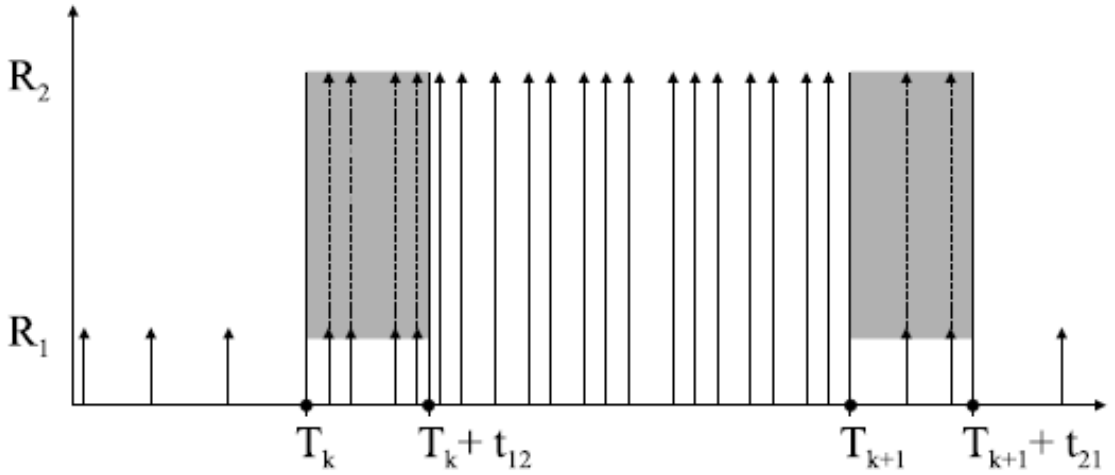
- Identification of the H ($M_1 \ldots M_H$) MMPP states
  (from the workload characterization)
- Identification of the arrival rates
  (from the workload characterization)
- Evaluation of the steady state probabilities for each state S_i of the MMPP
  (using standard results in queuing theory)
- Evaluation of the cumulative distribution function obtained as linear combination of the H steady states $M_i/M/1$
  (the service rate is assumed the same for all the states)

Unbiased approximation

The behavior of the MMPP/M/1 process is approximated adopting, as the weights of the linear combination, the probabilities, p_i, for the MMPP to stay (at steady state) in each state S_i

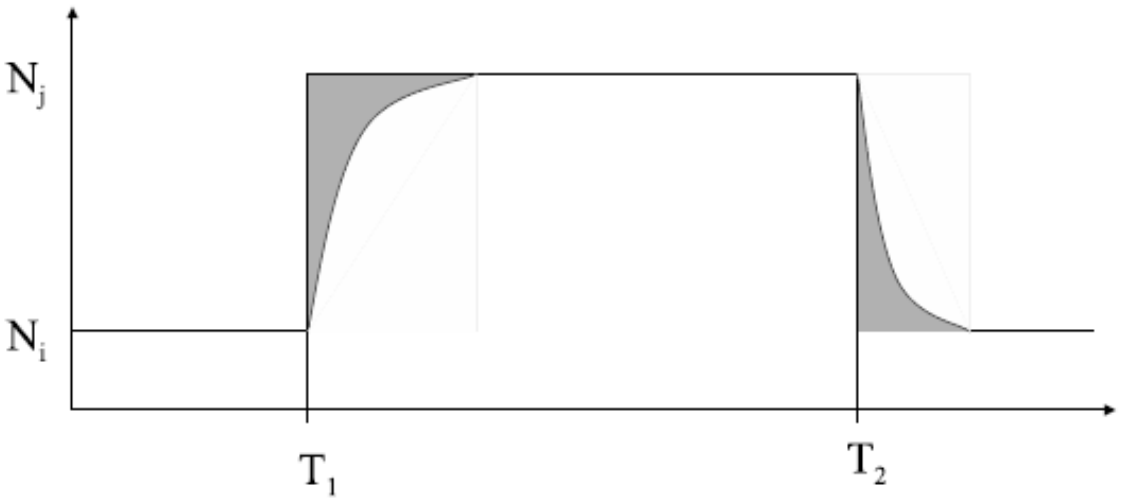


Unbiased approximation (response time)



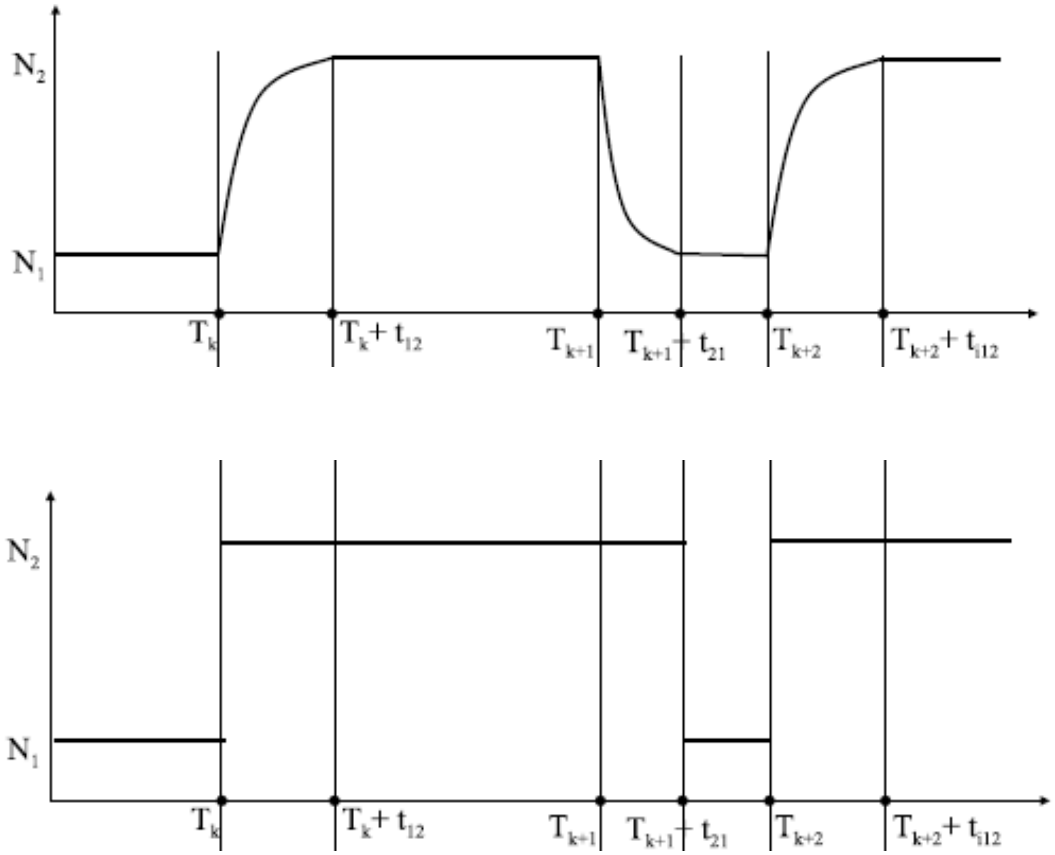
$$R = \sum_{i=1}^{H} w_i R_i \qquad w_i = \frac{p_i \lambda_i}{\sum_{j=1}^{H} p_j \lambda_j}$$

Unbiased approximation
(difference between unbiased and exact MMPP/M/1 behavior)



- The error is given by the difference between the areas comprised between state S_i to S_j (for the real MMPP/M/1) and the immediate transition to S_j (for the analytical approximation) and viceversa
- The two areas tend to cancel each other
- No possibility to guarantee the overstimation of the response time

Lower Bound approximation
(idea: systematic overstimation of the queue lenght during transient periods)



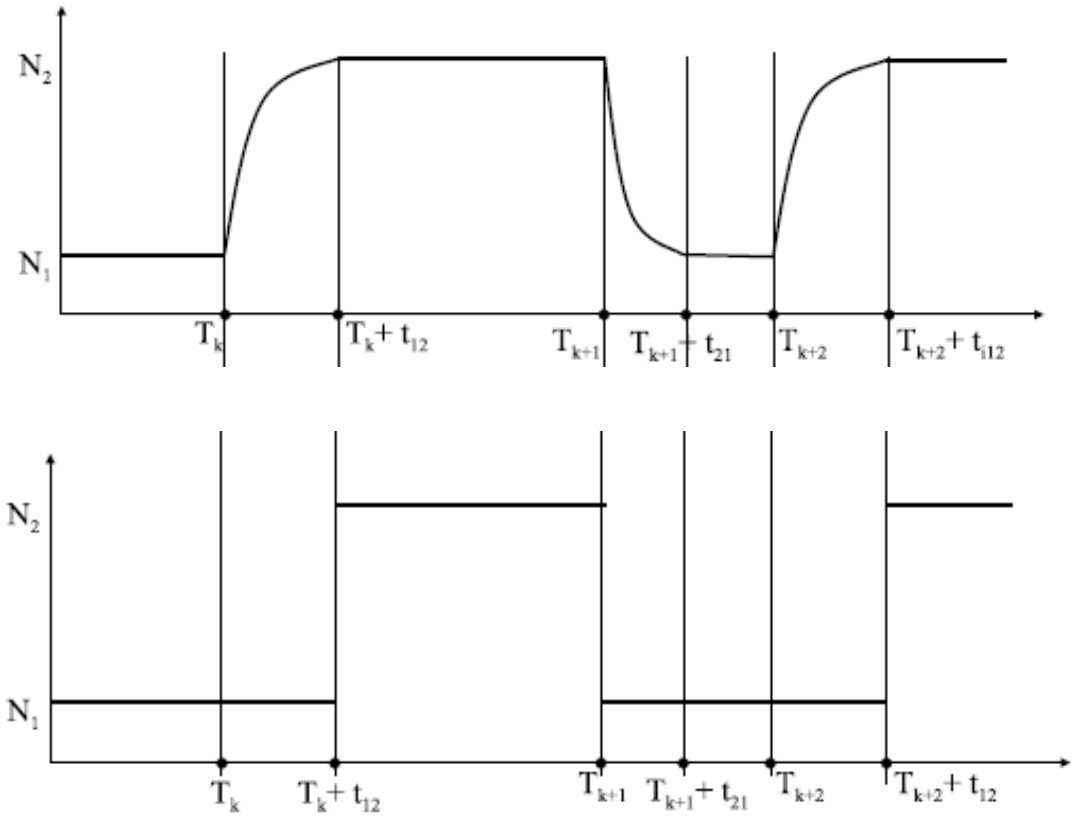
Lower Bound approximation (evaluation procedure)

- Evaluation of the steady state probabilities for each state S_i of the MMPP
  (using standard results in queuing theory)
- Evaluation of the transient phases durations
  (according to classical queuing theory)
- Evaluation of the modified probability

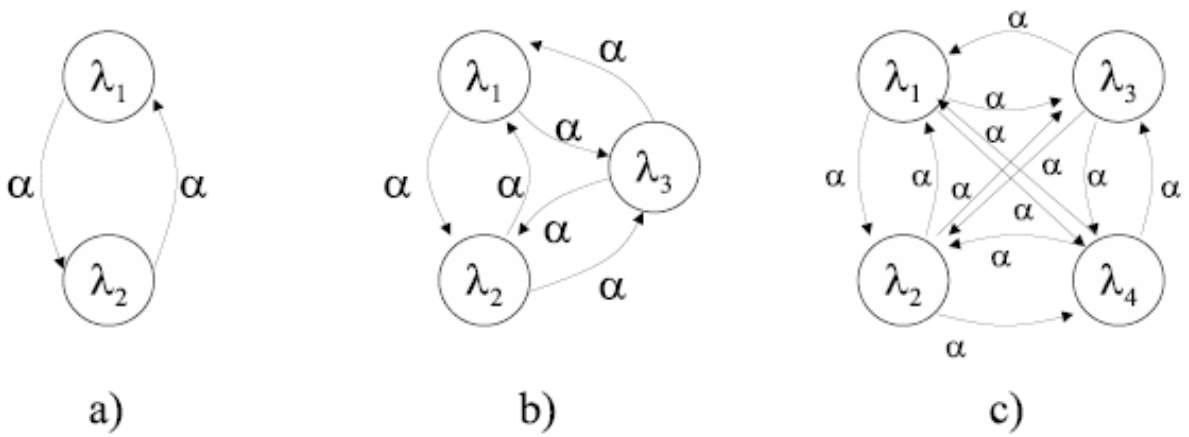
$$p'_i = p_i + \sum_j^{\lambda_i > \lambda_j} p_i \alpha_{ij} t_{ij} - \sum_j^{\lambda_i < \lambda_j} p_j \alpha_{ji} t_{ji}$$

- Generation of the lower bound process by performing a weighted superposition of the output process of the H steady states $M_i/M/1$

Upper Bound approximation
- to indentify the maximum error-
(idea: systematic understimation of the queue lenght during transient periods)



Validation
(synthetic benchmarks)



	2 States	3 States	4 States
α	0.1	0.05	0.0333...
P_i	0.5	0.3333....	0.25

Table 1: Transition rates for the adopted MMPP models.

Validation: case with heavy load

States	λ_1	λ_2	λ_3	λ_4	U_1	U_2	U_3	U_4
2	10	90	-	-	0.1	0.9	-	-
3	10	50	90	-	0.1	0.5	0.9	
4	10	35	65	90	0.1	0.35	0.65	0.9

Table 2: Interarrival rates (and corresponding Utilization Factors) for each MMPP model (Heavy Load).

	Mean Queue Length		Mean Response Time	
States	Lower Bound	Upper bound	Lower Bound	Upper bound
2	16%	7%	33%	14%
3	15%	6%	15%	7%
4	14%	5%	11%	4%

Table 3: Results of the simulator compared with error estimation (2 orders of magnitude, heavy load).

	Mean Queue Length		Mean Response Time	
States	Lower Bound	Upper bound	Lower Bound	Upper bound
2	1.6%	0.7%	3.3%	1.5%
3	1.3%	0.6%	1.5%	0.7%
4	1.3%	0.5%	1.1%	0.4%

Table 4: Results of the simulator compared with error estimation (3 orders of magnitude, heavy load).

Validation: case with light load

States	λ_1	λ_2	λ_3	λ_4	U_1	U_2	U_3	U_4
2	10	50	-	-	0.1	0.5	-	-
3	10	30	50	-	0.1	0.3	0.5	-
4	10	25	35	50	0.1	0.25	0.35	0.5

Table 5: Interarrival rates (and corresponding Utilization Factors) for each MMPP model (Light Load).

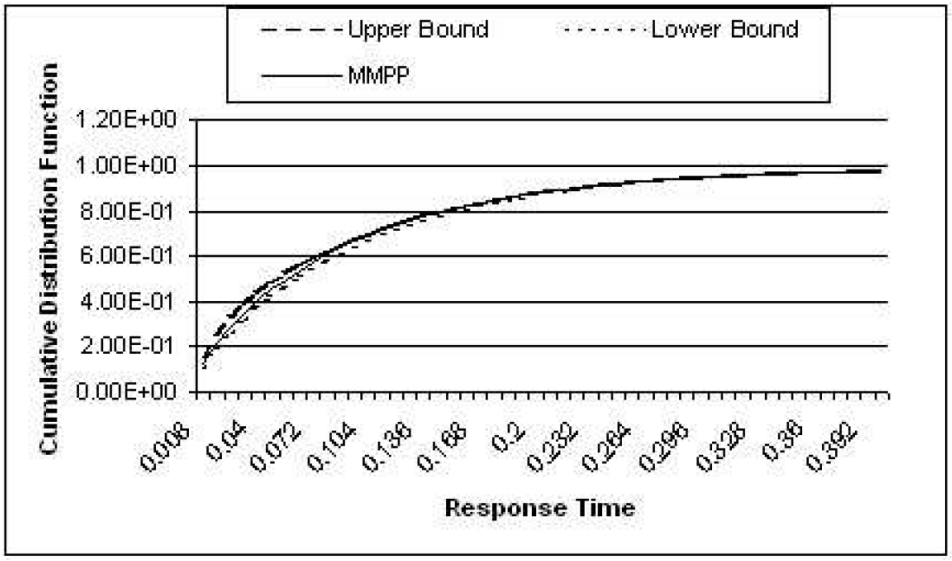
	Mean Queue Length		Mean Response Time	
States	Lower Bound	Upper bound	Lower Bound	Upper bound
2	4.0%	3.2%	4.8%	3.8%
3	2.4%	2.0%	1.6%	1.17%
4	2.3%	1.9%	1.0%	0.9%

Table 6: Results of the simulator compared with error estimation (2 orders of magnitude, light load).

	Mean Queue Length		Mean Response Time	
States	Lower Bound	Upper bound	Lower Bound	Upper bound
2	0.34%	0.30%	0.47%	0.38%
3	0.25%	0.19%	0.16%	0.11%
4	0.23%	0.20%	0.15%	0.07%

Table 7: Results of the simulator compared with error estimation (3 orders of magnitude, light load).

Cumulative distribution function
(case of 2 states)



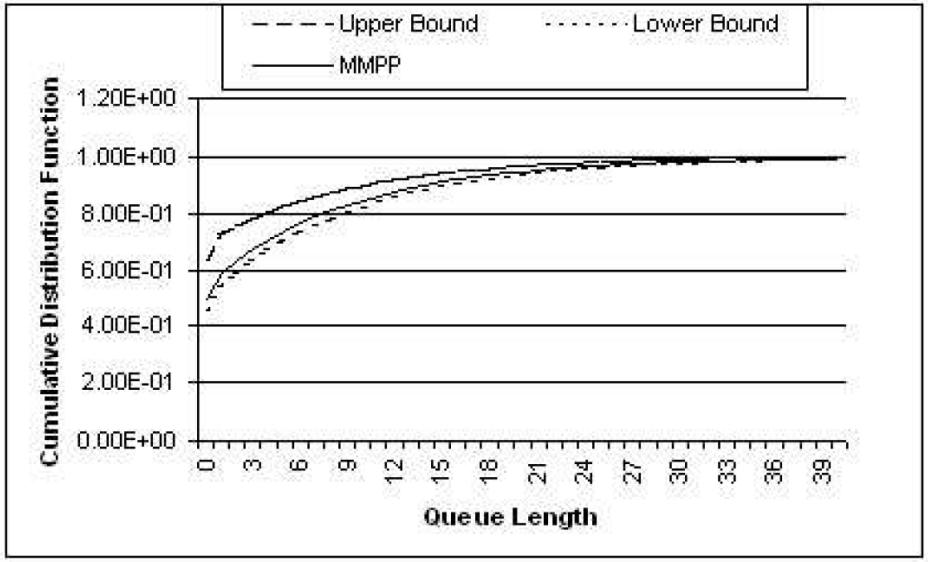
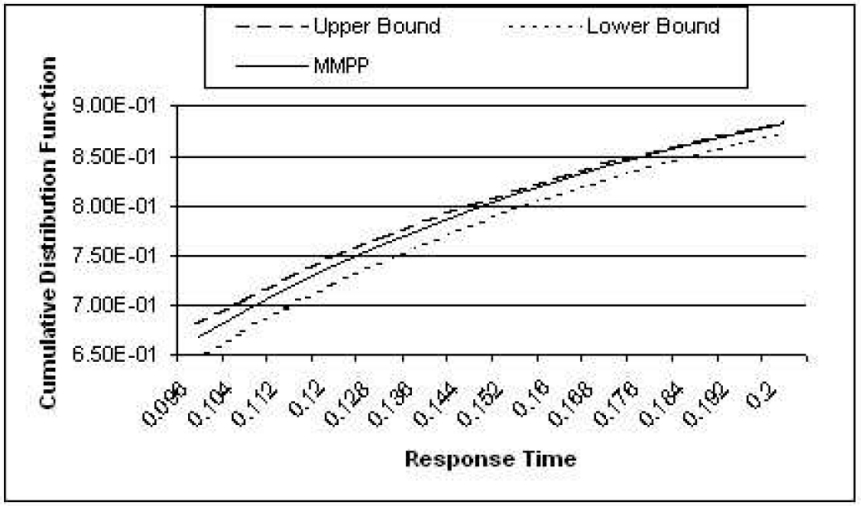


Figure 7: Distribution Functions for the 2-states model.

Cumulative distribution function

(case of 2 states)

zoom



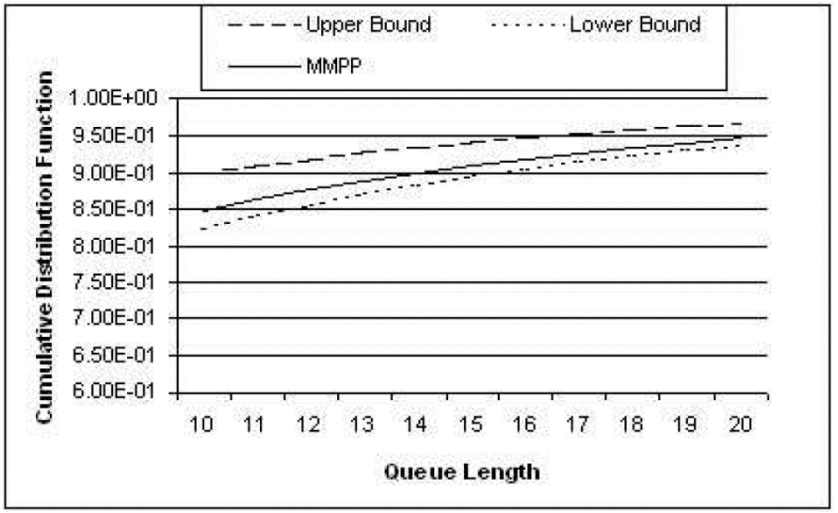
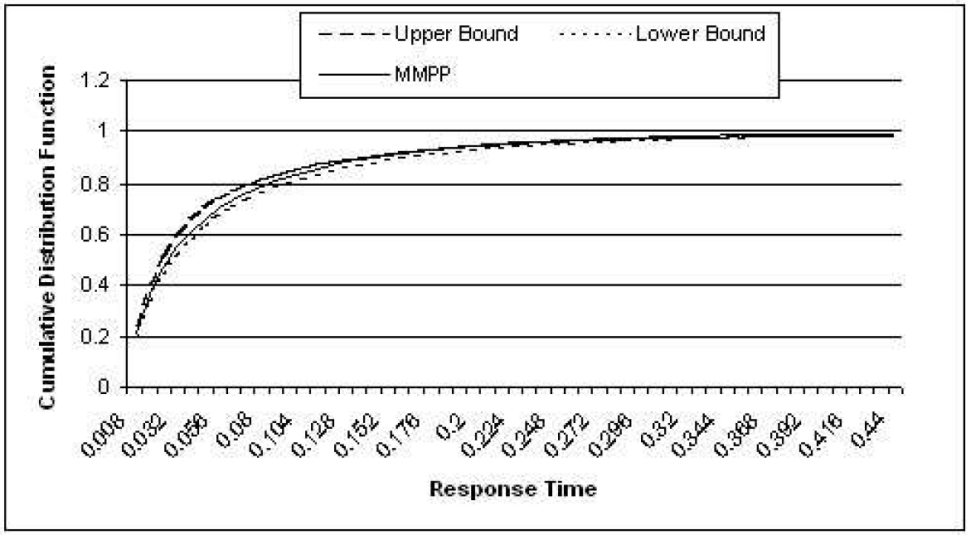


Figure 9: Zoom on Distribution Functions for the 2-states model.

Cumulative distribution function
(case of 4 states)



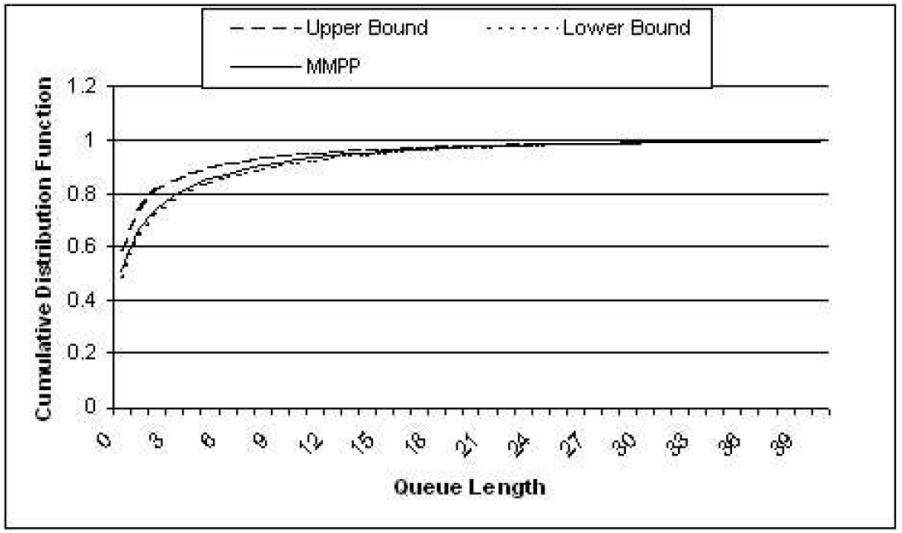
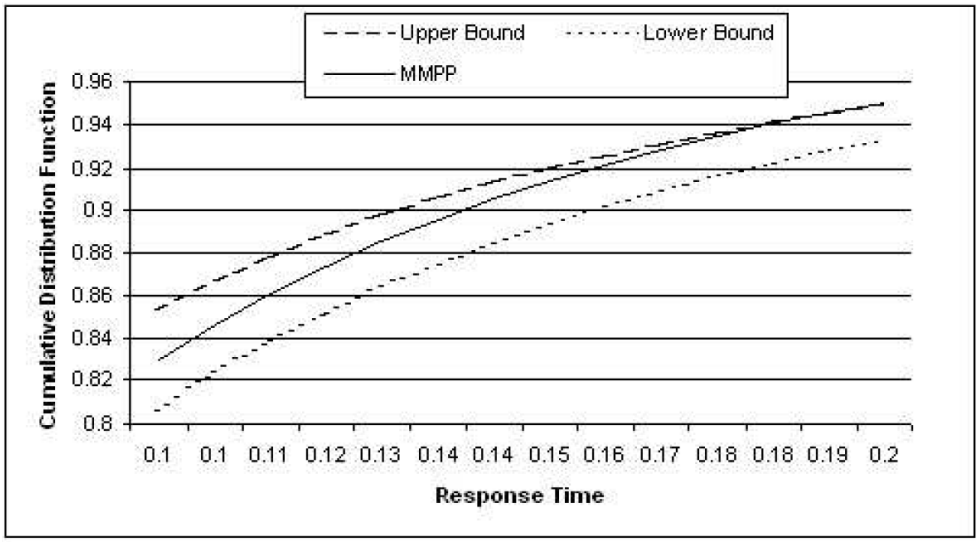


Figure 8: Distribution Functions for the 4-states model.

Cumulative distribution function

(case of 4 states)

zoom



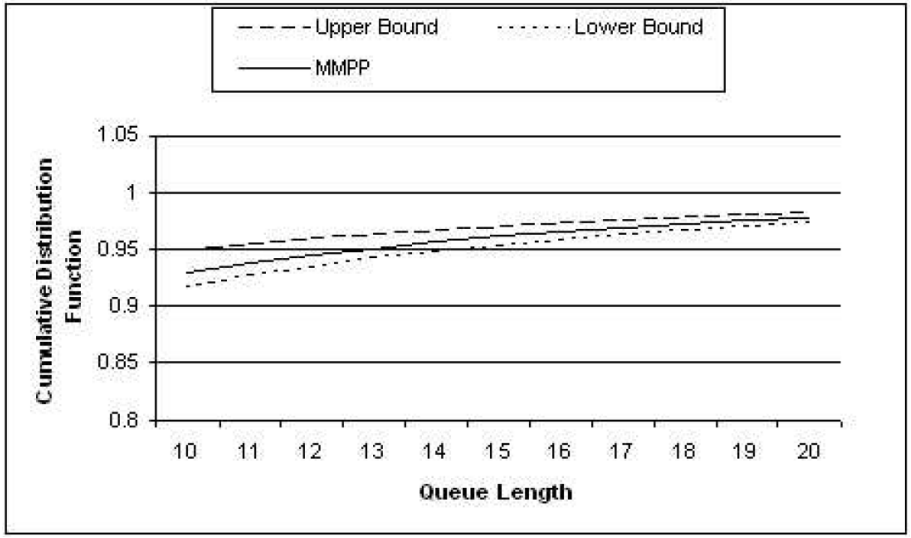


Figure 10: Zoom on Distribution Functions for the 4-states model.

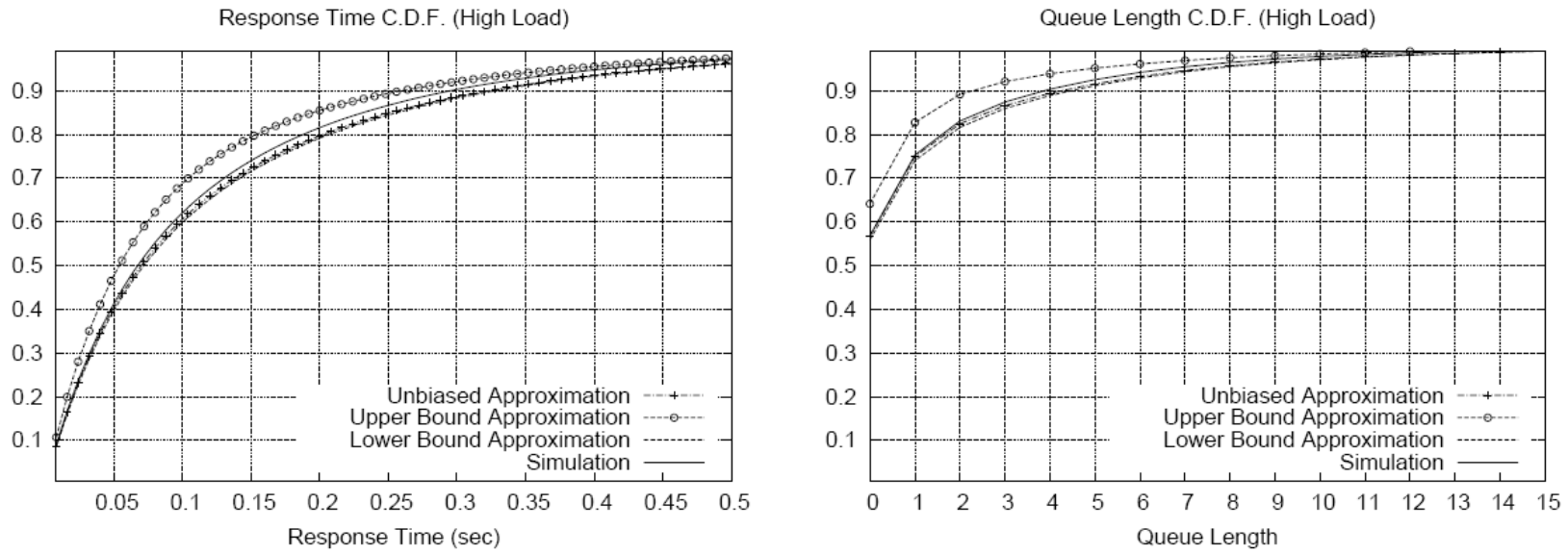
Observations

- The behavior of the MMPP/M/1 is overestimated and underestimated correctly by the upper and lower bound approximation model.
- The error is proportional to the utilization factor value gap (i.e. in the 2 state model the MMPP oscillates between two extremes (utilization factor of 0.1 and 0.9), while the other two models perform softer transitions.
- The error decrease with the increment of the magnitude order between the arrival (or service) rate and the transition rates of the MMPP states (with 2 order is acceptable, with 3 is negligible).
- The unbiased approximation is a good indicator of the real MMPP/M/1 behavior

Real case study
- Grid server -

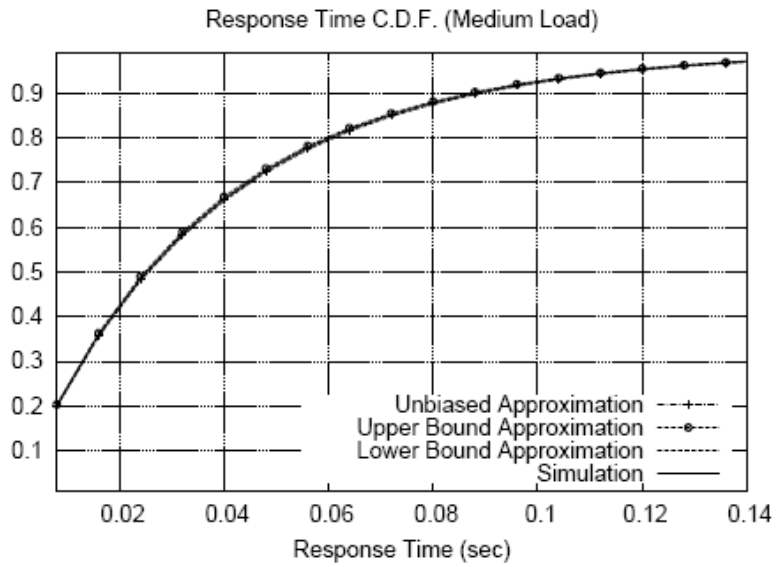
- MMPP/M/1 whose parameters come from real traces
- Incoming traffic requests modeled by a 2-state MMPP model
  - $\alpha_{12} = 0.17$, $\alpha_{21} = 0.08$
  - $\lambda_1 = 22.1$, $\lambda_2 = 7.1$

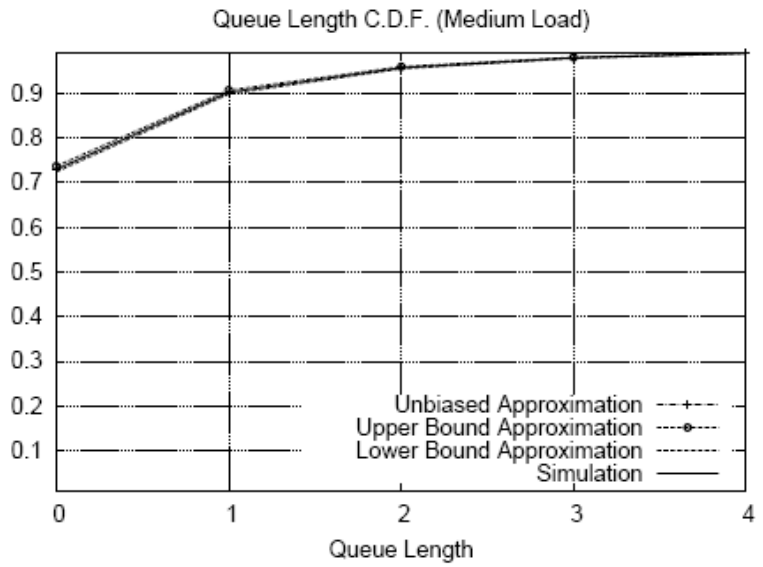
Heavy load (μ = 25, ρ = 0.884)



Cumulative Distribution Functions for Response Time and Queue Length (Heavy Load)

Medium load (μ = 33, ρ = 0.67)





Cumulative Distribution Functions for Response Time and Queue Length (Medium Load)

Conclusions and future work

- Deep analyze of the factors affecting the approximation error
- In some case service time presents heavy-tailed distributions, using Feldmann and Whitt's algorithm it is possible approximate a heavy-tailed distribution with a hyper-exponential distribution, so we will analyze the MMPP/H/1
- Analyze the performance behavior of load balancing policies for tasks with heavy tailed distributions