Queuing Theory References

Jain, R., *The Art of Computer System Performance Analysis*, 1991, Wiley, New York

Lazowska, E.D., et al, *Quantitative System Performance: Computer System Analysis Using Queueing Network Models*, 1984, Prentice-Hall, Englewood Cliffs, NJ

Kendall Notation

A/S/m/B/K/SD

A – Interarrival time distribution
S – Service time distribution
m – Number of servers
B – Number of buffers (system capacity)
K – Population size
SD – Service Discipline

Often, only the first 3 parameters are specified, in which case B and K are infinite, SD is FCFS

Kendall Notation

Common time distributions:

M – Exponential (Memoryless)
E_k – Erlang with parameter k
H_k – Hyperexponential with parameter k)
D – Deterministic
G – General

Example Queueing System

M/M/1

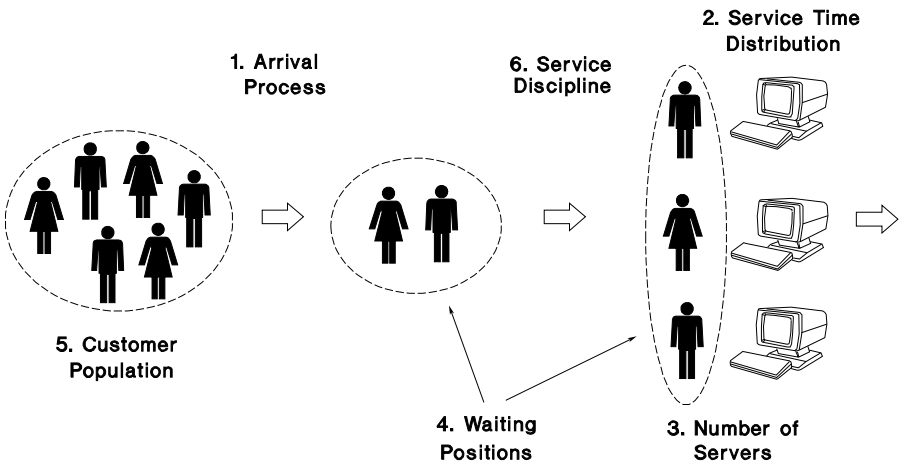
Exponential Arrival time
Exponential service time
1 server
Infinite buffers
Infinite population size
First Come, First Served service discipline

Example Queueing System

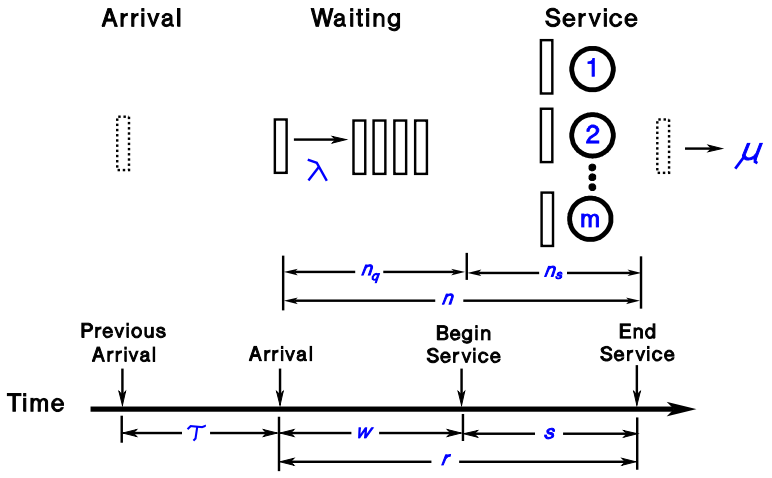
M/M/3/20/1500/FCFS

Exponential Arrival time
Exponential service time
3 servers
20 buffers (3 being services, up to 17 waiting)
Population size of 1500
First Come, First Served service discipline

Queueing System Components



Queue Terminology

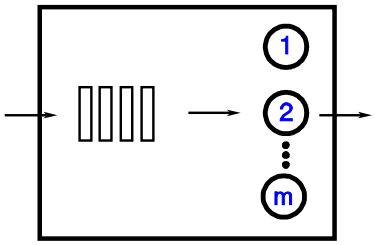


Queue070

Little's Law

For a stable system:

Mean number in system = arrival rate * mean response time



Mean number in queue = arrival rate * mean waiting time

Mean number in service = arrival rate * mean service time

Queue080