

Performance Modeling of Servers - Queueing Systems

Kai Shen

Dept. of Computer Science, University of Rochester

Some Queueing Theory

If the facts don't match the theory, change the facts.

- Albert Einstein

- Network server model in queueing systems
 - arrival process (inter-arrival time distribution if we assume IID): exponential (M, or memoryless), don't know (G)
 - service time distribution: exponential (M), don't know (G)
 - number of servers

URCS 573 - Spring 2008

2

Stochastic Processes

- Stochastic processes
 - Random functions of time sequences
 - input can be either continuous time or discrete sequences
 - output can be either continuous state or discrete state
 - a discrete-state process is also called a stochastic chain
- Examples
 - $n(t)$ - number of requests in the system at time t
 - $s(k)$ - service time of the k -th request
 - random walk visitation in a peer-to-peer network at step k

1/24/2008

URCS 573 - Spring 2008

3

Specific Stochastic Processes

- Markov processes
 - future state transition only depend on the current (independent of the past)
 - with continuous-time, it is sufficient to model the future while knowing the current state (not need to know long the system has been in current state) - memoryless distribution is good.
 - M/M/m queues can be modeled as Markov processes
 - specifically, $n(t)$ in M/M/m queues is a Markov chain
- Birth-death processes
 - a type of Markov chains
 - transitions are restricted to neighboring states in a chain
 - $n(t)$ in M/M/m queues is also a birth-death process

1/24/2008

URCS 573 - Spring 2008

4

Specific Stochastic Processes (more)

- Notations:
 - $a(t)$ - cumulative arrival number up to time t
 - $d(t)$ - cumulative departure number up to time t
- Poisson processes
 - if inter-arrival time is IID and exponential, $a(t)$ is a Poisson process
 - is a Poisson process also a birth-death process?
- Interesting properties
 - multiple Poisson processes merge into one Poisson process with summed mean rate
 - $d(t)$ of an $M/M/m$ queue is also a Poisson process with the same mean rate for $a(t)$

1/24/2008

URCS 573 - Spring 2008

5

Steady-State Probability

Birth-death process

- p_n is the steady-state probability of being in state n
- r_i is the transition rate from state i to state $i+1$
- u_i is the transition rate from state $i+1$ to state i
- then we have
 - $p_n = p_0 \times (r_0 \times \dots \times r_{n-1}) / (u_1 \times \dots \times u_n)$

Now look at $n(t)$ in an $M/M/1$ queue

- $r_0 = \dots = r_{n-1}; u_1 = \dots = u_n$
- then $p_n = (1-r/u) \times (r/u)^n$

1/24/2008

URCS 573 - Spring 2008

6

Open vs. Closed Systems

- Open system
 - arrival process is independent of service rate
- Closed system
 - each completed request trigger a new request (sometimes after certain thinktime)
- Which one is more realistic?
- Mixed model
 - Schroeder et al. NSDI2006

1/24/2008

URCS 573 - Spring 2008

7

Operational Laws

All professional men are handicapped by not being allowed to ignore things which are useless.

- Johann Wolfgang von Goethe

- Operational laws
 - those that don't require assumptions that cannot be validated through black-box operational observations
 - **operational metrics**: arrival and departure time of requests (no information on corresponding pairs), probably system busy periods, ...
- Little's Law
 - mean number in system = arrival rate \times mean time in system
 - assumption: number of arrivals equals number of completions

1/24/2008

URCS 573 - Spring 2008

8



Disclaimer

- Most materials in these slides were developed from the book "The Art of Computer Systems Performance Analysis", R. Jain, 1991, Wiley.