The Foreground-Background Queue
Nuijens, M.F.M.

Citation for published version (APA):
Nuijens, M. F. M. (2004). The Foreground-Background Queue

General rights
It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations
If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: http://uba.uva.nl/en/contact, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.
A model for a queue with one server consists of three ingredients: the times between two arrivals to the queue, the amount of service that customers require, and the manner in which the customers are treated: the \textit{(service) discipline}. The first two are described by probability distributions, which are assumed to be fixed. By means of the third ingredient, the discipline, the behaviour of the queue may be influenced. An important question is which discipline functions well for certain service-time distributions.

In this thesis we consider one possible choice for the service discipline, namely the \textit{Foreground-Background (FB) discipline}. This discipline works very well for certain service-time distributions. We describe properties of the queue that operates under the FB discipline, compare the FB queue with queues operating under other disciplines, and analyse the effect of different service-time distributions on the behaviour of the FB queue.

The thesis starts with a (introductory) description of the FB discipline, and an overview of the results on the FB queue in the literature. Then we study the maximum queue length in the FB queue, i.e. the maximum number of customers simultaneously present in the queue. This is an interesting variable in queues with a finite waiting room, or buffer. If one uses the queue length as performance measure, the FB discipline is optimal for a class of service-time distributions. This class contains many well-known distributions with so-called heavy tails, which seem to occur in internet traffic. For this class we show that the tail of the distribution of the maximum queue length decreases exponentially fast. This behaviour is opposite to that of queues with the classical \textit{FIFO} discipline.

In the ensuing part of the thesis we study the impact of the service-time distribution on the behaviour of the FB queue. We describe the effect of more variability
in the service times on certain characteristics of the queue. Surprisingly, more variability can have positive effects, for example on the maximum queue length and the stationary queue length. Taking a small detour, we describe the maximum queue length in queues with disciplines other than FB. For several disciplines we show that a certain ordering of the service times in two queues causes an ordering of the maximum queue lengths in the two queues. In particular, a more variable service-time distribution implies a shorter maximum queue length (in expectation).

Next, the thesis describes some properties of the stationary FB queue. First, we study the asymptotics of the sojourn-time distribution, the time between the arrival and departure of a customer. For so-called light-tailed distributions, the tail of the sojourn-time distribution is shown to be decreasing as slow as the tail of the busy-period length. This means that the decay rate of the sojourn time is minimal in the class of all work-conserving disciplines.

We then study the departure process. This consists of the departure times of the customers in the queue. We show that the tail of the distribution of the first departure from the stationary queue decreases exponentially fast. This means that large gaps in the departure process are unlikely. We further show that under mild conditions all moments of the stationary queue length are finite: this behaviour is again opposite to that of the FIFO queue. The behaviour of the expected queue length is studied in the case that the load $\rho$ of the queue converges to 1. Finally, we consider the queue length in the queue with constant service times: for these service times the FB discipline is the worst possible discipline.

The thesis concludes with a more theoretical chapter. In this chapter the deterministic queue is described in the form of a game. Further, we show that the queue length in a queue in continuous time may be approximated arbitrarily well by the queue lengths in a sequence of discrete-time queues. This approximation is used to prove a limiting argument that extends the optimality property (w.r.t. the queue length) of the FB discipline from discrete to continuous time.