Solving Enterprise Applications Performance Puzzles: Queuing Models to the Rescue

by Leonid Grinshpan

Book Review by Alex Podelko

Solving Enterprise Applications Performance Puzzles: Queuing Models to the Rescue1 by Leonid Grinshpan is a pretty interesting book about the application of queuing models to solving enterprise performance and I believe the book fills a few gaps in practical application of queuing theory. Another good name for this book could be "Building queuing models by example".

I spent a lot of time trying to use queuing models to solve practical performance issues and would testify that it is pretty challenging. There are a few areas where it was developed a little further (for example, around capacity planning of existing systems), but if you trying to do something else – you won’t find much help. You have a lot of books about systems’ performance, you have a lot of books about queuing theory with simple examples, but not much in between to solve practical tasks. And here Grinshpan’s book may help, especially if you are new in this area.

Chapter 1, Queuing Networks as Applications Models, is an introduction into the topic. It discusses how queuing theory may be used to model enterprise applications. A lot of analogies are used to introduce the subject.

Chapter 2, Building and Solving Application Models, is an overview of the whole process, including short discussions about essentials of queuing theory and using of tools to solve models.

Chapter 3, Workload Characterization and Transaction Profiling, discusses what the input data for models are and how to gather them.

Chapter 4, Servers, CPUs, and Other Building Blocks of Application Scalability, discusses scalability, bottlenecks, how to identify bottlenecks and ways to fix them (mostly with CPU and I/O examples).

Chapter 5, Operating System Overheads, discusses the main components of operating systems, where overheads come from, how to measure them, and their impact on transaction time.

Chapter 6, Software Bottlenecks, is devoted to software bottlenecks, which are rarely discussed in application to queuing models – while in practice software bottlenecks happen all the time. Memory bottlenecks and thread optimizations and their modeling are discussed in detail. Multiple other software bottlenecks are also reviewed.

Chapter 7, Performance and Capacity of Virtual Systems, is an overview of performance issues related to virtualization, their explanation with queuing theory, and a methodology of virtual machine sizing.

---

1 Available at http://www.amazon.com/Solving-Enterprise-Applications-Performance-Puzzles/dp/1118061578
Chapter 8, Model-Based Application Sizing: Say Good-Bye to guessing, explains why to use model-based sizing and discusses it step-by-step from gathering input data to model deliverables and what-if scenarios.

Chapter 9, Modeling Different Application Configurations, discusses several specials cases including geographical distribution of users, cross-platform modeling, remote terminal services, load balancing, and parallelization of transactions.

The book covers a lot of topics. However, to avoid disappointments, I'd like to point out what this book is not:

- It is not a textbook about queuing theory. The section 2.2 Essentials of Queuing Networks Theory has 5 pages in it.
- It is not a book about tools to solve queuing models. Available tools are listed and there are references, but they are just mentioned as a way to solve models (with one tool used as an illustration of the process). You don't need to know any tool to read the book (but you will need one when you try to solve your own models).
- It is not a comprehensive book about enterprise application performance. There is plenty of important information and practical recommendations about enterprise application performance in the book, but it is shared as needed to build models and analyze their results.

So the book is exactly what the title says: a practical book about building queuing models to investigate enterprise applications performance issues.