Open Models

Riccardo Pinciroli
Email: riccardo.pinciroli@polimi.it
Phone: 02 2399 3494

Open model solution technique

processing capacity : $\lambda_{sat} = 1 / D_{max}$

throughput : $X(\lambda) = \lambda$

utilization : $U_k(\lambda) = \lambda D_k$

residence time :

$$R_k(\lambda) = \begin{cases} D_k & \text{(delay centers)} \\ \dfrac{D_k}{1 - U_k(\lambda)} & \text{(queueing centers)} \end{cases}$$

queue length :

$$Q_k(\lambda) = \lambda R_k(\lambda) = \begin{cases} U_k(\lambda) & \text{(delay centers)} \\ \dfrac{U_k(\lambda)}{1 - U_k(\lambda)} & \text{(queueing centers)} \end{cases}$$

system response time : $R(\lambda) = \sum_{k=1}^{K} R_k(\lambda)$

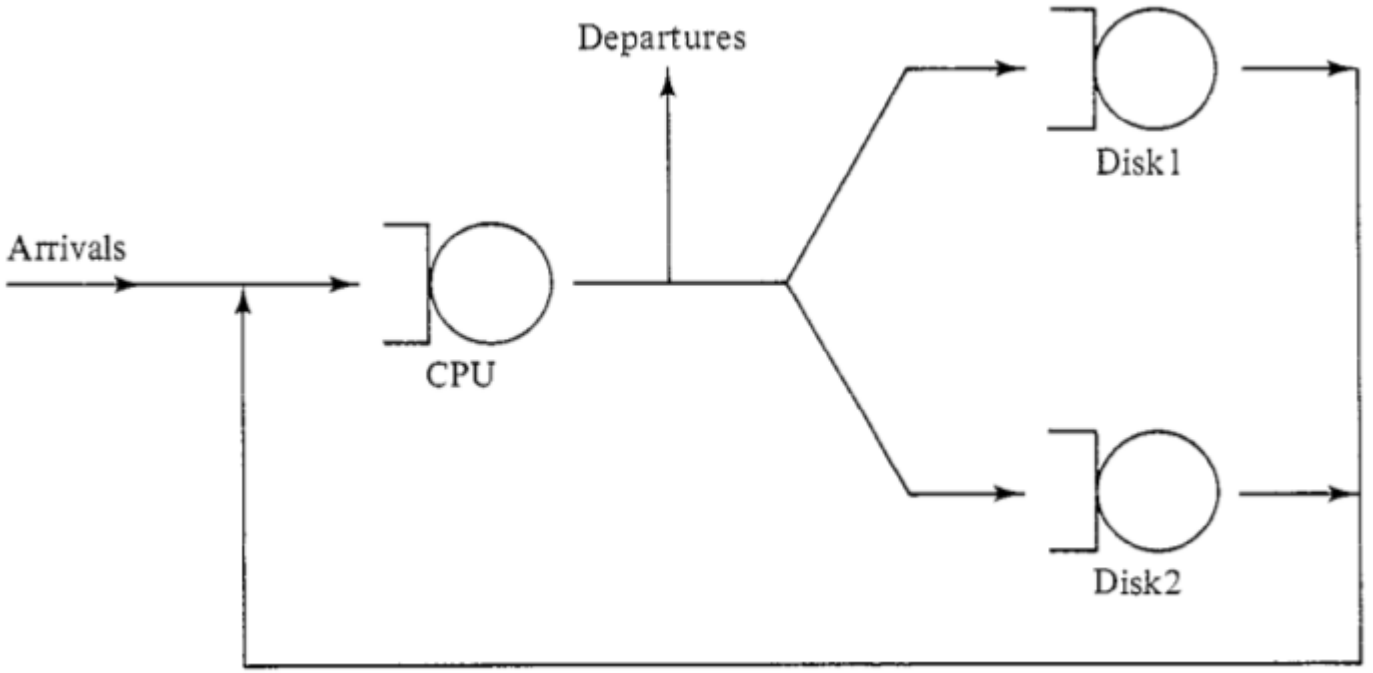
average number in system : $Q(\lambda) = \lambda R(\lambda) = \sum_{k=1}^{K} Q_k(\lambda)$

Exercise 1 (Lazowska, pag. 113)

$V_{CPU} = 121$ $V_{Disk1} = 70$ $V_{Disk2} = 50$

$S_{CPU} = .005$ $S_{Disk1} = .030$ $S_{Disk2} = .027$

$\lambda = 0.3$ jobs/sec.



Exercise 1 (Lazowska, pag. 113)

- D_k = ? (for each k)
- X_k, U_k, R^{es}_k, R_k, N_k = ? (for each k)
- N_k^{queue}, Q_k, Q_k^{visit} = ? (for each k)
- λ_{max}= ?
- R, N = ?

Exercise 2

An on-line store is deployed on a farm composed by three servers (thereafter named A,B,C). To improve the response time of the website, the system is measured and the following data are collected.

System jobs completed $C = 10000$
Server A completed jobs $C_A = 2200$
Server B completed jobs $C_B = 3000$
Server C completed jobs $C_C = 4800$

Server A busy time $B_A = 35$ min
Server B busy time $B_B = 24$ min
Server C busy time $B_C = 55$ min

Observation time $T = 2$ Hours

Exercise 2

(a) System throughput X and system arrival rate λ:
   i. Servers throughput X_A, X_B and X_C
   ii. Utilization U_A, U_B and U_C as well as servers demands D_A, D_B, D_C
   iii. Visits to each station V_A, V_B, V_C
   iv. the arrival rate that saturates the system

(b) Mean queue length Q_A (including the job currently served), the global response time R.

Consider now the model as a closed one with $Z = 1$.

(c) Determine the system bottleneck. With users $N = 6$, the system is light or heavy loaded?

(d) Draw and describe analytically the asymptotes $X(N)$ and $R(N)$.

Exercise 3

We have a single-class model with two service centers.
We have the following information about the system:

- Station 1 **response** time: 3 sec
- Station 2 **response** time: 8 sec
- Station 1 throughput: 3 transactions/sec
- Station 2 throughput: 4 transactions/sec
- System throughput: 2 transactions/sec

- D_1, D_2 = ?
- S_1, S_2 = ?

Exercise 4

A performance analysis study of a medium-scale intranet should be undertaken to determine if the services provided meet the performance targets. The intranet is modeled by an *open queuing network* and consists of three servers, namely A, B and C, representing the web, application and database servers respectively. To evaluate the performance, a 1 hour monitoring period has been carried out. The following data were collected:

- observation interval T= 1 hour
- number of completions of the intranet: $C = 720 jobs$
- average number of jobs in server A $N_A = 1 job$
- average number of jobs in server B $N_B = 4 jobs$
- utilization of server C $U_C = 0.90909$

Exercise 4

Compute:

(a) the service demands D_A, D_B, D_C of all the servers

(b) the number of jobs in server C

(c) the utilization of servers A and B

(d) the system response time R (the response time of the intranet)

(e) which servers should be replicated in order to support a throughput of the intranet of $\lambda = 0.3j/s$? Compute the number of replicas required. Assume that the arriving requests are uniformly distributed across the parallelized servers.

Suppose that in the intranet the average number of jobs N is constant and consider the closed queuing network composed of servers A, B and C and a delay station with Z=10s and the same service demands D_A, D_B, D_C of the open queuing network.

(f) describe (with graphs and equations) the asymptotic upper and lower bounds on system throughput X(N) and response time R(N) as a function of the population size N

(g) with 7 users it will be possible to have a response time of 15s? (*motivate the answer*)

Exercise 5

The Intranet of a medium scale company consists of three servers, namely A, B and C, which represent the web browser of the clients, the application server and the database server, respectively. The intranet is modelled by an open queueing network. In order to evaluate the performance of the system a 10 minutes monitoring phase has been performed. The following data have been collected:

C_A	Server A number of completions	500
C_B	Server B number of completions	150
C_C	Server C number of completions	300
C	Network (intranet) completions	75
B_A	Server A busy time	450s
B_B	Server B busy time	300s
B_C	Server C busy time	100s

Exercise 5

Compute:

a) the utilizations of all the servers
b) the service demands of all the servers, determine which is the bottleneck
c) the number of visits at server A
d) the service time at server A
e) the arrival rate that saturates the system
f) the response time of server A and the number of requests in A

Consider the closed queueing network composed of servers A, B and C and a delay with Z=10s and the same service demands of the open queueing network.

g) Represent the asymptotic bound X(N) and R(N) on system throughput and response time varying the population size N
h) Determine the minimum number of requests N such that after this value the system is in the heavy load condition.
i) Assume that the system works in the heavy load condition. What is the maximum and minimum system response time?

Exercise 6

The Intranet of a medium scale company consists of three servers, namely A, B and C, which represent the web server of the clients, the application server and the database server, respectively. The intranet is modelled by an open queueing network. In order to evaluate the performance of the system a 10 minutes monitoring phase has been performed. The following data have been collected:

C_B	Server B number of completions	150
C_C	Server C number of completions	300

C	Network (intranet) completions	100

B_B	Server B busy time	300s
B_C	Server C busy time	100s

It is also known that the maximum throughput achievable by the intranet is 0.2 trans/sec.

Exercise 6

Compute:

a) the service demands of all the servers and determine the server which should be upgraded to achieve the maximum gain of the network performance

b) the utilizations of all the servers

c) the number of visits at server B

d) the residence time of server A and the number of requests in A

e) the minimum number of identical web servers to put in parallel such that the residence time at the web server layer is less than 18 seconds. Assume that incoming requests are uniformly distributed across the web servers.

Consider the closed queueing network composed of servers A, B and C and a delay with Z=10s and the same service demands of the open queueing network without considering the servers added at point e).

f) Represent the asymptotic bound X(N) and R(N) on system throughput and response time varying the population size N

g) What is the maximum value of N which guarantees that the response time is less than 40 sec?