CP utilization in an exponential CP-Terminal system with equal think times and different job sizes

EINDHOVEN UNIVERSITY OF TECHNOLOGY

Department of Mathematics and Computing Science

Memorandum COSOR 82 - 13

CP utilization in an exponential CP-Terminal system with equal think times and different job sizes

by

J. van der Wal, Eindhoven

Eindhoven, the Netherlands

August 1982

CP UTILIZATION IN AN EXPONENTIAL CP-TERMINAL SYSTEM WITH EQUAL THINK TIMES AND DIFFERENT JOB SIZES

by

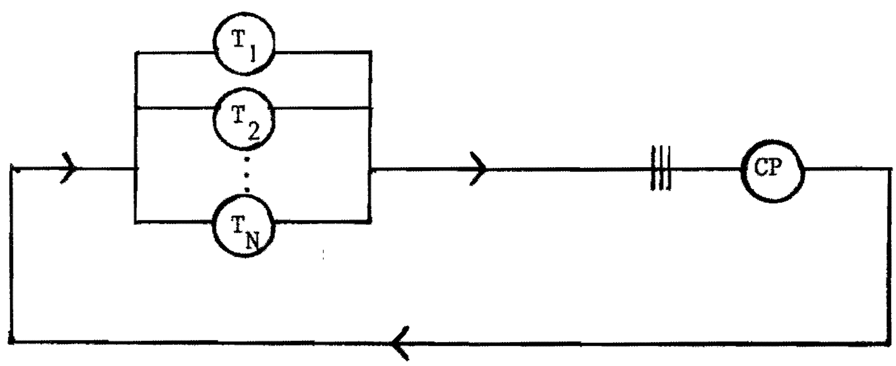
J. van der Wal, Eindhoven

Abstract.

This paper deals with the CP utilization in a CP-Terminal system with exponential job sizes and exponential think times. It is shown that if all terminals have equal think times then the CP utilization is completely independent of the scheduling of the jobs at the CP as long as preemptions are of the resume type.

1. Introduction.

We consider the following queueing system consisting of one CP (Central Processor) and N terminals, $T_1, T_2, \ldots, T_N$.



The system operates as follows. After having "thought" for an exponentially distributed amount of time with mean $1/\lambda$ a terminal produces a request for the CP and the terminal "goes to sleep". When the job has been served it starts to think about the next request. The duration of a terminal T_i job is exponentially distributed with mean $1/\mu_i$. All think times and job sizes are independent.

The question addressed in this paper is: how does, for this CP-terminal system with equal think times and different job sizes, the CP utilization factor depend on the order in which the jobs are served at the CP. At first sight one might expect some control limit type of scheduling rule, like smallest job first or largest job first, to be optimal. Somewhat surprisingly, however, it turns out that CP utilization is not at all influenced by the order of servicing at the CP.

The case of different think times where control limit polices do appear will be treated in a companion paper, see [2].

One way of attacking the problem would be to compute for each scheduling the limiting distribution to see how the probability of having all terminals thinking depends on the scheduling. We will not do this. Another line of reasoning is obtained as follows. First note that the duration of an idle period for the CP is not affected by the chosen scheduling. So the utilization factor of the CP is easily obtained from the expected duration of the busy period for the CP, and the utilization is maximized by the scheduling that maximizes the busy period duration. Our approach will be to consider the expected remaining lifetime of the busy period as function of the scheduling.
One might say that we reformulate the problem of maximizing CP utilization as a semi-Markov decision process concerning the expected remaining busy period duration. The notations introduced below are adapted from this area.

The state space of the queueing system under consideration is the set 2^N of all subsets of $I_N := \{1,2,\ldots,N\}$, i.e. state $A \subset I_N$ corresponds to the situation that all terminals T_i with $i \in A$ have delivered a job to the CP and the terminals T_j, $j \notin A$, are thinking.
A scheduling strategy f is a function on the nonempty subsets A of I_N such that $f(A) \in A$, i.e. f(A) is the index of the terminal whose request is to be served if the system is in state A. So the schedulings we allow for are those in which after each state change a job is appointed which will receive the full attention of the CP until the state of the system changes again. Further note that this type of strategy allows that the service of a job is interrupted when the state of the system changes.
We only consider preemptive resume type interruptions. For the time being we exclude service disciplines like FCFS (First Come First Served), LCFS

(Last Come First Served) and PS (Processor Sharing). In the sequel it will become clear that our approach also solves the problem for these special disciplines. Since we only consider the scheduling strategies of the type introduced above, and because of the preemptive resume assumption and the fact that all jobs and think times are exponential, the state of the queueing system immediately after a state change is completely characterized by the corresponding element in 2^N. (For the disciplines FCFS and LCFS this would not be the case).
Finally, we write $v^f(A)$ [$v^*(A)$] for the expected remaining lifetime of the busy period if the system is now in state A and strategy f [an optimal scheduling] is followed. Embedding the process on the time instants the state of the system changes we obtain a semi-Markov decision process. Form this it is immediate that the functions v^f and v^* satisfy the following recursive relations

$$(1)\quad \begin{cases} v^f(A) = \dfrac{1}{\mu_{f(A)} + \sum\limits_{j \notin A} \lambda} + \dfrac{\mu_{f(A)}}{\mu_{f(A)} + \sum\limits_{j \notin A} \lambda} v^f(A \setminus \{f(A)\}) + \\ \qquad\qquad + \sum\limits_{k \notin A} \dfrac{\lambda}{\mu_{f(A)} + \sum\limits_{j \notin A} \lambda} v^f(A \cup \{k\}) \\ v^f(\emptyset) = 0 \end{cases}$$

and

$$(2)\quad \begin{cases} v^*(A) = \max\limits_{i \in A} \left\{ \dfrac{1}{\mu_i + \sum\limits_{j \notin A} \lambda} \left(1 + \mu_i v^*(A \setminus \{i\}) + \sum\limits_{k \notin A} \lambda v^*(A \cup \{k\}) \right) \right\} \\ v^*(\emptyset) = 0 \end{cases}$$

The remainder of this paper is organized as follows. In section 2 first the case N = 2 is considered, which leads to the guess that $v^f(A)$ is independent of f. Next in section 3 the general case is considered. Assuming $v^f(A)$ to be independent of f we obtain a function v which solves both (1) for all f and (2), and thus proves that $v^f(A)$ is independent of f, implying that the utilizaton factor of the CP is independent of f.

2. The case N = 2.

In this section we consider the simplest version of the problem: the case N = 2. For N = 2 there is actually only one state where an action has to be chosen, namely the state {1,2}. So there are only two schedulings. The strategy which appoints 1 is denoted by f (so f({1,2}) = 1) the other is denoted by g. For strategy f the system (1) becomes

$$(3)\qquad \begin{cases} v^f(\{1\}) = \dfrac{1}{\mu_1+\lambda}\,[1+\lambda\, v^f(\{2\})] \\ v^f(\{2\}) = \dfrac{1}{\mu_2+\lambda}\,[1+\lambda\, v^f(\{1\})] \\ v^f(\{1,2\}) = \dfrac{1}{\mu_1} + v^f(\{2\}) \end{cases} .$$

Solving this system yields us

$$(4)\qquad \begin{cases} v^f(\{1\}) = \dfrac{1}{\mu_1} + \dfrac{\lambda}{\mu_1\mu_2} \\ v^f(\{2\}) = \dfrac{1}{\mu_2} + \dfrac{\lambda}{\mu_1\mu_2} \\ v^f(\{1,2\}) = \dfrac{1}{\mu_1} + \dfrac{1}{\mu_2} + \dfrac{\lambda}{\mu_1\mu_2} \end{cases} .$$

We can do the same for strategy g. Specifying (1) for strategy g yields a system very similar to (3). Except for the replacement of the superscript f by g the first two equations are the same and the third equation becomes

$$v^g(\{1,2\}) = \frac{1}{\mu_2} + v^g(\{1\}) \ .$$

Solving the system for strategy g gives

$$v^g(\{1\}) = \frac{1}{\mu_1} + \frac{\lambda}{\mu_1\mu_2}$$

$$v^g(\{2\}) = \frac{1}{\mu_2} + \frac{\lambda}{\mu_1\mu_2}$$

$$v^g(\{1,2\}) = \frac{1}{\mu_1} + \frac{1}{\mu_2} + \frac{\lambda}{\mu_1\mu_2} \ .$$

The solution is identical to the solution (4) of (3). So the expected remaining lifetime of a busy period is not affected by the scheduling strategy. (That the solution for g is identical to the solution (4) for f also follows from the policy iteration method, cf. Howard [1], namely $\frac{1}{\mu_2} + v^f(\{1\}) = v^f(\{1,2\})$).

From the identity of the solutions for f and g it follows that other scheduling disciplines like FCFS, LCFS and PS give exactly the same result. Hence the utilization factor of the CP cannot be influenced by the scheduling.

3. The general case.

Now let us consider the case of an arbitrary number of terminals. The results for N = 2 suggest that also for arbitrary N the scheduling will have no influence. In this section we will prove this.

First, let us rewrite the first equation in (1) as

$$\mu_{f(A)}(v^f(A) - v^f(A\setminus\{f(A)\})) = 1 + \sum_{j\notin A} \lambda\,(v^f(A \cup \{j\}) - v^f(A)) \,. \tag{5}$$

If the functions v^f are indeed independent of f, then they are all equal to a function v (clearly also equal to v^*) and this function v will satisfy for all $A \subset I_N$ and all $i \in A$

$$\mu_i(v(A) - v(A\setminus\{i\})) = 1 + \sum_{j\notin A} \lambda\,(v(A \cup \{j\}) - v(A)) \,. \tag{6}$$

On the other hand if a function v satisfying (6) for all A and $i \in A$ exists, then this function also satisfies (5) and (1) for all A. But for each f the solution of (1) is unique since for fixed f the states $A \subset I_N$ constitute a Markov chain with $\emptyset$ as absorbing state, which is ultimately reached from any initial state. So the function v is equal to v^f for all f, implying that v^f is independent of f.

Below we will construct, using the equations (6), a function v for which we ultimately verify that it satisfies (6) for all A and $i \in A$.

Let v be a function satisfying (6), then

$$v(A) - v(A\setminus\{i\}) = \frac{1}{\mu_i} + \frac{\lambda}{\mu_i} \sum_{j\notin A} (v(A \cup \{j\}) - v(A))$$

$$= \frac{1}{\mu_i} + \frac{\lambda}{\mu_i} \sum_{j\notin A} \left[\frac{1}{\mu_j} + \frac{\lambda}{\mu_j} \sum_{k\notin A\cup\{j\}} (v(A \cup \{j,k\}) - v(A \cup \{j\})) \right]$$

$$= \frac{1}{\mu_i} + \frac{\lambda}{\mu_i} \sum_{j\notin A} \frac{1}{\mu_j} + \frac{\lambda^2}{\mu_i} \sum_{j\notin A} \frac{1}{\mu_j} \sum_{k\notin A\cup\{j\}} (v(A \cup \{j,k\}) - v(A \cup \{j\}))$$

(7)

$$= \ldots = \frac{1}{\mu_i} + \frac{\lambda}{\mu_i} \sum_{j\notin A} \frac{1}{\mu_j} + \frac{\lambda^2}{\mu_i} \sum_{\substack{j,k\notin A \\ j\neq k}} \frac{1}{\mu_j \mu_k}$$

$$+ \ldots + \frac{\lambda^{N-|A|}}{\mu_i} \sum_{\substack{j_1,\ldots,j_{n-|A|}\notin A \\ j_m \neq j_n \\ m\neq n}} \frac{1}{\mu_{j_1},\ldots,\mu_{j_{N-|A|}}}$$

In order to simplify the expressions in (7) let us introduce the following notation for all A and $\ell = 0,1,\ldots,N$

(8) $$C(\ell,A) := \{B \subset I_N \mid |B| = \ell \text{ , } A \cap B = \emptyset\} \text{ .}$$

So $C(0,A) = \{\emptyset\}$ and $C(\ell,A) = \emptyset$ for $\ell > N - |A|$.

Now (7) can be rewritten as

$$v(A) - v(A\setminus\{i\}) = \frac{1}{\mu_i} + \frac{\lambda}{\mu_i} \sum_{B\in C(1,A)} \prod_{j\in B} \frac{1}{\mu_j} +$$

(9) $$+ \ldots + \frac{\lambda^{N-|A|}(N-|A|)!}{\mu_i} \sum_{B\in C(N-|A|,A)} \prod_{j\in B} \frac{1}{\mu_j}$$

$$= \frac{1}{\mu_i} + \frac{1}{\mu_i} \sum_{\ell=1}^{N-|A|} \lambda^{\ell} \ell! \sum_{B\in C(\ell,A)} \prod_{j\in B} \frac{1}{\mu_j} \text{ .}$$

Now consider some arbitrary set $A = \{i_1, i_2, \dots, i_m\}$. Then we can write

$$\begin{aligned} v(A) &= [v(\{i_1,\dots,i_m\}) - v(\{i_1,\dots,i_{m-1}\})] + \\ &+ [v(\{i_1,\dots,i_{m-1}\}) - v(\{i_1,\dots,i_{m-2}\})] + \\ &+ \dots + [v(\{i_1,i_2\}) - v(\{i_1\})] + [v(\{i_1\}) - v(\emptyset)] , \end{aligned} \tag{10}$$

with $v(\emptyset) = 0$. With the shorthand notation $A_k := \{i_1,\dots,i_k\}$, $k = 1,\dots,m$, $A_0 = \emptyset$, (10) becomes

$$\begin{aligned} v(A) &= \sum_{k=1}^{m} (v(A_k) - v(A_{k-1})) \\ &= \sum_{k=1}^{m} \left[\frac{1}{\mu_{i_k}} + \frac{1}{\mu_{i_k}} \sum_{\ell=1}^{N-|A_k|} \lambda^{\ell} \ell! \sum_{B \in C(\ell,A_k)} \prod_{j \in B} \frac{1}{\mu_j} \right] . \end{aligned} \tag{11}$$

Now define $D(\ell,A)$ for all A and $\ell = 1,2,\dots,N$ by

$$D(\ell,A) := \{B \subset I_N \mid |B| = \ell , A \cap B \neq \emptyset\} . \tag{12}$$

Substituting the notation (12) in (11) yields

$$\begin{aligned} v(A) &= \sum_{B \in D(1,A)} \prod_{j \in B} \frac{1}{\mu_j} + \lambda \sum_{B \in D(2,A)} \prod_{j \in B} \frac{1}{\mu_j} + \dots + \\ &+ \lambda^{N-1} (N-1)! \sum_{B \in D(N,A)} \prod_{j \in B} \frac{1}{\mu_j} \\ &= \sum_{\ell=0}^{N-1} \lambda^{\ell} \ell! \sum_{B \in D(\ell+1,A)} \prod_{j \in B} \frac{1}{\mu_j} . \end{aligned} \tag{13}$$

By (13) a function v has been defined.

Finally it has to be verified that the function v defined by (13) satisfies (6) for all A and all $i \in A$. First consider the left hand side in (6).

(14)
$$\mu_i(v(A) - v(A\setminus\{i\})) = \mu_i\left[\sum_{\ell=0}^{N-1} \lambda^\ell \ell! \sum_{B\in D(\ell+1,A)} \prod_{j\in B} \frac{1}{\mu_j} - \sum_{\ell=0}^{N-1} \lambda^\ell \ell! \sum_{B\in D(\ell+1,A\setminus\{i\})} \prod_{j\in B} \frac{1}{\mu_j}\right] .$$

Using

$$D(\ell+1,A)/D(\ell+1,A\setminus\{i\})$$
$$= \{B \subset I_N \mid |B| = \ell+1 , A \cap B \neq \emptyset , (A\setminus\{i\}) \cap B = \emptyset\}$$
$$= \{B \subset I_N \mid |B| = \ell+1 , A \cap B = \{i\}\}$$
$$= \{B \cup \{i\} \mid B \in C(\ell,A)\}$$

(14) simplifies to

(15)
$$\mu_i(v(A) - v(A\setminus\{i\})) = \mu_i \sum_{\ell=0}^{N-1} \lambda^\ell \ell! \sum_{B\in C(\ell,A)} \frac{1}{\mu_i} \prod_{j\in B} \frac{1}{\mu_j}$$
$$= \sum_{\ell=0}^{N-1} \lambda^\ell \lambda! \sum_{B\in C(\ell,A)} \prod_{j\in B} \frac{1}{\mu_j} = \sum_{\ell=0}^{N-|A|} \lambda^\ell \ell! \sum_{B\in C(\ell,A)} \prod_{j\in B} \frac{1}{\mu_j} .$$

Remains to consider the right hand side in (6). Substituting (13) yields

(16)
$$1 + \sum_{j\notin A} \lambda(v(A \cup \{j\}) - v(A))$$
$$= 1 + \sum_{j\notin A} \frac{\lambda}{\mu_j} \sum_{\ell=0}^{N-1} \lambda^\ell \ell! \sum_{B\in C(\ell,A\cup\{j\})} \prod_{k\in B} \frac{1}{\mu_k} =$$

$$= 1 + \sum_{\ell=0}^{N-|A|-1} \lambda^{\ell+1}\, \ell! \sum_{j \notin A} \frac{1}{\mu_j} \sum_{B \in C(\ell, A \cup \{j\})} \prod_{k \in B} \frac{1}{\mu_k}$$

$$= 1 + \sum_{\ell=0}^{N-|A|-1} \lambda^{\ell+1}\, \ell!\,(\ell + 1) \sum_{B \in C(\ell+1, A)} \prod_{k \in B} \frac{1}{\mu_k}$$

$$= \sum_{\ell=0}^{N-|A|} \lambda^{\ell}\, \ell! \sum_{B \in C(\ell, A)} \prod_{k \in B} \frac{1}{\mu_k} \; .$$

Comparing (15) and (16) we see that (6) holds indeed for all A and $i \in A$ for the function v defined by (13). So, as argued before, $v = v^f = v^*$ for all schedulings f. Hence the utilization factor of the CP is independent of the scheduling. And this is not only the case for the scheduling strategies considered in this section but, as one easily shows, also for disciplines like FCFS, LCFS and PS.

A consequence of this result is that we can use for instance a discipline like smallest job first or a dynamic priority rule based on already received attention without affecting the utilization factor of the CP.

References.

[1] Howard, R.A., Dynamic programming and Markov processes, Wiley, New York, 1960.

[2] Wal, J. van der, The maximization of CP utilization in an exponential CP-Terminal system with different think times and different job sizes, Eindhoven Univ. of Technology, Dept. of Math. and Comp. Sci., Memorandum COSOR 82 - 14