



# HETEROGENEOUS FINITE-SOURCE RETRIAL QUEUEING SYSTEMS

B. Almási\*, G. Bolch, J. Sztrik\*

University of Erlangen, Germany

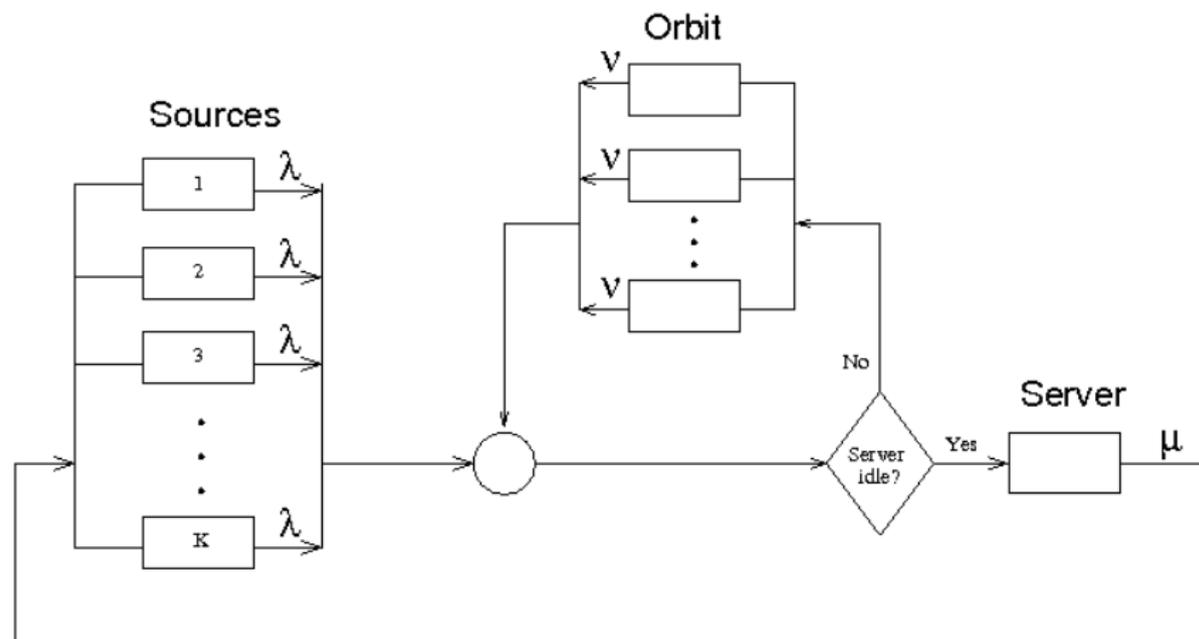
\* University of Debrecen, Faculty of Informatics, Hungary

<http://irh.inf.unideb.hu/user/jsztrik>

# Outline

- 1 The queueing model, applications
- 2 Case studies
- 3 Bibliography

# The queueing model



# Applications

- magnetic disk memory systems
- local area networks with CSMA/CD protocols
- collision avoidance local area network modeling

## Mathematical model

$$P(0; 0) = \lim_{t \rightarrow \infty} P(C(t) = 0, N(t) = 0)$$

$$P(j; 0) = \lim_{t \rightarrow \infty} P(\alpha_1 = j, N(t) = 0), \quad j = 1, \dots, K$$

$$P(0; i_1, \dots, i_k) = \lim_{t \rightarrow \infty} P(C(t) = 0, \beta_1 = i_1, \dots, \beta_k = i_k), \quad k = 1, \dots, K-1$$

$$P(j; i_1, \dots, i_k) = \lim_{t \rightarrow \infty} P(\alpha_1 = j, \beta_1 = i_1, \dots, \beta_k = i_k), \quad k = 1, \dots, K-1.$$

# Performance measures

Once we have obtained these limiting probabilities the **main system performance measures** can be derived in the following way.

## 1. The server utilization with respect to source $j$

$$U_j = P(\text{ the server is busy with source } j)$$

that is, we have to summarize all the probabilities where the first component is  $j$ . Formally

$$U_j = \sum_{k=0}^{K-1} \sum_{i_1, \dots, i_k \neq j} P(j; i_1, \dots, i_k).$$

Hence the **server utilization**

$$U = E[C(t) = 1] = \sum_{j=1}^K U_j.$$

Let us denote by  $P_W^{(i)}$  the steady state probability that request  $i$  is waiting ( staying in the orbit ). It is easy to see that

$$P_W^{(i)} = \sum_{j=0, j \neq i}^K \sum_{k=1}^{K-1} \sum_{i \in (i_1, \dots, i_k)} P(j; i_1, \dots, i_k).$$

Similarly, it can easily be seen, that the steady state probability  $P^{(i)}$  that request  $i$  is in the service facility (it is under service or waiting in the orbit) is given by

$$P^{(i)} = P_W^{(i)} + U_i.$$



and

$$P_W^{(i)} = P^{(i)} - U_i = 1 - \frac{\mu_i + \lambda_i}{\lambda_i} U_i.$$

Alternatively, by the help of (2) we can express the mean response time  $E[T_i]$  for request  $i$  in terms of  $U_i$  as

$$E[T_i] = \frac{P^{(i)}}{\lambda_i(1 - P^{(i)})} = \frac{1 - \frac{\mu_i}{\lambda_i} U_i}{\mu_i U_i} = \frac{\lambda_i - \mu_i U_i}{\lambda_i \mu_i U_i}. \quad (3)$$

### 3. Mean waiting time of source $i$

The mean waiting time of request  $i$  is given by

$$E[W_i] = E[T_i] - \frac{1}{\mu_i} = \frac{1}{\gamma_i} - \frac{1}{\lambda_i} - \frac{1}{\mu_i} = \frac{\lambda_i - (\mu_i + \lambda_i) U_i}{\lambda_i \mu_i U_i}. \quad (4)$$



## 6. Mean rate of generation of primary calls

$$\bar{\lambda} = \sum_{i=1}^K \gamma_i = \sum_{i=1}^K \lambda_i (1 - P^{(i)}) = \sum_{i=1}^K \mu_i U_i.$$

## 7. Blocking probability of primary call $i$

$$B_i = \frac{\lambda_i \sum_{j=1, j \neq i}^K \sum_{k=0}^{K-1} \sum_{i \neq i_1, \dots, i_k} P(j; i_1, \dots, i_k)}{\bar{\lambda}}.$$

Hence **blocking probability of primary calls**

$$B = \sum_{i=1}^K B_i.$$

In particular, in the case of **homogeneous calls**

$$U_i = E[C(t)]/K, \quad i = 1, \dots, K, \quad N = K - \frac{(\lambda + \mu)U}{\lambda},$$

$$\bar{\lambda} = \lambda E[K - C(t) - N(t)] = \mu U,$$

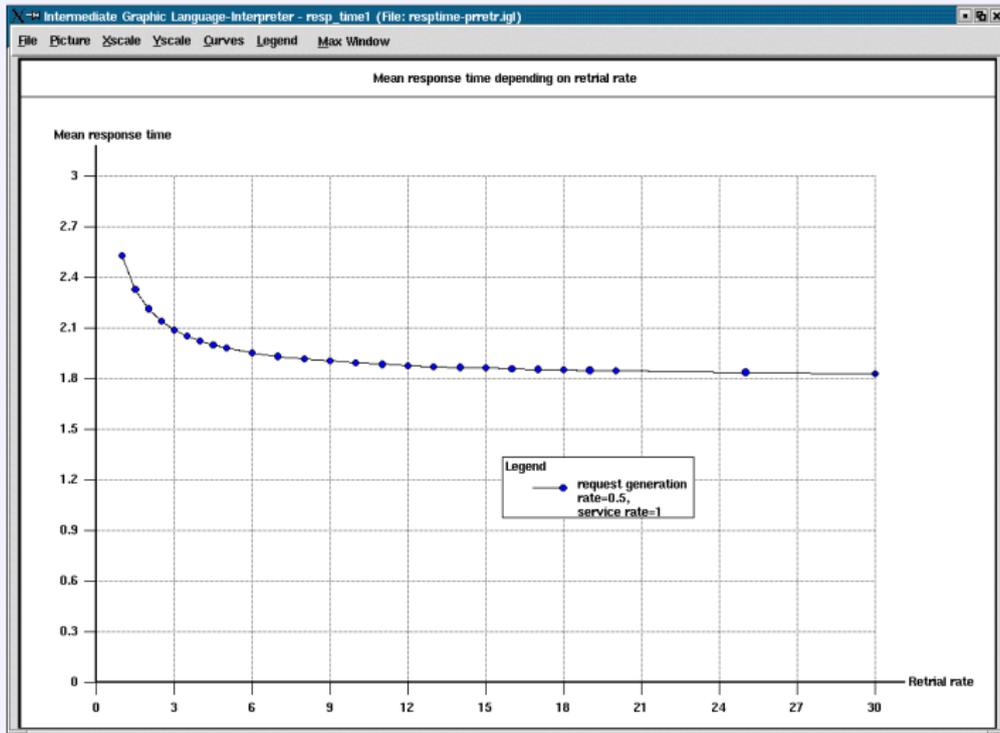
$$E[W] = \frac{N}{\bar{\lambda}} = K - \frac{1}{\mu U} - \frac{1}{\lambda} - \frac{1}{\mu},$$

$$B = \frac{\lambda E[K - C(t) - N(t); C(t) = 1]}{\lambda E[K - C(t) - N(t)]}.$$

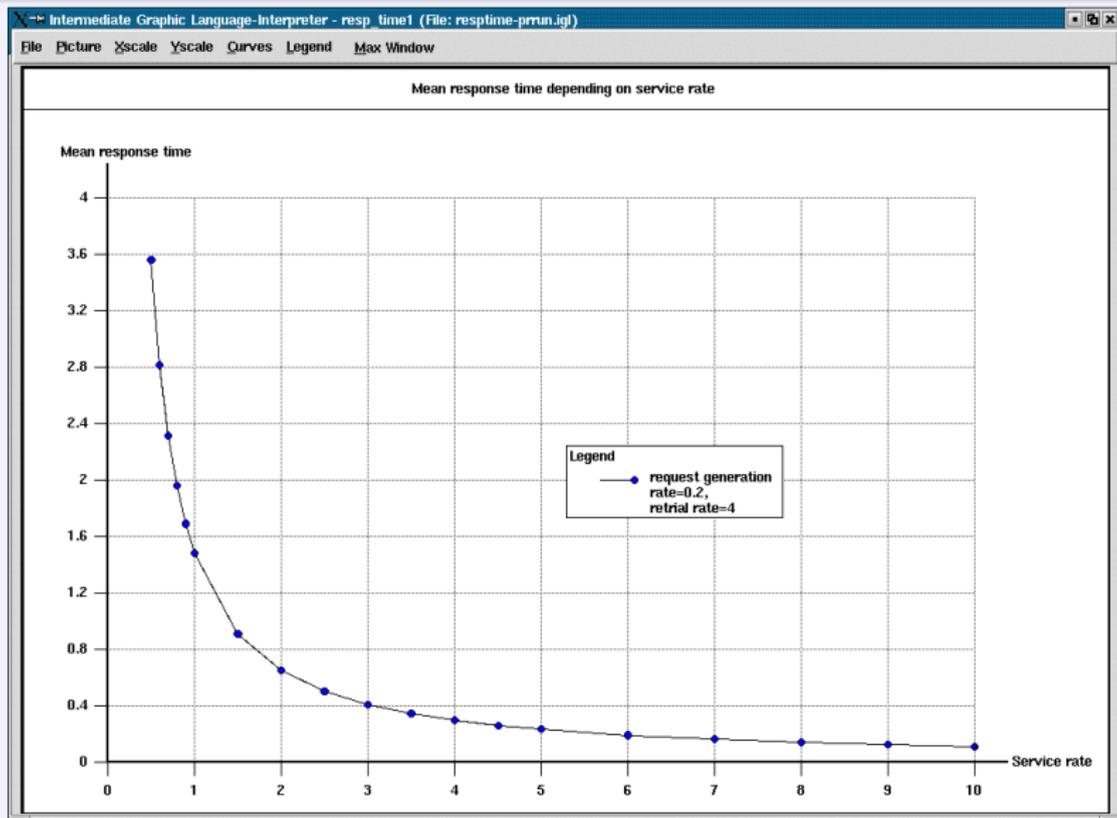
# Evaluation Tool MOSEL

**MOSEL ( Modeling, Specification and Evaluation Language )**  
developed at the University of Erlangen, Germany, is used to  
formulate and solved the problem.

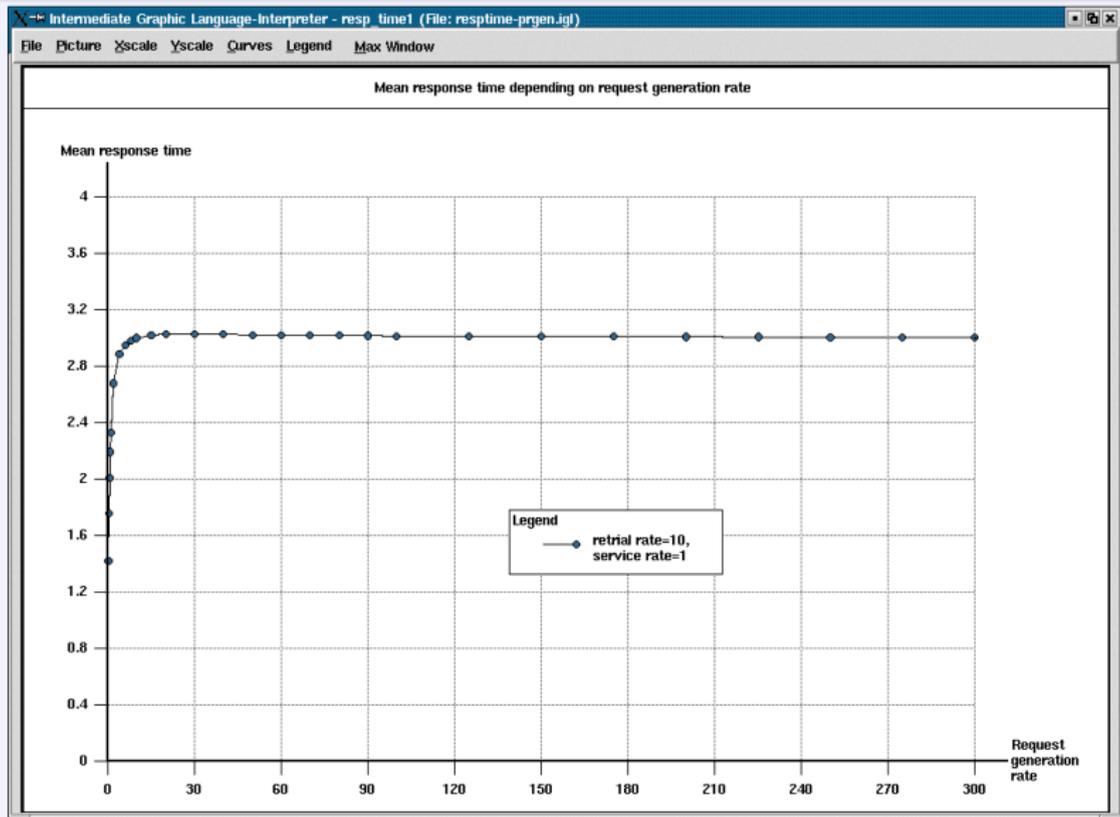
# Case studies



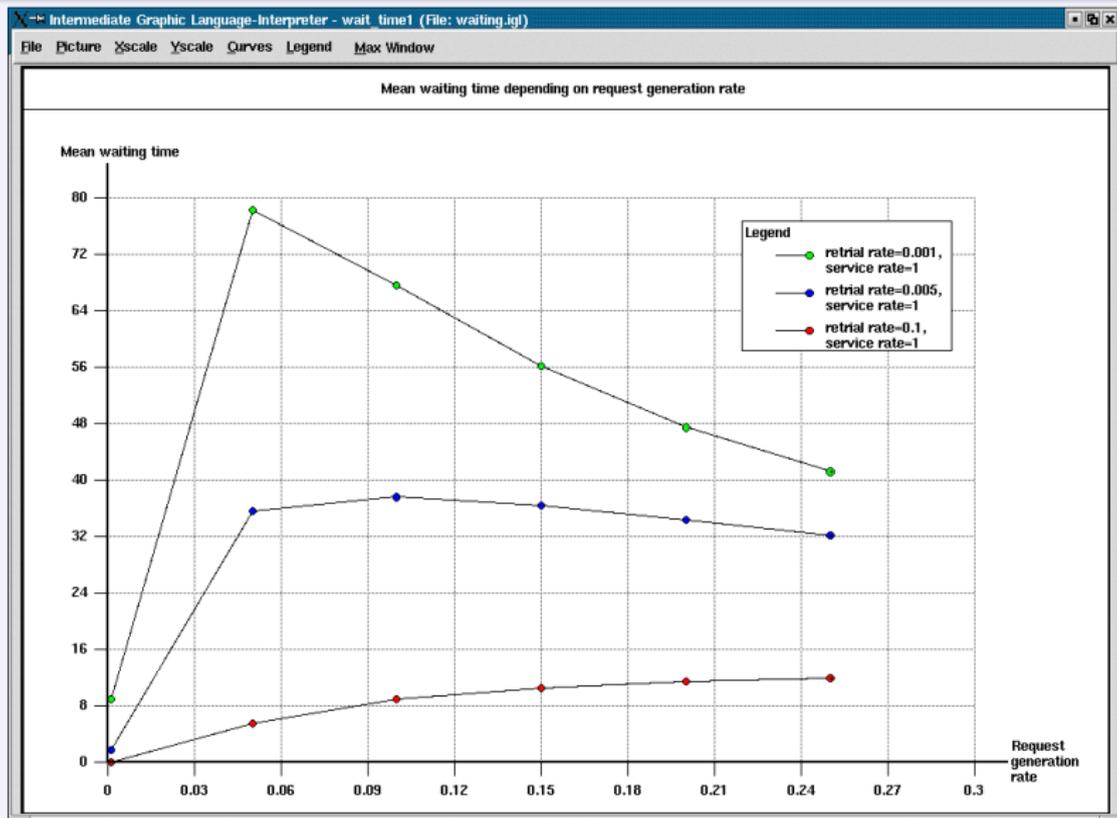
$E[T]$  versus retrial rate



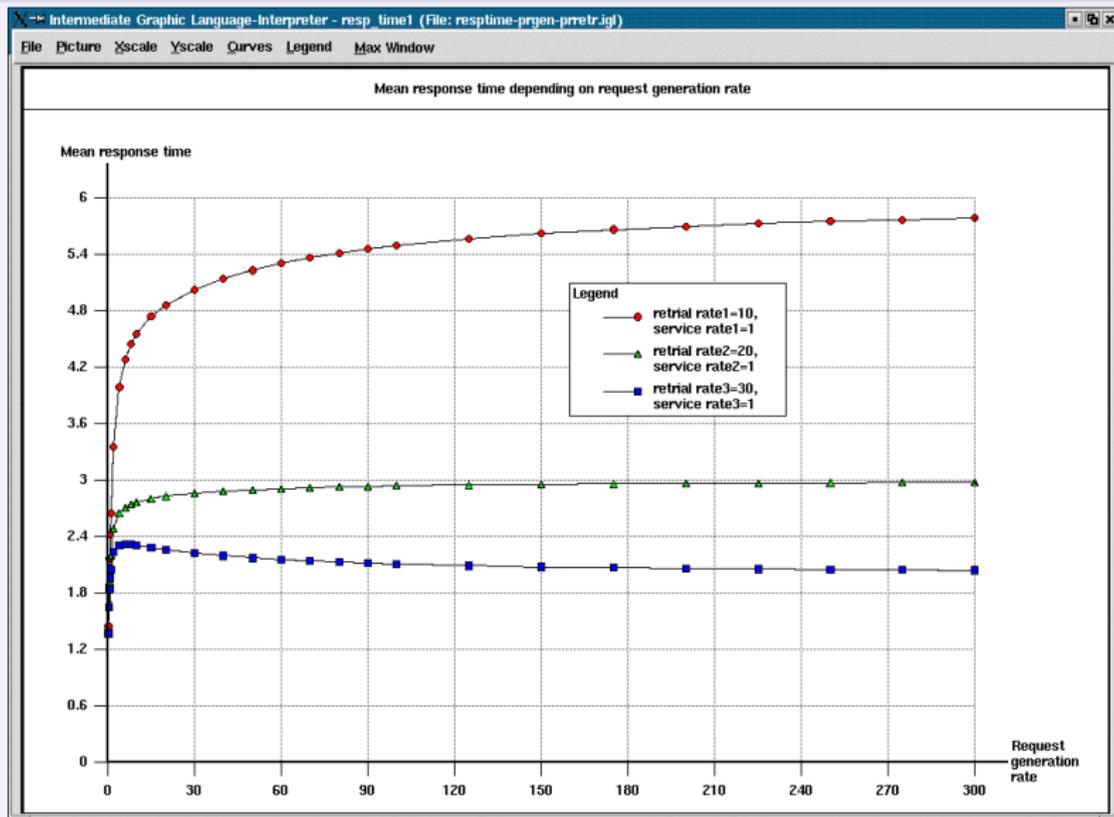
$E[T]$  versus service rate



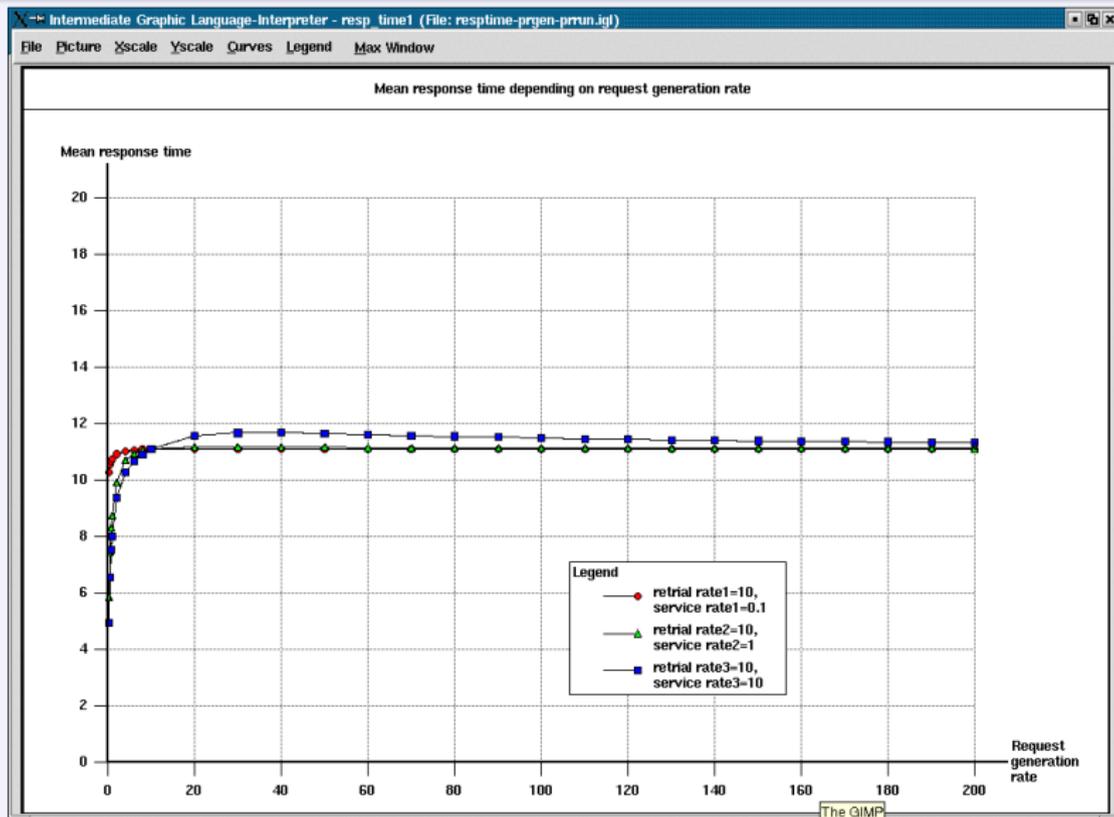
$E[T]$  versus primary request generation rate



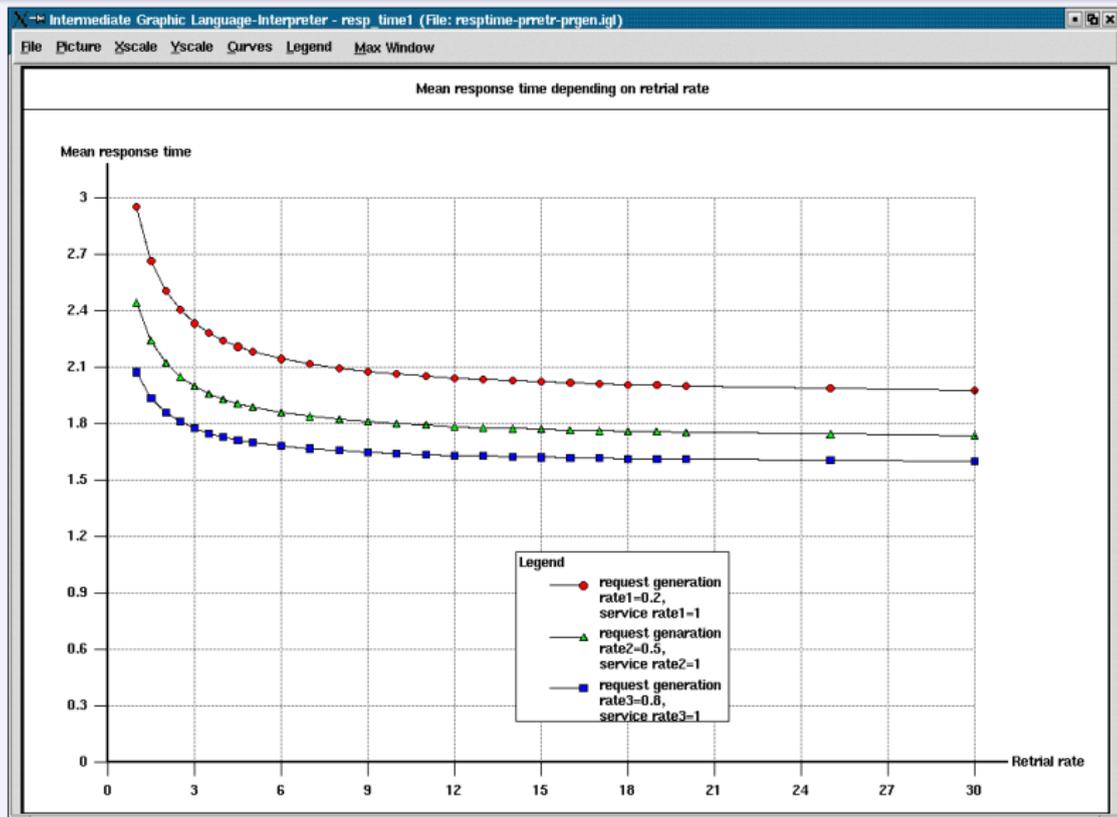
$E[T]$  versus primary request generation rate



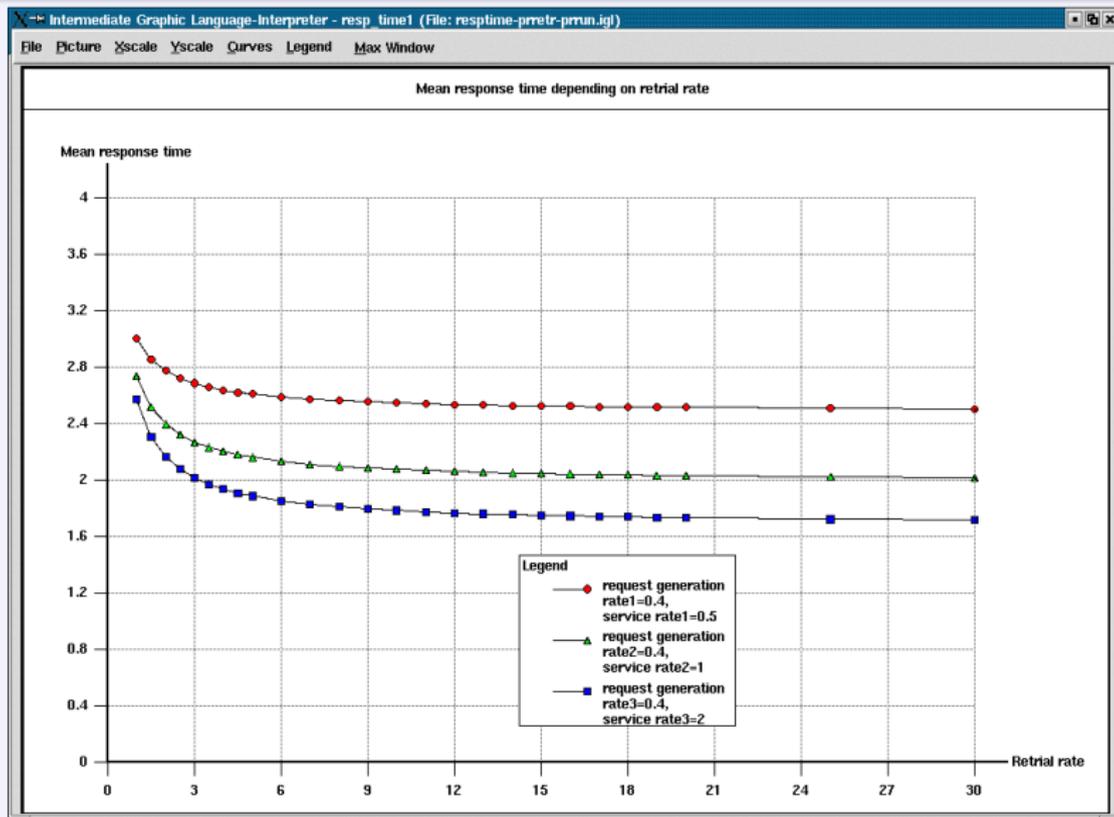
$E[T]$  versus primary request generation rate with homogeneous service and heterogeneous retrial



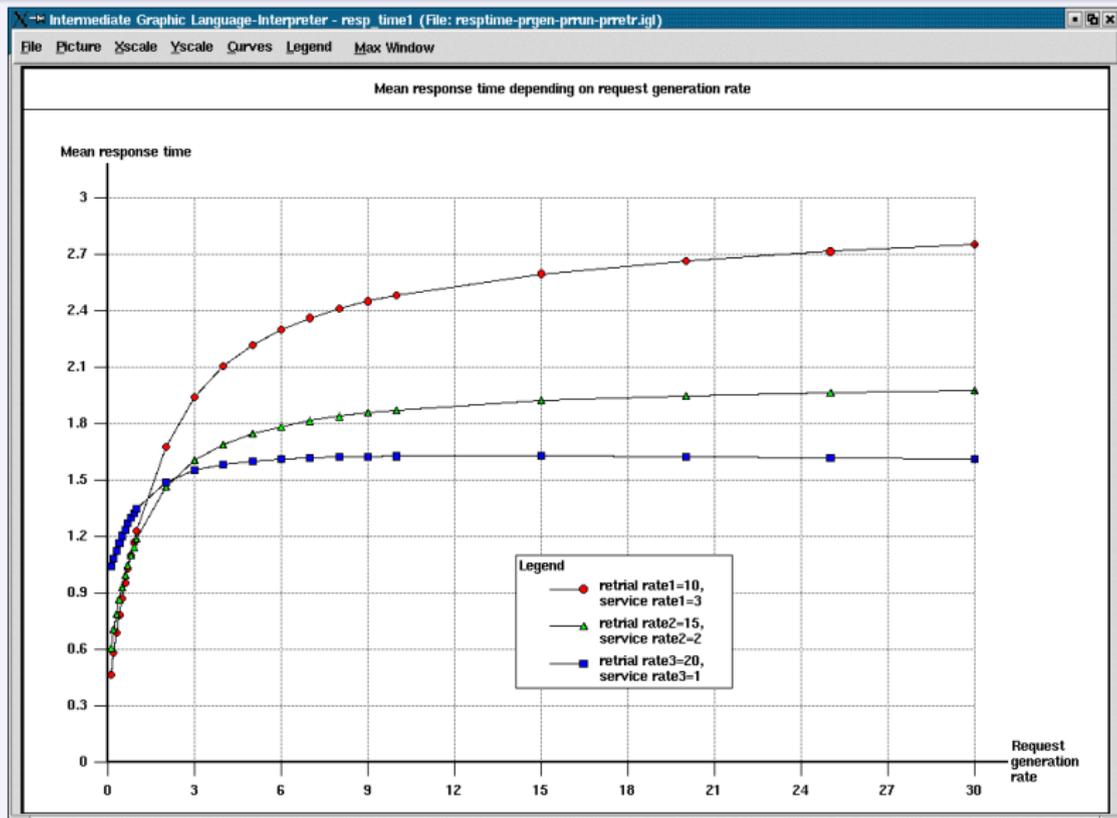
$E[T]$  versus primary request generation rate with homogeneous retrial and heterogeneous service



$E[T]$  versus retrieval rate with homogeneous service and heterogeneous primary request generation



$E[T]$  versus retrieval rate with homogeneous primary request generation and heterogeneous service



$E[T]$  versus primary request generation rate with heterogeneous service and heterogeneous retrial

# Conclusions

- Finite-source retrial queueing system with heterogeneous requests
- Markovian model via MOSEL
- Effect of parameters on performance measures in steady-state
- Graphical presentations

# Bibliography

-  ARTALEJO J.R.: Retrial queues with a finite number of sources, *J. Korean Math. Soc.* 35(1998) 503-525.
-  BEGAIN K., BOLCH G., HEROLD H.: *Practical Performance Modeling, Application of the MOSEL Language*, Kluwer Academic Publisher, Boston, 2001.
-  FALIN G.I. AND TEMPLETON J.G.C: *Retrial queues*, Chapman and Hall, London, 1997.
-  FALIN G.I. AND ARTALEJO J.R.: A finite source retrial queue, *European Journal of Operational Research* 108(1998) 409-424.
-  ALMÁSI B., BOLCH G., SZTRIK J.: Heterogeneous finite-source retrial queues, *Journal of Mathematical Sciences* 121 (2004) 2590-2596.

*Thank You  
for Your  
Attention*