# Queueing networks with station breakdowns and globally gated service regime*

D. Kofman[a] and U. Yechiali[b]

[a]Département Réseaux, E.N.S.T, 75634 Paris Cedex 13, France

[b]Department of Statistics and Operations Research, School of Mathematical Sciences, Sackler Faculty of Exact Sciences, Tel Aviv University, Tel Aviv 69978, Israel

Local and Metropolitan Area Networks are often modeled as Polling Systems with a single server cycling between the stations. Many works in the literature studied various aspects of those models, but only a few tackled the problem of station breakdowns. Recently Kofman and Yechiali [7] analyzed polling systems with failing nodes under the Gated and the Exhaustive protocols. In this work we investigate such a system under the Globally Gated service regime, where each station is subject to a random failure process. If a node fails during a visit of the server, the breakdown is discovered only at the end of the service (e.g. transmission of a message), whereupon the work at that station stops and the server moves to the next node. The service of the interrupted job will be re-attempted during the next (or a future) visit of the server.

We present an analysis of this important model and derive various performance measures needed for efficient design, development and operation of such networks.

## 1. INTRODUCTION

Local and Metropolitan Area Networks are commonly studied as polling systems (see e.g. [8–12]), where a single server moves among the stations in a cyclic order. However, only a few works in the literature tackled the important problem of station breakdowns. Recently [7], we investigated systems with failing nodes under the Gated and Exhaustive service disciplines. In the present work we analyze a polling system under the Globally-Gated service regime, where each station (node, channel) is subject to a random failure process.

We consider two modes of station breakdowns: node-dependent, and general. If the server enters a non-empty station and starts serving customers (jobs, messages) present there, the station may fail due to a node-dependent breakdown process. We assume that if a breakdown occurs, it is observed only at the end of a service (e.g. transmission of a message), whereupon the server switches to the next station, while the interrupted job (message) will have to be processed (transmitted) again during the next (or a future) visit

*This work was supported by a grant under the French-Israeli Scientific Cooperation in Information Superhighways and Information Technology 1996-1998.

of the server. If the node-dependent failure process is Exponential, the system resembles the so-called Cyclic-service Queue with Bernoulli Schedule (see Tedijanto [11] and references there). Under that discipline, when visiting a non-empty channel $i$ ($i = 1, 2, \ldots, N$), the server always serves the first customer. Then, after every service completion, it serves the next customer (if the queue is still not empty) with probability $a_i$, or moves to the next station with probability $1-a_i$. The Bernoulli Schedule reduces to the classical Exhaustive regime if $a_i = 1$ ($i = 1, 2, \ldots, N$), or to the 1-Limited, if $a_i = 0$.

Our breakdown model differs from the Bernoulli Schedule in three aspects: (i) there is no guarantee that the first customer will be *successfully* served; (ii) a breakdown is observed only at the *end* of a service; and (iii) the service of the job 'suffering' a breakdown will have to be repeated.

In addition to the node-dependent failure process we further assume that due to an overall general 'noise' in the system at large, whenever the server switches to station $i$, the station is found to be "*up*" with probability $p_i$, and "*down*" with probability $(1 - p_i)$. These probabilities are independent of whether the station was up or down when last visited. If the server finds that the station is down, it immediately switches on to the next node.

As mentioned, the present work also differs from the models studied in [7] or [11] in that it considers a polling system with breakdowns under the *Globally Gated* discipline. This service regime was introduced by Boxma, Levy and Yechiali [4], and further studied by Boxma, Weststrate and Yechiali [5], Altman, Khamisy and Yechiali [1], and Armony and Yechiali [3]. Under this regime, all stations' gates are 'closed' simultaneously at every new beginning of a cycle, and the server, upon visiting a station, attempts to serve only those customers that were present there at the *start* of the cycle. Thus, while under the Bernoulli Schedule the number of services in any given queue is (probabilistically) unconstrained, under the Globally Gated regime the number of service attempts is bounded by the number of customers present in that queue at the *beginning* of the cycle.

The paper is organized as follows. After describing the assumptions and introducing the notation, we define in section 2 the evolution equations of the system-state at polling instants and derive the Laplace-Stieltjes transform (LST) and mean length of the busy period at each queue. In section 3 we obtain a set of equations for the Probability Generating Functions (PGFs) of the system-state at polling instants. The PGF of the system-state when polling station $(i+1)$ is given in terms of the PGF when polling station $i$ ($i = 1, 2, \ldots, N$). In section 4 we derive a set of equations for the calculation of the mean number of jobs at the various queues at polling instants. In section 5 we calculate the mean cycle time, while in section 6 we derive the PGF of the system-state at an *arbitrary* moment, which leads to an expression for the LST of the waiting time of a customer in any given queue $i$. We will use the following notation and assumptions:

- $N$ = number of stations (channels, nodes, queues), indexed $i = 1, 2, \ldots, N$.
- $A^i(t)$ = number of customers (jobs, messages) arriving to node $i$ during a time interval of length $t$. $A^i(t)$ is a Poisson process with rate $\lambda_i$.
- $T_i$ = time to breakdown of station $i$. $T_i$ is Exponentially distributed with parameter $\gamma_i$. This station-dependent process is regenerated at the beginning of every new

visit of the server.

- $(B_{im})_{m=1}^{\infty}$ = sequence of i.i.d. service times at queue $i$, all distributed like $B_i$ with mean $b_i$.

- $D_i$ = switch-over time from queue $i$ to queue $i+1$, with $E[D_i] = d_i$.

- $S^*(w) \equiv E[e^{-wS}]$ = LST of a non-negative continuous random variable $S$.

- The server visits the stations in a *cyclic* order, i.e.
  $1 \to 2 \to \cdots \to i \to (i+1) \to \cdots \to (N-1) \to N \to 1 \to 2 \to \cdots$

- 'Polling instant' of node $i$ is the moment when the server enters station $i$, after completion of a switch-over period, $D_{i-1}$, from station $i-1$ to station $i$.

- $X_i^j$ = number of jobs at queue $j$ $(j = 1, 2, \ldots, N)$ when node $i$ is polled.

- $\underline{X}_i \equiv (X_i^1, X_i^2, \ldots, X_i^j, \ldots, X_i^N)$ = system-state vector at polling instant of node $i$.

- $\widehat{X}(z) \equiv E[z^X]$ = PGF of a non-negative, integer-valued, random variable $X$.

- $Y_i(k)$ = number of customers that are *successfully* served in queue $i$ during a visit of the server, given that node $i$ is up when polled and $k$ customers were present in queue $i$ at the beginning of the cycle. Clearly, $Y_i(0) = 0$, but for $k \geq 1$, $Y_i(k) = k$ only if station $i$ does *not* break down during the visit of the server.

- $\widehat{Y}_i(k) = \min(Y_i(k) + 1, k)$. $\widehat{Y}_i(k)$ is the number of service attempts if node $i$ is up when polled and there were $k$ customers present in queue $i$ at the start of the cycle.

## 2. THE GLOBALLY GATED REGIME

Let $\underline{Q} = (Q^1, Q^2, \ldots, Q^j, \ldots, Q^N)$ denote the state-vector of the number of jobs at the various queues at the *start* of a cycle. Assuming (without loss of generality) that cycles start at station 1, it is clear that $Q^j = X_1^j$, $j = 1, 2, \ldots, N$. Nevertheless, to avoid possible ambiguity, we shall retain the separate notation $Q^j$.

As described in the Introduction, under the Globally Gated service regime, at each cycle the server attempts to serve at queue $j$ only those $Q^j$ jobs present there at the *start* of the cycle. Thus for any given $\underline{Q}$, the cycle's duration is (stochastically) determined, and is given by

$$C = \sum_{i=1}^{N} \left[ \left( \sum_{m=1}^{\widehat{Y}_i(Q_i)} B_{im} \right) \cdot 1_{\{\text{station } i \text{ is up when polled}\}} + D_i \right] .$$

### 2.1. Evolution Equations

The evolution equations for the system state $\underline{X}$ at polling instants depends on whether a node is *up* or is *down* when polled.

If node $i$ is *down* when polled, the server immediately moves to the next station, and hence

$$X_{i+1}^j \stackrel{d}{=} X_i^j + A^j(D_i) \qquad i, j = 1, 2, \ldots, N \; . \tag{2.1}$$

If node $i$ is *up* when polled, then

$$X_{i+1}^{j} \overset{d}{=} \begin{cases} X_i^j + A^j\Big(\sum\limits_{m=1}^{\widehat{Y}_i(Q^i)} B_{im}\Big) + A^j(D_i), & j \neq i \\ X_i^i + A^i\Big(\sum\limits_{m=1}^{\widehat{Y}_i(Q^i)} B_{im}\Big) + A^i(D_i) - Y_i(Q^i), & j = i \end{cases} \tag{2.2}$$

where, if $i = N$, by "station $N+1$" we mean "station 1".

Let $\theta_{Q^i}^i$ denote the length of a busy period in station $i$, *starting* with $Q^i$ jobs. Clearly,

$$\theta_{Q^i}^i = \sum_{m=1}^{\widehat{Y}_i(Q^i)} B_{im} \ .$$

As seen from Equation (2.2), $\theta_{Q^i}^i$ plays an important role in the analysis. We turn to derive its distribution and mean.

### 2.2. The Busy Period $\theta_{Q^i}^i$

For fixed $r$, $r = 0, 1, 2 \ldots$ let $\theta_r^i$ be the duration of a busy period in station $i$, *initiated* with $r$ jobs.

Let $\sigma_k^i \equiv \sum\limits_{i=1}^{k} B_{im}$ be the duration of $k$ consecutive services in station $i$, $k = 1, 2, 3 \ldots$, and set $\sigma_0^i = 0$. Define

$$\Phi_r^i(w) = E\Big[e^{-\omega\theta_r^i}\Big], \quad \Phi_0^i(\omega) = 1 \ . \tag{2.3}$$

Suppressing $i$, we have $\theta_r = \sum\limits_{k=1}^{r} (\sigma_k) 1_{(\sigma_{k-1}, \sigma_k]}(T) + (\sigma_r) 1_{(\sigma_r, \infty)}(T)$

where, $1_{(a,b]}(T) = \begin{cases} 1, & \text{if } a < T \leq b \\ 0, & \text{otherwise} \end{cases}$. Now,

$$\begin{aligned} &E\Big[e^{-\omega\sigma_k} 1_{(\sigma_{k-1},\sigma_k]}(T)\Big] = E\Big[e^{-\omega\sigma_k} P(\sigma_{k-1} < T \leq \sigma_k | B_1, \ldots, B_k)\Big] \\ &\quad = E\Big[e^{-\omega\sigma_k}(e^{-\gamma\sigma_{k-1}} - e^{-\gamma\sigma_k})\Big] = E\Big[e^{-(\omega+\gamma)\sigma_{k-1}} \cdot e^{-\omega B_k} - e^{-(\omega+\gamma)\sigma_k}\Big] \\ &\quad = \sigma_{k-1}^*(\omega+\gamma) B^*(\omega) - \sigma_k^*(\omega+\gamma) = (B^*(\omega+\gamma))^{k-1}[B^*(\omega) - B^*(\omega+\gamma)] \ . \end{aligned}$$

We further have,

$$\begin{aligned} &E\Big[e^{-\omega\sigma_r} 1_{(\sigma_r,\infty)}(T)\Big] = E\Big[e^{-\omega\sigma_r} \cdot P(T > \sigma_r | \sigma_r)\Big] \\ &\quad = E\Big[e^{-\omega\sigma_r} e^{-\gamma\sigma_r}\Big] = \sigma_r^*(\omega+\gamma) = (B^*(\omega+\gamma))^r \ . \end{aligned}$$

Thus,

$$\begin{aligned} \Phi_r(\omega) &= \sum_{k=1}^{r} [B^*(\omega) - B^*(\omega+\gamma)](B^*(\omega+\gamma))^{k-1} + (B^*(w+\gamma))^r \\ &= [B^*(\omega) - B^*(\omega+\gamma)] \left(\tfrac{1-(B^*(\omega+\gamma))^r}{1-B^*(\omega+\gamma)}\right) + (B^*(\omega+\gamma))^r \ . \end{aligned} \tag{2.4}$$

To summarize,

$$\Phi_r(\omega) = \varphi_1(\omega) + (1 - \varphi_1(\omega))(B^*(\omega+\gamma))^r \tag{2.5}$$

where

$$\varphi_1(\omega) = \frac{B^*(\omega) - B^*(\omega+\gamma)}{1 - B^*(\omega+\gamma)} \ . \tag{2.6}$$

Finally, taking expectation over $Q^i$, we obtain

$E[\Phi^i_{Q^i}(\omega)] = \varphi^i_1(\omega) + (1 - \varphi^i_1(\omega))\widehat{Q}^i(B^*_i(\omega+\gamma))$

where $\varphi^i_1(\omega)$ is obtained from (2.5) by inserting back the index $i$ for $B^*(\cdot)$. Now,

$\varphi^i_1(0) = 1; \quad \frac{d}{d\omega}\varphi^i_1(\omega)|_{\omega=0} = \frac{-b_i}{1-B^*_i(\gamma_i)}$ . Thus,

$$E[\theta^i_r] = -\frac{d}{d\omega}\Phi^i_r(\omega)|_{\omega=0} = \frac{b_i}{1-B^*_i(\gamma_i)}\Big[(1-B^*_i(\gamma_i))^r\Big] \ . \tag{2.7}$$

Finally,

$$E[\theta^i_{Q^i}] = \frac{b_i}{1-B^*_i(\gamma_i)}\Big[1 - \widehat{Q}_i(B^*_i(\gamma_i))\Big] \ . \tag{2.8}$$

It is just left to calculate the term $\widehat{Q}_i(B^*_i(\gamma_i))$. This will be done in section 4, in conjunction with the calculation of the first-order moments of the $X^j_i$s.

### 2.3. The distribution of a successful and unsuccessful service attempt

Suppressing $i$, let $S$ be the duration of a *successful* service attempt, i.e. the duration of a service that has been completed before $T$. Then $S \sim B\Big|_{B<T}$. Now,

$$\begin{aligned} E\Big[e^{-\omega S}\Big] \equiv S^*(\omega) &= E\Big[e^{-\omega B} \mid B < T\Big] = \frac{E\Big[e^{-\omega B}\cdot 1_{[B<T]}\Big]}{P(B<T)} \\ &= \frac{E\Big[e^{-\omega B}\cdot P(B<T|B)\Big]}{E[P(B<T \mid B)]} = \frac{E\Big[e^{-\omega B}\cdot e^{-\gamma B}\Big]}{E[e^{-\gamma B}]} = \frac{B^*(\omega+\gamma)}{B^*(\gamma)} \ . \quad \text{Thus,} \end{aligned}$$

$$S^*_i(w) = \frac{B^*_i(\omega+\gamma_i)}{B^*_i(\gamma_i)} \tag{2.9}$$

$$E[S_i] = -\frac{B^{*\prime}_i(\gamma_i)}{(B^*_i(\gamma_i))} = \frac{E\Big[Be^{-\gamma_i B_i}\Big]}{B^*_i(\gamma_i)} \ . \tag{2.10}$$

Let $\widehat{S} \sim B\Big|_{B\geq T}$ be the duration of an *unsuccessful* service attempt. Then

$$\begin{aligned} E\Big[e^{-\omega\widehat{S}}\Big] \equiv \widehat{S}^*(\omega) &= E\Big[e^{-\omega B} \mid B \geq T] = \frac{E\Big[e^{-\omega B}\cdot 1_{[B\geq T]}\Big]}{P(B\geq T)} \\ &= \frac{E\Big[e^{-\omega B}P(B\geq T|B)\Big]}{E[P(B\geq T|B]} = \frac{E\Big[e^{-\omega B}(1-e^{-\gamma B})\Big]}{E[1-e^{-\gamma B}]} \ . \quad \text{Thus,} \end{aligned}$$

$$\widehat{S}^*_i(\omega) = \frac{B^*_i(\omega) - B^*_i(\omega+\gamma_i)}{1-B^*_i(\gamma_i)} \ , \tag{2.11}$$

$$E[\widehat{S}_i] = \frac{b_i + B^{*\prime}_i(\gamma_i)}{1-B^*_i(\gamma_i)} \ . \tag{2.12}$$

Using (2.10) and (2.12), it is easy to check that

$$E[B_i] = E[S_i]P[B_i < T_i] + E[\widehat{S}_i]P[B_i \geq T_i] = b_i \ .$$

### 2.4. An alternative derivation of $E[\theta^i_{Q^i}]$

With the above expressions for the distributions and means of $S_i$ and $\widehat{S}_i$ one can derive $E[\theta^i_{Q^i}]$ 'directly' as follows. The probability of a successful service in station $i$ is

$$a_i \equiv P[T_i > B_i] = E[P[T_i > B_i|B_i]] = E\left[e^{-\gamma_i B_i}\right] = B_i^*(\gamma_i) \ .$$

Because of the memoryless property of the breakdown process, the probability distribution of $Y_i(Q^i)$, the number of successful service completions during a visit of the server in station $i$, is given by

$$P[Y_i(Q^i) = \ell] = \begin{cases} a_i^\ell(1-a_i) \ , & \ell = 0,1,2,\ldots,Q^i-1 \\ a_i^{Q^i} \ , & \ell = Q^i \end{cases} \tag{2.13}$$

Clearly, if $Q^i = 0$ then $P[Y_i(0) = 0] = P[\widehat{Y}_i(0) = 0] = 1$.

Now, if $Y_i(Q^i) = \ell < Q^i$, there were exactly $\ell$ *successful* services and *one* unsuccessful service attempt, and if $Y_i(Q^i) = Q^i$, then *all* service attempts were successful. Hence,

$$\begin{aligned} E\left[\theta^i_{Q^i}|Q^i\right] &= \sum_{\ell=0}^{Q^i-1}\left[\ell\cdot E[S_i] + E[\widehat{S}_i]\right]a_i^\ell(1-a_i) + Q^iE[S_i]a_i^{Q^i} \\ &= E[S_i]\left[\frac{a_i(1-a_i^{Q^i-1})}{1-a_i} - (Q^i-1)a_i^{Q^i}\right] + E[\widehat{S}_i](1-a_i^{Q^i}) + Q_i^iE[S_i]a_i^{Q^i} \\ &= E[S_i]\left(\frac{a_i(1-a_i^{Q^i})}{1-a_i}\right) + E[\widehat{S}_i](1-a_i^{Q^i}) \ . \end{aligned}$$

Substituting the values of $E[S_i]$ and $E[\widehat{S}_i]$ from (2.10) and (2.12), we get

$$E\left[\theta^i_{Q^i}|Q^i\right] = \frac{-B_i^{*\prime}(\gamma_i)}{a_i}\left(\frac{a_i - a_i^{Q^i+1}}{1-a_i}\right) + \frac{b_i + B^{*\prime}(\gamma_i)}{1-a_i}(1-a_i^{Q^i}) = \frac{b_i}{1-a_i}\left[1-a_i^{Q^i}\right] \ .$$

It follows that $E[\theta^i_{Q^i}] = \frac{b_i}{1-a_i}\left[1-\widehat{Q}^i(B_i^*(\gamma_i))\right]$, as given by Equation (2.8).

Applying (2.13) we get

$$\begin{aligned} E\left[Y_i(Q^i)\right] &= E\left[\sum_{\ell=0}^{Q^i-1}\ell a_i^\ell(1-a_i) + Q^ia_i^{Q^i}\right] \\ &= E\left[\frac{a_i}{1-a_i}\left(1-a_i^{Q^i}\right)\right] = \frac{a_i}{1-a_i}\left[1-\widehat{Q}^i(a_i)\right] \ . \end{aligned} \tag{2.14}$$

We also wish to calculate $E[\widehat{Y}_i(Q^i)]$, where $\widehat{Y}_i(Q^i)$ is the total number of service attempts during a visit of station $i$ (before the server leaves, either because all $Q^i$ jobs are successfully completed, or the station breaks down).

$$\begin{aligned} E\left[\widehat{Y}_i(Q^i)\right] &= E\left[\sum_{\ell=0}^{Q^i-1}(\ell+1)a_i^\ell(1-a_i) + Q^ia_i^{Q^i}\right] \\ &= E\left[\frac{1}{1-a_i}\left(1-a_i^{Q^i}\right)\right] = \frac{1}{1-a_i}\left[1-\widehat{Q}^i(a_i)\right] \ . \end{aligned} \tag{2.15}$$

Indeed, comparing with (2.8), $E[\theta^i_{Q^i}] = b_i E[\widehat{Y}_i(Q^i)]$ .

## 3. GENERATING FUNCTIONS AT POLLING INSTANTS

We extend the system's state vector to include both $\underline{Q}$ and $\underline{X}_i$ such that, at a polling instant of station $i$, the state is given by

$(\underline{Q};\underline{X}_i) = (Q^1, Q^2, \ldots, Q^N; X^1_i, X^2_i, \ldots, X^N_i)$ .

The Probability Generation Function (PGF) at that instant is given (independently whether the station is up or down) by

$$G_i(\underline{\omega};\underline{z}) = E\left[\prod_{j=1}^{N} \omega_j^{Q^j} \prod_{j=1}^{N} z_j^{X^j_i}\right] . \tag{3.1}$$

Observe that the PGF of the number of arrivals to the various queues during the switching time $D_i$ from station $i$ to station $i+1$ is given by $E\left[\prod_{j=1}^{N} z_j^{A^j(D_i)}\right] = D^*_i(\Sigma)$ where $\Sigma = \sum\limits_{j=1}^{N} \lambda_j(1-z_j)$. Thus, by substituting the evolution equations (2.1) and (2.2) into (3.1), we express $G_{i+1}(\cdot)$ in terms of $G_i(\cdot)$ as follows.

For $i = 1, 2, \ldots, N-1$

$$\begin{aligned} G_{i+1}(\underline{\omega};\underline{z}) &= D^*_i(\Sigma)\Bigg\{(1-p_i)G_i(\underline{\omega};\underline{z}) \\ &\quad + p_i E\left[\left(\prod_{j=1}^{N} \omega_j^{Q^j}\right)\left(\prod_{j=1}^{N} z_j^{X^j_i}\right)\left(\prod_{j=1}^{N} z_j^{A^j(\theta^i_{Q^i})}\right) z_i^{-Y_i(Q^i)}\right]\Bigg\} \\ &= D^*_i(\Sigma)\Bigg\{(1-p_i)G_i(\underline{\omega};\underline{z}) \qquad\qquad (3.2) \\ &\quad + p_i E\left[\left(\prod_{j=1}^{N} \omega_j^{Q^j}\right)\left(\prod_{j=1}^{N} z_j^{X^j_i}\right) E\left[\left(\prod_{j=1}^{N} z_j^{A^j(\theta^i_{Q^i})}\right) z_i^{-Y_i(Q^i)} \Big| \underline{Q}, \underline{X}_i\right]\right]\Bigg\} . \end{aligned}$$

Suppressing the index $Q^i$ from $\theta^i_{Q^i}$ we write

$$\begin{aligned} K &\equiv E\left[\left(\prod_{j=1}^{N} z_j^{A^j(\theta^i)}\right) z_i^{-Y_i(Q^i)} \Big| \underline{Q}\right] = E\left[E\left[\left(\prod_{j=1}^{N} z_j^{A^j(\theta^i)}\right) \Big| \widehat{Y}_i\right] z_i^{-Y_i(Q^i)} \Big| \underline{Q}\right] \\ &= E\left[S^*_i(\Sigma)^{Y_i} \widehat{S}^*_i(\Sigma) z_i^{-Y_i} \cdot 1_{\{Y_i<Q^i\}} \Big| \underline{Q}\right] + E\left[S^*_i(\Sigma)^{Y_i} z_i^{-Y_i} \cdot 1_{\{Y_i=Q^i\}} \Big| \underline{Q}\right] \equiv K_1 + K_2 \end{aligned}$$

where $S^*_i(\cdot)$ and $\widehat{S}^*_i(\cdot)$ are given by (2.9) and (2.11), respectively.

$$\begin{aligned} K_1 &= \sum_{\ell=0}^{Q^i-1} S^*_i(\Sigma)^{\ell} \widehat{S}^*_i(\Sigma) z_i^{-\ell} a_i^{\ell}(1-a_i) = \widehat{S}^*_i(\Sigma)(1-a_i)\frac{1-(a_i S^*_i(\Sigma)/z_i)^{Q^i}}{1-a_i S^*_i(\Sigma)/z_i} \\ &= (B^*_i(\Sigma) - B^*_i(\Sigma+\gamma_i))\frac{1-[B^*_i(\Sigma+\gamma_i)z_i^{-1}]^{Q^i}}{1-B^*_i(\Sigma+\gamma_i)z_i^{-1}} \\ K_2 &= S^*_i(\Sigma)^{Q^i} z_i^{-Q^i} a_i^{Q^i} = \left[B^*_i(\Sigma+\gamma_i)z_i^{-1}\right]^{Q^i} . \end{aligned}$$

Substituting $K_1$ and $K_2$ into (3.2) yields

$$\begin{aligned} G_{i+1}(\underline{\omega};\underline{z}) &= D_i^*(\Sigma)\Bigg\{(1-p_i)G_i(\underline{\omega};\underline{z}) \\ &+p_i\Bigg[E\Bigg[\left(\prod_{j=1}^{N}\omega_j^{Q^j}\right)\left(\prod_{j=1}^{N}z_j^{X_i^j}\right)\Bigg[(B_i^*(\Sigma)-B_i^*(\Sigma+\gamma_i))\left(\frac{1-[B_i^*(\Sigma+\gamma_i)z_i^{-1}]^{Q^i}}{1-B_i^*(\Sigma+\gamma_i)z_i^{-1}}\right) \\ &+[B_i^*(\Sigma+\gamma_i)z_i^{-1}]^{Q^i}\Bigg]\Bigg]\Bigg]\Bigg\} \ . \quad \text{Thus,} \end{aligned}$$

$$\begin{aligned} G_{i+1}(\underline{\omega};\underline{z}) &= D_i^*(\Sigma)\Bigg\{(1-p_i)G_i(\underline{\omega};\underline{z})+p_i\Bigg[\eta_i(\Sigma,\gamma_i)G_i(\underline{\omega};\underline{z}) \\ &+(1-\eta_i(\Sigma,\gamma_i))G_i\left(\omega_1,\omega_2,\ldots,\omega_{i-1},B_i^*(\Sigma+\gamma_i)\frac{\omega_i}{z_i},\omega_{i+1},\ldots,\omega_N;\underline{z}\right)\Bigg]\Bigg\} \end{aligned} \tag{3.3}$$

where

$$\eta_i(\Sigma,\gamma_i)=\frac{B_i^*(\Sigma)-B_i^*(\Sigma+\gamma_i)}{1-B_i^*(\Sigma+\gamma_i)z_i^{-1}} \quad (i=1,2,\ldots,N-1)\ . \tag{3.4}$$

For $i=N$,

$$\begin{aligned} G_1(\underline{\omega};\underline{z}) &= D_N^*(\Sigma)\Bigg\{(1-p_N)G_N(\underline{1};\omega_1z_1,\omega_2z_2,\ldots,\omega_Nz_N) \\ &+p_NE\left[\left(\prod_{j=1}^{N}(\omega_jz_j)^{X_N^j}\right)\left(\prod_{j=1}^{N}(\omega_jz_j)^{A^j(\theta^N)}\right)\cdot(\omega_Nz_N)^{-Y_N(Q^N)}\right]\Bigg\} \ . \end{aligned}$$

Applying again the results for $K_1$ and $K_2$ (with $i=N$) and setting $\Sigma^+=\sum_{j=1}^{N}\lambda_j(1-\omega_jz_j)$, we obtain

$$\begin{aligned} G_1(\underline{\omega};\underline{z}) &= D_N^*(\Sigma)\Bigg\{(1-p_N)G_N(\underline{1};\omega_1z_1,\omega_2z_2,\ldots,\omega_Nz_N) \\ &+p_N\Bigg[\eta_N^+(\Sigma^+,\gamma_N)G_N(\underline{1};\omega_1z_1,\omega_2z_2,\ldots,\omega_Nz_N) \\ &+(1-\eta_N^+(\Sigma^+,\gamma_N))\cdot G_N\left(1,1,\ldots,1_{N-1},\frac{B_N^*(\Sigma^++\gamma_N)}{\omega_Nz_N};\omega_1z_1,\omega_2z_2,\ldots,\omega_Nz_N\right)\Bigg]\Bigg\} \end{aligned} \tag{3.5}$$

where

$$\eta_N^+(\Sigma^+,\gamma_N)=\frac{B_N^*(\Sigma^+)-B_N^*(\Sigma^++\gamma_N)}{1-B_N^*(\Sigma^++\gamma_N)/(\omega_Nz_N)} \ . \tag{3.6}$$

In principle, by taking derivatives of (3.3) and (3.5), one can obtain sets of equations from which various-order moments of $X_i^j$ may be calculated. We turn to obtain 'directly' such a set.

## 4. FIRST-ORDER MOMENTS

We now derive a set of equations for the calculation of the mean number of jobs at the various queues at polling instants.

Define $f_i^j = E[X_i^j]$. Then, for any fixed $i$ ($i = 1, 2, \ldots, N$), taking expectations of both sides of (2.1) and (2.2) leads to

$$f_{i+1}^j = f_i^j + \lambda_j d_i + p_i \begin{cases} \lambda_j b_i E[\widehat{Y}_i(Q^i)] \ , & i \neq j \\ \left\{\lambda_i b_i E[\widehat{Y}_i(Q^i)] - E[Y_i(Q^i)]\right\} \ , & i = j \end{cases} \tag{4.1}$$

Substituting results (2.15) and (2.14) in (4.1) yields.

$$f_{i+1}^j = f_i^j + \lambda_j d_i + p_i \begin{cases} \frac{\lambda_j b_i}{1-a_i}[1 - \widehat{Q}^i(a_i)] \ , & i \neq j \\ \frac{\lambda_i b_i - a_i}{1-a_i}[1 - \widehat{Q}^i(a_i)] \ , & i = j \end{cases} \tag{4.2}$$

**Remark.** If $\gamma_i = 0$, i.e. $E[T_i] = 1/\gamma_i = \infty$, then $a_i = B^*(\gamma_i) = 1$ and $\widehat{Q}_i^i(a_i) = 1$. Thus, by using L'Hospital rule,

$\lim_{a_i \to 1} \frac{1-\widehat{Q}^i(a_i)}{1-a_i} = \frac{d}{da_i}\widehat{Q}^i(a_i)\Big|_{a_i=1} = E[Q^i] = f_1^i$ .

Hence, if $p_i = 1$, Equation (4.2) reduces to

$$f_{i+1}^j = \begin{cases} f_i^j + \lambda_j d_i + \lambda_j b_i f_1^i, & j \neq i \\ f_i^i + \lambda_i d_i + \lambda_i b_i f_1^i - f_1^i, & j = i \end{cases}$$

Next we turn to find the unknown quantity $\widehat{Q}^i(a_i)$ appearing in (4.2).

Let $x_i \equiv \frac{p_i b_i}{1-a_i}[1 - \widehat{Q}^i(a_i)] = p_i b_i E[\widehat{Y}_i(Q^i)]$ (see (2.15)), and $\alpha_i \equiv \frac{a_i}{\lambda_i b_i}$. That is, $x_i$ is the expected duration of a visit of the server in station $i$. Dividing by $\lambda_j$, Equation (4.2) can be written as

$$f_{i+1}^j/\lambda_j = f_i^j/\lambda_j + d_i + \begin{cases} x_i & j \neq i \\ (1-\alpha_j)x_j & j = i \end{cases} \tag{4.3}$$

Now, for every $j$, by summing (4.3) over all $i$, we obtain

$$0 = d + \sum_{\substack{i=1 \\ i \neq j}}^{N} x_i + (1-\alpha_j)x_j \quad j = 1, 2, \ldots, N \tag{4.4}$$

or

$$A\underline{x} = d \cdot \underline{1} \tag{4.5}$$

where $d = \sum_{j=1}^{N} d_j$ and $A$ is an $N \times N$ matrix

$$A = \begin{bmatrix} (\alpha_1 - 1) & -1 & -1 & \cdots & -1 \\ -1 & (\alpha_2 - 1) & -1 & \cdots & -1 \\ -1 & -1 & (\alpha_3 - 1) & \cdots & -1 \\ & & & & \\ -1 & -1 & -1 & & (\alpha_N - 1) \end{bmatrix}$$

while $\underline{1}$ and $\underline{x}$ are each an $N$-dimensional vector: $\underline{x}^T = [x_1, x_2, \ldots, x_N]$, $\underline{1}^T = [1, 1, \ldots, 1]$. Since $a_i < 1$, $x_i > 0$. It follows that a necessary condition for stability of station $i$ is $\alpha_i - 1 > 0$, i.e. $a_i > \lambda_i b_i$. Indeed, since $a_i = B_i^*(\gamma_i)$ is the probability of a successful service attempt at node $i$, the expected number of attempts until service completion is $1/a_i$. Thus, the mean amount of work 'flowing' to node $i$ per unit time is $\lambda_i b_i / a_i$. This quantity should not exceed 1.

It is clear that with $\alpha_i > 1$ for every $i = 1, 2, \ldots, N$, Equation (4.5) has a *unique* solution $\underline{x}^0$. Hence, the unknown quantity $\widehat{Q}^i(a_i)$ in Equation (4.2) is given by

$$\widehat{Q}^i(a_i) = 1 - \frac{1 - B_i^*(\gamma_i)}{p_i b_i} x_i^0 \ . \tag{4.6}$$

## 5. CYCLE TIME

The Cycle duration, $C$, is given by

$$C = \sum_{i=1}^{N} D_i + \sum_{i=0}^{N} p_i E\Big[\widehat{Y}_i(Q^i)\Big] b_i \ . \tag{5.1}$$

Using (2.15), the mean cycle time is

$$E[C] = d + \sum_{i=1}^{N} \frac{p_i b_i}{1 - a_i} \Big[1 - \widehat{Q}^i(a_i)\Big] = d + \sum_{i=1}^{N} x_i^0 \ . \tag{5.2}$$

It follows from (4.4) that, for every $j$,

$$\alpha_j x_j^0 = d + \sum_{i=1}^{N} x_i^0 = E[C] \ . \tag{5.3}$$

Indeed, the mean number of arrivals to node $j$ during a cycle is $\lambda_j E[C]$, each requiring $b_j/a_j$ amount of work. Thus, the mean visit time in node $j$ is $x_j^0 = \lambda_j E[C] \cdot \frac{b_j}{a_j} = E[C]/\alpha_j$. Alternatively, from (2.14), the mean number of jobs served per cycle in station $i$ is

$$p_i E\Big[Y_i(Q^i)\Big] = \frac{a_i}{b_i} x_i^0 = \lambda_i \alpha_i x_i^0 = \lambda_i E[C] \ . \tag{5.4}$$

Equation (5.4) simply says that the average outflow of jobs from a node, per cycle, equals the average inflow of jobs to that node. Now, from (5.3),

(i) $\quad x_j^0 = E[C]/\alpha_j = E[C] \frac{\lambda_j b_j}{a_j}$, $\quad$ (ii) $\quad \sum_{j=1}^{N} x_j^0 = E[C] - d$.

Hence, $\sum_{j=1}^{N} x_j^0 = E[C] \sum_{j=1}^{N} \frac{\lambda_j b_j}{a_j} = E[C] - d$, implying that

$$E[C] = \frac{d}{1 - \sum_{j=1}^{N} \lambda_j (\frac{b_j}{a_j})} \equiv \frac{d}{1 - \rho} \tag{5.5}$$

where $\rho \equiv \sum_{j=1}^{N} \rho_j = \sum_{j=1}^{N} \lambda_j \left(\frac{b_j}{a_j}\right)$ is the total inflow rate of work to the system. It is clear that a necessary condition for system's stability is $\rho = \sum_{j=1}^{N} \lambda_j (b_j/a_j) < 1$.

## 6. PGF OF THE SYSTEM STATE AT AN ARBITRARY MOMENT

Let $F(\underline{z}) = E\Big[\prod_{j=1}^{N} z_j^{X_j}\Big]$ be the PGF of the system state $\underline{X} = (X_1, X_2, \ldots, X_N)$ at an *arbitrary* moment. Then, as in Eisenberg [6], Altman and Yechiali [2],

$$F(\underline{z}) = \frac{1}{E[C]}\Big\{\sum_{i=1}^{N}\Big[g_i^i b_i F\Big(\underline{z}\Big|\begin{matrix}\text{service}\\ \text{period } i\end{matrix}\Big) + d_i F\Big(\underline{z}\Big|\begin{matrix}\text{switch} - \text{over}\\ \text{period } i\end{matrix}\Big)\Big]\Big\} \tag{6.1}$$

where $g_i^i = p_i[E\widehat{Y}_i(Q^i)]$ is the mean number of service beginnings in node $i$ per cycle. Let $\overline{G}_i(\underline{z})$ be the PGF of the system-state at the *end* of a visit to node $i$. Then,

$$G_{i+1}(\underline{1};\underline{z}) = \overline{G}_i(\underline{z}) D_i^*(\Sigma)\ . \tag{6.2}$$

Let $V_i(\underline{z})$ ($\overline{V}_i(\underline{z})$) be the PGF of the system-state at a customer's service-beginning (service-endings) instant in queue $i$. Then

$$\begin{aligned}\overline{V}_i(\underline{z}) &= a_i V_i(\underline{z})\frac{B_i^*(\Sigma)}{z_i} + (1-a_i)V_i(\underline{z})B_i^*(\Sigma)\\ &= V_i(\underline{z})B_i^*(\Sigma)\Big[\frac{a_i}{z_i} + (1-a_i)\Big] \qquad (6.3)\\ F\Big(\underline{z}\Big|\begin{matrix}\text{service}\\ \text{period } i\end{matrix}\Big) &= V_i(\underline{z})\frac{1-B_i^*(\Sigma)}{b_i\cdot\Sigma} \qquad (6.4)\end{aligned}$$

where $\frac{1-B_i^*(\Sigma)}{b_i\Sigma}$ is the PGF of the number of arrivals to the various stations during the time elapsed from the start of the current service. Similarly,

$$F\Big(\underline{z}\Big|\begin{matrix}\text{switch} - \text{over}\\ \text{period } i\end{matrix}\Big) = \overline{G}_i(\underline{z})\frac{1-D_i^*(\Sigma)}{d_i\cdot\Sigma}\ . \tag{6.5}$$

All the above PGFs are related through Eisenberg's [6] observation that the union of visit beginnings and all service endings in node $i$ coincides with the union of all service beginnings and visit endings there. This observation leads to

$$G_i(\underline{1};\underline{z}) + g_i^i\overline{V}_i(\underline{z}) = g_i^i V_i(\underline{z}) + \overline{G}_i(\underline{z})\ . \tag{6.6}$$

Combining with (6.2) and (6.3), equation (6.6) is written as

$$V_i(\underline{z}) = \frac{\big[G_i(\underline{1};\underline{z})D_i^*(\Sigma) - G_{i+1}(\underline{1};\underline{z})\big]}{D_i^*(\Sigma)g_i^i\Big[1 - B_i^*(\Sigma)\Big(\frac{a_i}{z_i} + (1-a_i)\Big)\Big]} \tag{6.7}$$

Using (6.4), (6.5) and the above, we finally have for (6.1)

$$\begin{aligned}F(\underline{z}) &= \frac{1}{E[C]}\Big[\sum_{i=1}^{N}\Big[g_i^i b_i V_i(\underline{z})\cdot\frac{1-B_i^*(\Sigma)}{b_i\cdot\Sigma} + d_i\frac{G_{i+1}(\underline{1};\underline{z})}{D_i^*(\Sigma)}\cdot\frac{1-D_i^*(\Sigma)}{d_i\cdot\Sigma}\Big]\Big]\\ &= \frac{1}{E[C]}\Big[\sum_{i=1}^{N}\Big[\frac{[G_i(\underline{1};\underline{z})D_i^*(\Sigma) - G_{i+1}(\underline{1};\underline{z})]}{D_i^*(\Sigma)\Big[1 - B_i^*(\Sigma)\Big(\frac{a_i}{z_i} + (1-a_i)\Big)\Big]}\cdot\frac{1-B_i^*(\Sigma)}{\Sigma}\\ &\qquad + \frac{G_{i+1}(\underline{1};\underline{z})}{D_i^*(\Sigma)}\cdot\frac{1-D_i^*(\Sigma)}{\Sigma}\Big]\Big]\ . \qquad (6.8)\end{aligned}$$

The relation between $G_{i+1}(\underline{1};\underline{z})$ and $G_i(\underline{1};\underline{z})$ is expressed in Eq. (3.3).

**Waiting Times**

The distribution of waiting times in each queue is obtained in the classical way. Let $W_i$ be the waiting time of an arbitrary job at station $i$, with $W_i^*(\cdot)$ its LST. Then, since the distribution of the number of jobs in the system is the same at arrival and departure instants, and with the use of PASTA,

$$F(1,1,\ldots,1,z_i,1,\ldots,1) = W_i^*(\lambda_i(1-z_i)) \ . \tag{6.9}$$

## REFERENCES


1. E. Altman, A. Khamisy and U. Yechiali, "On elevator polling with globally gated regime", Queueing Systems 11 (1992) 85-90.
2. E. Altman and U. Yechiali, "Polling in a closed network", Probability in the Engineering and Informational Sciences 8 (1994) 327-343.
3. R. Armony and U. Yechiali, "Polling systems with permanent and transient customers", Technical Report, School of Math. Sciences, Tel Aviv University (1996), to be published in Stochastic Models.
4. O. Boxma, H. Levy and U. Yechiali, "Cyclic reservation schemes for efficient operation of multiple-queue single-server systems", Annals of Operations Research 35 (1991) 187-208.
5. O. Boxma, J.W. Weststrate and U. Yechiali, "A globally gated polling system with server interruptions, and applications to the repairmen problem", Probability in the Engineering and Informational Sciences 7 (1993) 187-208.
6. M. Eisenberg, "Queues with periodic service and changeover time", Operations Research 20 (1972) 440-451.
7. D. Kofman and U. Yechiali, "Polling systems with station breakdowns", Performance Evaluation 27& 28 (1996) 647-672.
8. H. Takagi, Analysis of Polling Systems, MIT Press, 1986.
9. H. Takagi, "Queueing analysis of polling models: An update". In: Stochastic Analysis of Computer and Communications Systems, (H. Takagi, ed.), pp. 267-318, North Holland, 1990.
10. H. Takagi, "Queueing analysis of polling models: Progress in 1990-1993". Technical Report No. 584, Institute of Socio-Economic Planning, University of Tsukuba, Japan, 1994.
11. Tedijanto, "Exact results for the cyclic-service queue with a Bernoulli schedule", Performance Evaluation 11 (1990) 107-115.
12. U. Yechiali, "Analysis and control of polling systems". In: Performance Evaluation of Computer and Communications Systems (L. Donatiello and R. Nelson, eds.), pp. 630-650, Springer Verlag, 1993.