

## IEOR 8100: Spring 2012, Professor Whitt

### Performance Analysis: Looking at Systems and System Data Through the Lens of Queueing Models

Queueing models are useful in performance analysis, but there is much more to successful performance analysis than queueing models and their analysis. There is also much more than looking at system data. When considering queueing models and looking at system data, it is good to be aware of the typical full performance analysis process.

Steps in Performance Analysis, Exploiting System Data:

1. What is **the system** to consider? [It might be a factory, a computer, a communication network or a service facility. Among service facilities, it might be a bank, call center or a hospital, or a single ward in the hospital.] How is the system supposed to operate? How does the system actually operate? [Directly experience the system and consult system experts.] What are the key resources? What are the relative costs of these resources? What are the bottlenecks? What are the key customer (or job) classes? Do different classes have high or low priorities? What are the flows? What decisions are made in system operation. e.g., scheduling and routing? What are the precedence constraints? What are the sources of uncertainty? What model(s) should (or might) be appropriate? What models have been used in the past, for the specific system and for this class of systems in the literature? [This step might involve a flow analysis and a rough network of queues.]
2. What are **the performance questions**? Is there a clear, well-defined goal? What are the performance goals? What questions do we want to answer? Are we designing a new system, changing an existing system, or managing an existing system? Are there existing performance problems that need correcting? Are we seeking greater efficiency, greater revenue or do we want to improve the quality of service? Or all of the above? Or are we just checking that an ongoing system is operating OK?
3. What **queueing models** might provide insight? Are variants of the basic single-server ( $M/M/1$  and  $GI/GI/1$ ) models relevant? Are variants of the Erlang  $B$ ,  $C$  and  $A$  models (i.e.,  $M/M/s/0$ ,  $M/M/s/\infty$  or  $M/M/s/\infty + M$ ) relevant, including the non-Markovian extensions (i.e.,  $G/GI/s/0$ ,  $G/GI/s/\infty$  or  $G/GI/s/\infty + GI$ )? Are stationary models appropriate? Or, instead, are time-varying arrival rates important?
4. What is revealed by looking at **system data**? [When we look at system data, it is good to have already considered the issues above.] Do the detailed system data support the initial view of the system (step 1 above)? Do the detailed system data support the initial concerns and goals (step 2 above)? Do data reveal new issues beyond the ones anticipated in advance? What do the data indicate about appropriate queueing models (step 3 above)? What statistical analyses are needed to better understand what we can properly deduce from the data? For example, when we estimate parameters by sample averages, what are the associated confidence intervals?
5. How can we use the data to help **pick useful queueing models**? Do we need to consider non-stationary models to account for time-varying arrival rates? Can a first-order deterministic model be useful or do we need a stochastic model? If we plan to use a stochastic model, what is the simplest model consistent with the data? In general,

we typically want to balance ability to analyze the model with the need for realistic representation. On the one hand, we need to be able to analyze the model, but on the other hand, we want the model to realistically represent the system? [We need to know something about what models can be analyzed when we make this choice. Success here depends on a reasonable understanding of *both* (i) queueing models and their analysis, and (ii) the actual system and the system data.]

6. Given various candidate models, what does **the analysis of these queueing models** reveal?
7. Are the results of the model analysis **consistent with system data**? Should we exploit simulation to further investigate and confirm the model analysis?
8. What are **tentative answers** to the original performance questions?
9. **Implementation.** Improvements need to be implemented to make a difference. Performance analysis is intended to yield performance improvement. [Good analysis in all the steps above can fail because this critical step is not understood and managed properly.]
10. **Follow up.** Check to see if system changes based on performance analysis actually achieve their goals. That calls for the careful analysis of new system data after the performance changes have been implemented.
11. Employ the **scientific method**. Keep cycling through the entire process, raising new questions, considering new models, collecting new data, analyzing models, performing simulations, making new changes, etc. etc. Aim for continuous process improvement.