Preface

The origin of queueing theory and its application traces back to Erlang's historical work for telephony networks as recently celebrated by the Erlang Centennial, 100 Years of Queueing, Copenhagen, recalling his first paper in 1909. Ever since, the simplicity and fundamental flavour of Erlang's famous expressions, such as his loss formula for an incoming call in a circuit switched system to be lost, has remained intriguing. It has motivated the development of results with similar elegance and expression power for various systems modeling congestion and competition over resources.

A second milestone was the step of queueing theory into queueing networks as motivated by the first so-called product form results for assembly type networks in manufacturing in the nineteen fifties (R.R.P. Jackson 1954, J.R. Jackson 1957, and E. Koenigsberg 1958, 1959). These results revealed that the queue lengths at nodes of a network, where customers route among the nodes upon service completion in equilibrium can be regarded as independent random variables, that is, the equilibrium distribution of the network of nodes factorizes over (is a product of) the marginal equilibrium distributions of the individual nodes as if in isolation. These networks are nowadays referred to as Jackson networks.

A third milestone was inspired by the rapid development of computer systems and brought the attention for service disciplines such as the Processor Sharing discipline introduced by Kleinrock in 1967. More complicated multi server nodes and service disciplines such as First-Come-First-Served, Last-Come-First-Served and Processor Sharing, and their mixing within a network have led to a surge in theoretical developments and a wide applicability of queuing theory.

Queueing networks have obtained their place in both theory and practice. New technological developments such as Internet and wireless communications, but also advancements in existing applications such as manufacturing and production systems, public transportation, logistics, and health care, have triggered many theoretical and practical results.

Queueing network theory has focused on both the analysis of complex nodes, and the interaction between nodes in networks. This handbook aims to highlight fundamental, methodological and computational aspects of networks of queues to

provide insight and unify results that can be applied in a more general manner. Several topics that are closely related are treated from the perspective of different authors to also provide different intuition that, in our opinion, is of fundamental importance to appreciate the results for networks of queues. Of course, applications of modern queueing networks are manifold. These are illustrated in the concluding chapters of this handbook. The handbook is organized in five parts.

Part 1. Exact analytical results, chapters 1–7

Product form expressions for the equilibrium distribution of networks are by far leading and have been most dominant in the literature on exact analytical results for queueing networks. In recent years, features such as batch routing, negative customers and signals have been introduced to enhance the modeling power of this class of networks. A unified theory from different perspectives is contained in the first part of this handbook. Topics include

- a characterization of product forms by physical balance concepts and simple traffic flow equations,
- classes of service and queue disciplines such as Invariant Disciplines and Order Independent queues that allow a product form,
- a unified description of product forms for discrete time queueing networks,
- insights for insensitivity from the classical Erlang loss model up to Generalised Semi Markov Processes and partially insensitive networks,
- aggregation and decomposition results that allow subnetworks to be aggregated into single nodes to reduce computational burden.

These product form results encompass a number of intriguing aspects that are not only most useful for practical purposes but also indicate a variety of open problems which remain to be tackled.

Part 2. Monotonicity and comparison results, chapters 8–9

Exact (product form) results are only available for a limited class of networks. These exact results, however, may also be invoked to obtain bounds for performance measures for intractable queueing networks. Two basic approaches can be identified:

- stochastic monotonicity and ordering results based on the ordering of the generators of the processes,
- comparison results and explicit error bounds based on an underlying Markov reward structure which leads to ordering of expectations of performance measures.

There is a clear trade-off for applying either of these two approaches. Stochastic monotonicity yields stronger results such as with non-exponential service times.

The Markov reward approach in turn is applicable under less stringent conditions, particularly with more complex structures as in a queueing network. These results are not only of theoretical and qualitative interest by themselves, but also motivate the derivation of exact analytical results to enable bounds.

Part 3. Diffusion and fluid results, chapters 10–12

Limiting regimes often allow for amenable expressions for performance measures in systems that are otherwise intractable. Two particular regimes are of interest: the fluid regime and the diffusion regime that are illustrated through the following topics:

- fluid limits for analysis of system stability,
- diffusion approximation for multi-server systems,
- system fed by Gaussian traffic to model variation in the arrival process.

These topics illustrate a rich class of systems that may be analyzed in the limiting regime and identify an important area of current research.

Part 4. Computational and approximate results, chapters 13–15

Practical applications such as in manufacturing, computer performance and communications rapidly prove to be beyond analytical solvability due to e.g. non-exponential service times, capacity constraints, synchronization or prioritization. Numerically exact or approximate approaches for averages or distributions of performance measures have been developed in literature. An illustration is provided via the following topics:

- MVA (mean value analysis) and QNA (queueing network analyzer) focusing on mean and variance of performance measures such as queue length and sojourn times,
- numerical approximation of response time distributions
- approximate decomposition results for large open queueing networks.

The numerical approach to performance analysis is a lively research community that considerably contributes to the success of queueing theory in applications as it allows for explicit numerical results for performance measures.

Part 5. Selected applications, chapters 16–18

Applications of queueing networks are manifold. To illustrate the application power of queueing theory, some special application areas and their specific queueing network aspects are enlightened:

- loss networks as originating from circuit switched telecommunications applications,
- capacity sharing as originating from packet switching in data networks,
- hospital logistics.

The first two applications have a theoretical nature as they illustrate a typical class of queueing networks. The last application illustrates a typical approach for application of queueing theory in a practical environment.

Despite the fundamental theoretical flavour of this book, it is to be kept in mind that the area of queueing theory would not have existed and would not have progressed so strongly had it not been driven by application areas that led to the various fundamental questions. The intertwined progress of theory and practice will remain to be most intriguing and will continue to be the basis of further developments in queueing theory. You are highly invited to step in.

Acknowledgments

We would like to express some words of gratitude. First of all, we are most grateful to all authors involved for their positive reactions and enthusiasm right from the start, for their cooperation and for their invaluable contributions. We wish to apologize to those authors who had already submitted at an early stage but who had to wait for such a long time until its final publication. We like to apologize to some other authors for our persistently hampering on their shoulders. We hope that you can appreciate our intention and final outcome.

We deeply owe our gratefulness to the editor-in-chief Dr. F.S. Hillier. Had it not been for his stimulation, patience and confidence in its finalization, we would not have been able to complete this book successfully.

The support by the Dutch Organisation for Scientific Research (NWO) by which the outcome of this book could be realized has been highly appreciated.

Enschede, *Richard J. Boucherie*
Amsterdam, *Nico M. Van Dijk*
June 2010

http://www.springer.com/978-1-4419-6471-7

Queueing Networks
A Fundamental Approach
Boucherie, R.J.; van Dijk, N.M. (Eds.)
2011, XXIV, 800 p., Hardcover
ISBN: 978-1-4419-6471-7