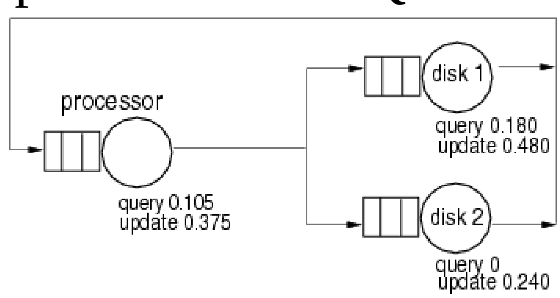
CS 672 – Multiclass Queuing Networks

Dr. Daniel A. Menascé
http://www.cs.gmu.edu/faculty/menasce.html
Department of Computer Science
George Mason University

1

A Simple Two Class QN Model



Server with N transactions in execution

Questions:

- What is the predicted increase in the throughput of query transactions if the load of update transactions is moved to off-peak hours?
- How will the response time change if the total I/O load of query transactions is moved to disk 2?

2

Notation for Closed Multiclass QNs

$R:$ number of classes

$\vec{N} = (N_1, \ldots, N_r, \ldots, N_R):$ population vector

$\vec{1}_r = (0, \ldots, 1, \ldots, 0):$ vector in which all elements except for the r-th, which is 1, are zero.

$R'_{i,r}(\vec{N}):$ avg. residence time at device i for class r customers.

$R_{i,r}(\vec{N}):$ avg. response time at device i for class r customers.

$X_{0,r}(\vec{N}):$ avg. system throughput for class r customers.

$\bar{n}_{i,r}(\vec{N}):$ avg. queue length at device i for class r customers.

$\bar{n}_i(\vec{N}):$ avg. queue length at device i for all classes.

$\bar{n}^{A}_{i,r}(\vec{N}):$ avg. queue length at device i seen by an arriving class r customers.

3

The BCMP Theorem

- BCMP: Baskett, Chandy, Muntz, and Palacios, "Open, closed, and mixed networks of queues with different classes of customers," JACM, April 1975.
- Network state: $\vec{n} = (\vec{n}_1, \ldots, \vec{n}_K)$

$$\vec{n}_i = \left(n_{i,1}, \ldots, n_{i,r}, \ldots, n_{i,R}\right)$$

no. of class r customers at device i.

- BCM theorem: condition for product form solution:

$$p(\vec{n}) = p(\vec{n}_1) \times \ldots \times p(\vec{n}_i) \times \ldots p(\vec{n}_K)$$

4

BCMP Theorem Assumptions

- FCFS: service time distribution is exponential with the same mean for all classes. Visit ratios may be different. The service rate can be load dependent but it can only depend on the total number of customers at the device.
- PS: processor-sharing discipline. Each class may have a distinct service time distribution.
- IS: infinite servers (or ample number of servers or delay server). No queuing.
- LCFS-PR: Last-Come First-Served Preemptive Resume. Each class may have a distinct service time distribution.

MVA Formulas for Multiclass QNs

$$\text{Arrival Theorem}: \overline{n}^{A}_{i,r}(\vec{N}) = \overline{n}_i(\vec{N} - \vec{1}_r)$$

$$\text{So}: R'_{i,r}(\vec{N}) = D_{i,r}\left[1 + \overline{n}_i(\vec{N} - \vec{1}_r)\right]$$

$$X_{0,r}(\vec{N}) = N_r \Big/ \sum_{i=1}^{K} R'_{i,r}(\vec{N})$$

$$\overline{n}_{i,r}(\vec{N}) = X_{0,r}(\vec{N}) R'_{i,r}(\vec{N})$$

$$\overline{n}_i(\vec{N}) = \sum_{r=1}^{R} \overline{n}_{i,r}(\vec{N}) = \sum_{r=1}^{R} X_{0,r}(\vec{N}) R'_{i,r}(\vec{N})$$

MVA Formulas for Multiclass QNs

$$R_{i,r}(\vec{N}) = S_{i,r}\left[1+\bar{n}^{A}_{i,r}(\vec{N})\right]$$

$$R'_{i,r}(\vec{N}) = V_{i,r}R_{i,r}(\vec{N}) = V_{i,r}S_{i,r}\left[1+\bar{n}^{A}_{i,r}(\vec{N})\right] = D_{i,r}\left[1+\bar{n}^{A}_{i,r}(\vec{N})\right]$$

$$X_{0,r}(\vec{N}) = N_r \,/ \sum_{i=1}^{K} R'_{i,r}(\vec{N})$$

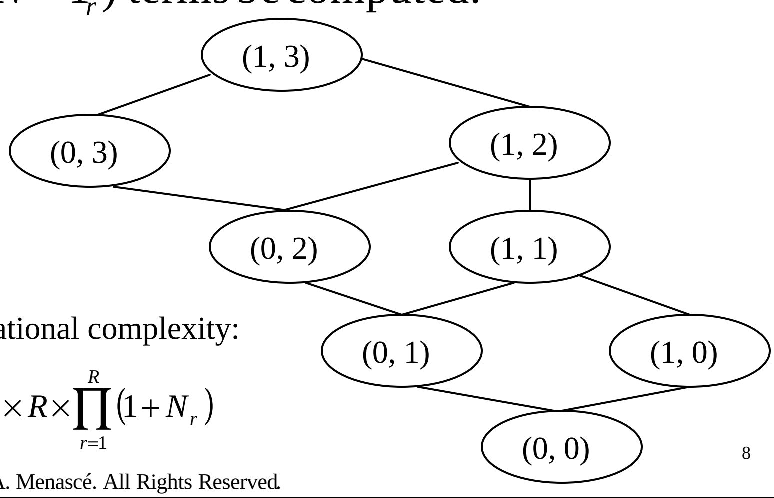
$$\bar{n}_{i,r}(\vec{N}) = X_{i,r}(\vec{N})R_{i,r}(\vec{N}) = X_{0,r}(\vec{N})V_{i,r}R_{i,r}(\vec{N}) = X_{0,r}(\vec{N})R'_{i,r}(\vec{N})$$

$$\bar{n}_{i}(\vec{N}) = \sum_{r=1}^{R} \bar{n}_{i,r}(\vec{N}) = \sum_{r=1}^{R} X_{0,r}(\vec{N})R'_{i,r}(\vec{N})$$

7

MVA Formulas for Multiclass QNs

The term $\bar{n}_i(\vec{N}-\vec{1}_r)$ requires that all R $\bar{n}_{i,r}(\vec{N}-\vec{1}_r)$ terms be computed.



Computational complexity:

$$K \times R \times \prod_{r=1}^{R}(1+N_r)$$

8

Approximate MVA for Multiclass Closed QNs

Bard - Schweitzer Approximation :

$$\bar{n}_{i,r}(\vec{N}-\vec{1}_r) = \frac{N_r - 1}{N_r}\bar{n}_{i,r}(\vec{N})$$

$$\text{So, } \bar{n}_i(\vec{N}-\vec{1}_r) = \frac{N_r - 1}{N_r}\bar{n}_{i,r}(\vec{N}) + \sum_{s=1 \& s \neq r}^{R} \bar{n}_{i,s}(\vec{N}) =$$

Therefore, need $\bar{n}_{i,r}(\vec{N})$ to compute $\bar{n}_i(\vec{N}-\vec{1}_r)$.

Solution : start with initial value for $\bar{n}_{i,r}(\vec{N})$ as :

$$\bar{n}^e_{i,r}(\vec{N}) = N_r / K_r \text{ where } K_r = \left|\{i \mid D_{i,r} \neq 0\}\right|$$

9

Approximate MVA for Multiclass QNs

$$\text{Step 1}: \bar{n}^e_{i,r}(\vec{N}) = N_r / K_r \quad \forall \quad i,r$$

$$\text{Step 2}: \bar{n}_i(\vec{N}-\vec{1}_r) = \frac{N_r - 1}{N_r}\bar{n}^e_{i,r}(\vec{N}) + \sum_{s=1 \& s \neq r}^{R} \bar{n}^e_{i,s}(\vec{N}) \quad \forall \quad i,r$$

$$\text{Step 3}: R'_{i,r}(\vec{N}) = D_{i,r}\left[1 + \bar{n}_i(\vec{N}-\vec{1}_r)\right] \quad \forall \quad i,r$$

$$\text{Step 4}: X_{0,r}(\vec{N}) = N_r / \sum_{i=1}^{K} R'_{i,r}(\vec{N}) \quad \forall \quad r$$

$$\text{Step 5}: \bar{n}_{i,r}(\vec{N}) = X_{0,r}(\vec{N}) R'_{i,r}(\vec{N}) \quad \forall \quad i,r$$

If $\max_{i,r}\left\{\text{abs}\left(\left[\bar{n}^e_{i,r}(\vec{N}) - \bar{n}_{i,r}(\vec{N})\right] / \bar{n}_{i,r}(\vec{N})\right)\right\} > e$

then $\bar{n}^e_{i,r}(\vec{N}) = \bar{n}_{i,r}(\vec{N}) \quad \forall \quad i,r$ and go to step 2.

10

Notation for Open Multiclass QNs

R: number of classes

$\vec{l} = (l_1, ..., l_r, ..., l_R)$: arrival rate vector

$R'_{i,r}(\vec{l})$: avg. residence time at device i for class r customers.

$R_{i,r}(\vec{l})$: avg. response time at device i for class r customers.

$X_{0,r}(\vec{l})$: avg. system throughpu t for class r customers.

$\bar{n}_{i,r}(\vec{l})$: avg. queue length at device i for class r customers.

$\bar{n}_i(\vec{l})$: avg. queue length at device i for all classes.

$\bar{n}^{A}_{i,r}(\vec{l})$: avg. queue length at device i seen by an arriving class r customers.

11

Formulas for Open Multiclass QNs

Arrival Theorem: $\bar{n}^{A}_{i,r}(\vec{l}) = \bar{n}_i(\vec{l})$

So: $R'_{i,r}(\vec{l}) = D_{i,r}\left[1 + \bar{n}_i(\vec{l})\right]$

$X_{0,r}(\vec{l}) = l_r$

Little's Law: $\bar{n}_{i,r}(\vec{l}) = X_{0,r}(\vec{l}) R'_{i,r}(\vec{l})$

$$\Rightarrow \bar{n}_{i,r}(\vec{l}) = l_r R'_{i,r}(\vec{l}) = l_r D_{i,r}\left[1 + \bar{n}_i(\vec{l})\right]$$

$$= U_{i,r}(\vec{l})\left[1 + \bar{n}_i(\vec{l})\right]$$

For any two classes r and s:

$$\frac{\bar{n}_{i,r}(\vec{l})}{\bar{n}_{i,s}(\vec{l})} = \frac{U_{i,r}(\vec{l})}{U_{i,s}(\vec{l})}$$

Since $\bar{n}_i(\vec{l}) = \sum_{s=1}^{R} \bar{n}_{i,s}(\vec{l})$

$$\Rightarrow \bar{n}_{i,r}(\vec{l}) = \frac{U_{i,r}(\vec{l})}{1 - U_i(\vec{l})}$$

$$\boxed{\text{Little's Law}: R'_{i,r}(\vec{l}) = \frac{D_{i,r}}{1 - U_i(\vec{l})}}$$

12

Formulas for Open QN Models

$$U_{i,r} = l_r \times D_{i,r}$$

$$U_i = \sum_{r=1}^{R} U_{i,r}$$

$$R'_{i,r} = \frac{D_{i,r}}{1 - U_i}$$

$$R_r = \sum_{i=1}^{K} R'_{i,r}$$

Mixed Class Models

- Classes 1, …, O are open and classes O+1, O+2, …, O+C are closed.
- The O open classes are characterized by the vector of arrival rates $\vec{l} = (l_1, ..., l_O)$
- The C closed classes are characterized by the population vector $\vec{N} = (N_{O+1}, ..., N_{O+C})$

Solution of Mixed Class Models

Step 1: Solve the open submodels and obtain:

$$U_{i,r}(\vec{l}) = l_r D_{i,r} \quad \forall \quad i,r$$

Step 2: Find the total utilization of all open classes

$$U_i^{open} = \sum_{r=1}^{O} U_{i,r}(\vec{l}) \quad \forall \quad i$$

Step 3: Elongate service demands of closed classes:

$$D_{i,r}^{e} = \frac{D_{i,r}}{1 - U_i^{open}} \quad \forall \quad i,r$$

Step 4: Use MVA and solve closed model to find:

$$R'_{i,r}(\vec{N}), \overline{n}_{i,r}(\vec{N}), X_{O,r}(\vec{N}) \quad \forall \quad r = O+1, \ldots, O+C$$

Step 5: Find $\overline{n}_{i,closed}(\vec{N}) = \sum_{r=O+1}^{O+C} \overline{n}_{i,r}(\vec{N}) \quad \forall \quad i$

Step 6: Compute metrics for open submodel:

$$R'_{i,r}(\vec{l}) = \frac{D_{i,r}\left[1 + \overline{n}_{i,closed}(\vec{N})\right]}{1 - U_i^{open}}$$

$$\overline{n}_{i,open}(\vec{l}) = \sum_{r=1}^{O} l_r R'_{i,r}(\vec{l})$$

15