# INTRODUCTION TO COMPUTER AND NETWORK PERFORMANCE ANALYSIS USING QUEUING SYSTEMS


## PROF. VARSHA APTE
Department Of Computer Science & Engineering
IIT Bombay


**PRE-REQUISITES :** At least 3rd year UG in CSE Pre-requisite Courses : Operating Systems and Computer Networks
**INTENDED AUDIENCE :** 4th year undergraduates or post-graduates in Computer Science and Engineering, IT industry professionals engaged in computer applications performance testing and evaluation
**INDUSTRY SUPPORT :** Any IT company

**COURSE OUTLINE :**
Analyzing the performance of any computer or networked system: such as Web application servers, packet scheduling disciplines, operating system schedulers, cellular telephony networks is an important step in the design and deployment of such systems. Performance of many systems can be measured, but a sound basis in queuing system models is required for two important aspects: 1) to ensure that the performance tests are performance field data is correct, and 2) for predicting the performance of a system in a scenario that cannot be measured. The theory of queues is a mathematical theory that helps model a wide range of computing and networking systems with a common abstraction of a queuing system. In this course, you will learn the most basic results in queuing systems, in an intuitive way, and learn how to apply them to computer and network systems performance. The focus in this short course will be on being able to reason about asymptotic values of performance metrics at high loads and low loads, which is very useful in validating results of performance tests. Practical examples of performance analysis of networked servers (e.g. a Web server), of packet network links, of cellular networks etc, will be provided. A case study of interpreting the load test results of a simple Web Server using the framework of queuing systems will be covered throughout the course.

**ABOUT INSTRUCTOR :**
Prof. Varsha Apte is a Professor in the Department of Computer Science and Engineering at IIT Bombay. She works in the areas of performance analysis of computing systems and networks. She completed her M.Sc. in from Pune University in 1989, and PhD. from Duke University in 1994, both in Computer Science. After her Ph.D., she joined the Teletraffic Theory and Performance Analysis department in AT&T Bell Labs as Member, Technical Staff, which later became the Network Design and Performance Analysis department of AT&T Labs. She joined as faculty in the CSE Dept, IIT Bombay in 2002. During the academic years 2009-2011, she was Visiting Faculty at the Computer Science and Automation Department, Indian Institute of Science, Bangalore. While in Bangalore, she also worked at IBM Research Labs as part-time Visiting Researcher from June 2009-May 2010. For more information, please visit: https://www.cse.iitb.ac.in/~varsha/

**COURSE PLAN :**
**Week 1 :**
1.1 Introduction, why do delays happen, contention for resources
1.2 Performance metrics and parameters
1.3 Intro to queuing system: Standard parameters and metrics, Kendall Notation. Metrics of open queuing systems
1.4 Intro to Memorylessness
1.5 Operational Laws,Utilization Law, Throughput, stability of a queuing system

**Week 2 :**
2.1 Asymptotic Analysis of G/G/1, G/G/1/K queues (Values of metrics at low load and high load asymptotes)
2.2 Asymptotic Analysis of G/G/c/K queues, Examples
2.3 Little's Law - Intro and discussion
2.4 ExDmples for Little's LDw Dnd D CDse Study of DpplicDtion of queing theory (open systems)
2.5 Some results for M/G/1 queues and Memoryless arrivals

**Week 3 :**

3. 1 Case Study:Experimental Performance Measurement of a Web Server (open load)atch.
3.2 Open queuing networks - tandem queuing network
3.3 Open queuing networks - general jackson queuing network
3.4 Open queuing networks - examples
3.5 Closed Queuing Systems. Metrics, parameters. Analysis of simplest closed queueing system

**Week 4 :**

4.1 Closed Queuing System: Low Load and High Load Asymptotes of all metrics. Response Time linear asymptote, Kleinrock's Saturation Number Heuristic
4.2 Case Study: Experimental Performance Measurement of a Web Server (closed load)
4.3 General formulation of Jacksonian Closed Queuing Networks
Arrival Theorem, Mean Value Analysis (Derivation)
4.4 Mean Value Analysis - more explanation
4.5 Mean Value Analysis examples: concluding Case Study of a Load test on a web server. Discuss applications and limitations of queueing systems based modeling