ABSTRACT

Considering queueing systems in data networks one can identify a special feature of the communication system: the service time, which is proportional to message length, is known in advance. This is a precondition for queueing strategies such as SPT (Shortest Processing Time) or SRPT (Shortest Remaining Processing Time), which offers the shortest mean delay time among all conceivable strategies. For a practical evaluation of these disciplines it is essential to include the influence of overhead time, especially for the preemptive SRPT strategy. This paper analyses the SRPT strategy with constant overhead time $T_O$ modified by a constant preemption distance $T_P > T_O$ for $M/G/1$ systems. The comparison of strategies: FIFO, SPT, SRPT, and Round Robin for different service time distributions and overhead show a considerable reduction of the mean delay time $T_D$ for the SPT and SRPT strategies, especially for service time distributions with high coefficients of variation. This result indicates the potential advantages of the SRPT strategy in data network applications.

The main intent of this paper is to investigate the potential reduction of the mean delay time for these strategies in comparison with other preemptive and non-preemptive strategies.

For a valid practical evaluation of preemptive strategies it is necessary to include overhead time to account for the time needed to handle the preemption. In this paper a modified SRPT strategy with overhead time is presented.

A short description of the methods used for the analysis and computation is given in section 3 and 4 /4/. Additional formulae are given in the appendix. The modified SRPT strategy with overhead time is then compared to the FIFO, SPT, and Round Robin (RR) strategy with overhead time /5/. The comparison is based on a stationary $M/G/1$ model and service time distributions with different coefficients of variation.

2. QUEUEING STRATEGIES WITH OVERHEAD

Considering preemptive queueing strategies with overhead in general, one is led to the definition of at least four different types of overhead:

- set up time for a new job;
- take down time for a preempted job;
- set up time for a preempted job;
- take down time at the end of a job.

To reduce the number of parameters in the analysis and comparison, we restrict our model to constant set up times $T_V$ for new and preempted jobs. Simulation results show that a good approximation for a model with constant set up times $T_1$ and constant take down times $T_2$ can be achieved with the TV model with $TV=T_1+T_2$.

In the subsequent discussion we confine ourselves to $M/G/1$ queueing systems with constant set up or overhead time $TV$ and the following notation:

- $\lambda$ : input intensity
- $TA$ : service (processing) time random variable
- $F_A(t)$ : probability distribution function of $TA$
- $f_A(t)$ : probability density function of $TA$
- $T_A$ : service intensity ($=\lambda/TA$)
- $c_A$ : coefficient of variation for $TA$.
2.1 Non-Preemptive Strategies

In the case of non-preemptive strategies overhead times do not present a difficult problem. The results known for these strategies /6/ can be applied after adding the overhead time $TV$ to the service time random variable $TA$. The new service time random variable is then defined as $TA' = TA + TV$. Introducing overhead time also increases the utilization factor of the system. The basic load $p = \lambda / \mu$ is used as a basis of comparison. For non-preemptive strategies the utilization factor is

$$p' = \lambda \cdot TA' = p + \lambda \cdot TV$$  (1)

The mean delay time $TD$ is equal for all strategies that select jobs in a way that is independent of their service time. In our comparison we use the FIFO strategy as a representative for these strategies. The SPT strategy is also used in the comparison as the optimal non-preemptive strategy.

Fig.1 M/G/1 model for non-preemptive queueing strategies with overhead time $TV$.

2.2 Preemptive Strategies

2.2.1 Modified SRPT with Overhead (SRPT/TV/TP)

The analysis of preemptive strategies with overhead is more difficult, because the overhead occurs for every preemption, hence it occurs a variable number of times during a service depending on the strategy and system load.

Starting out with the optimal strategy for systems without overhead, i.e. the Shortest Remaining Processing Time (SRPT) discipline /1,2/ our main interest is the evaluation of this strategy with overhead. To account for the overhead induced by a preemption a constant preemption distance $TP$ is defined and the strategy is modified to allow a preemption at time $t$ only if the difference between the rest service time at time $t$ of the job in service $TR(t)$ and the service time of a new job $TA$ is greater than the preemption distance:

$$TR(t) - TA > TP; TP >= TV$$  (2)

The structure of this system is shown in fig.2. The queue of the system can be viewed as a list sorted according to the rest service times of the jobs. Whenever the server is empty the job with the shortest remaining processing time (rest service time) is processed next.

Fig.2 M/G/1 model for the SRPT strategy with overhead time $TV$. (TR: rest service time, TU: undivided service time)

If a new job causes a preemption according to the preemption rule eq.(2), the preempted job is returned to the queue. A job cannot be preempted during the execution of the overhead time. In this case the preemption is delayed. This corresponds to the definition of overhead as a supervisory task, that would leave the system in an undefined state on preemption. The dashed lines in fig.2 show the possibilities for preemptions. When preempted during its actual service time, the job is returned to the queue with its rest service time reduced by the amount of service time obtained ($TR := TR - TU$).

2.2.2 Round Robin with Overhead (RR)

The second preemptive strategy considered in this paper is the Round Robin strategy with overhead /5/. Fig.3 shows the structure of this system.

Fig.3 M/G/1 model for the Round Robin strategy with overhead time $TV$. (TS: maximum time slice, TR: rest service time, $TQ = \min \{TR, TS\}$)

There are two main differences to the SRPT strategy. First, a job is always preempted after it has been served for a constant time slice $TS$, i.e. preemptions are independent of arrivals. Second, the order within the queue is FIFO. If a job finishes within the maximum time slice, it leaves the system ($TR <= TS$).
3. METHOD OF ANALYSIS

This section describes the methods used to derive results concerning the mean delay time $TD$ in the SRPT/TV/TP system. First, the service time random variable is transformed to include the overhead for a specific job. Then, the two main constituents of the delay time, the initial waiting time $TW$ and the completion time $TC$, are studied.

3.1 Service Time with Overhead

Considering a job with given service time $a$, the new random variable $TA^*(a)$ can be defined as the service time including the overhead time needed to start and restart this job: thus $TA^*(a)$ contains the constant overhead time $TV$ at least once for the first start of the job and once for every preemption. With $NV(a)$ defined as the number of overhead times for a job of given service time $a$, we can write $TA^*(a)$ as

$$TA^*(a) = a + NV(a) \cdot TV$$

Restricting ourselves to Poisson input with parameter $\lambda$, it can then be shown [4] that preemptions for a job with rest service time $r$ also arrive according to a Poisson distribution with parameter $\lambda(r) = \lambda \cdot F_A((r-TP)^-)$

$$\lambda(r) = \lambda \cdot F_A((r-TP)^-)$$

where $F_A(t^-)$ is a notation for taking the limit from the left. We now divide the given service time $a$ into $k$ small intervals of length $\Delta t$ and obtain the number of overhead times for every interval $\Delta t$, restricting preemptions in the $j$-th interval to jobs with service times smaller than $(j-1)\Delta t-TP$. Taking the limit of the sum of random variables or product of the Laplace transforms for $\Delta t \to 0$ and $k \to \infty$ we obtain the Laplace transform of the probability density of $TA^*(a)$:

$$L_{TA^*(a)}(s) = \frac{1}{g(a)}$$

Using this result we find the first moment of $TA^*(a)$:

$$TA^*(a) = -L_{TA^*(a)}'(s) \bigg|_{s=0} = a + TV \cdot \exp(\lambda(a) \cdot TV) \int_0^a \frac{\lambda(t) \cdot TV \cdot \exp(\lambda(t) \cdot TV) dt}{\lambda(t)}$$

consisting of the given service time $a$, a term accounting for the first overhead time and its repetitions, and a term accounting for the overhead times induced by later preemptions. The total mean service time with overhead and the utilization factor $\rho^*$ are then given by:

$$TA^* = \int_0^\infty f_A(t) \cdot TA^*(t) dt$$

$$\rho^* = \lambda \cdot TA^*$$

Moments of higher order can be derived in a similar fashion. Especially the second moment is needed to determine the mean residual service time with overhead of a partially served job, see eq. (A9,A10,A13,A14).

3.2 Mean Delay Time

The delay time $TD(a)$ of a job with given service time $a$ corresponds to the total time of the job in the system. For further analysis the delay time is divided into two parts. The first part, called initial waiting time $TW(a)$, describes the time from the entry of the job in the system until it is served for the first time. The second part, called completion time $TC(a)$, includes the service and overhead time as well as the waiting time due to preemptions. The delay time depending on the given service time $a$ is the sum:

$$TD(a) = TW(a) + TC(a)$$

and the total mean delay is given by

$$\overline{TD} = \int_0^\infty f_A(t) \cdot (\overline{TW(t)} + \overline{TC(t)}) dt$$

3.2.1 Mean Initial Waiting Time

For the derivation of the mean initial waiting time $TW(a)$ we use the results obtained for non-preemptive priority systems with the transformed service time $TA^*$ instead of $TA$. The initial waiting time is not affected by the order of service of jobs served first according to the strategy and we can use the non-preemptive approach, because the effect of preemptions has already been considered with $TA^*$.

A job has priority $i$ if its service time is greater than $(i-1)\Delta t$ and smaller than or equal to $i\Delta t$. The observed job $A$ has priority $m$ with $m\Delta t = a$. Priority $m+1$ contains all jobs with service time greater than $m\Delta t$. With $m+1$ priorities we obtain our result by taking the limit $\Delta t \to 0$ and $m \to \infty$ with $m^\Delta t = a$. The main complication here is the derivation of the mean residual service time with overhead $\overline{TRA}$ of the job $B$ in service when the observed job $A$ has given service time $a$ and priority $m$. Several classes of $B$-jobs have to be distinguished according to the rest service time $r$ and original service time $b$ as seen in table 1.

For classes 1,2 and 5 the mean residual service times $\overline{TRA}_1(a)$, $\overline{TRA}_2(a)$, $\overline{TRA}_5(a)$ have been derived
eq. (A13, A14, A15). Classes 3 and 4 present the difficulty that the arriving job A with service time \(-a-\) does not preempt the job in service B with rest service time \(r\), but another job C arriving later with smaller service time could preempt job B during the interval \(a < r < a + TP\), so that job B would return to the queue and would exit the system after the observed job A.

A lower and upper bound for \(TW(a)\) can be determined by a minimum value \(TW_{\text{min}}(a)\), which is obtained by neglecting classes 3 and 4 altogether, and a maximum value \(TW_{\text{max}}(a)\), which is obtained by neglecting the preemptions of classes 3 and 4 eq. (A16-A20). For relatively small values of \(TP\) there is only a small difference between the lower and the upper bound, which gives us a good approximation of \(TW(a)\).

### Table 1 Residual Service Time Classes

<table>
<thead>
<tr>
<th>class</th>
<th>original priority</th>
<th>original service time</th>
<th>rest service time (r)</th>
<th>original service time (b)</th>
<th>system state</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>(i = m)</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>2</td>
<td>(i = m + 1)</td>
<td>(r = a)</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>3</td>
<td>(i = m + 1)</td>
<td>(a &lt; r &lt; a + TP)</td>
<td>(b = a + TP)</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>4</td>
<td>(i = m + 1)</td>
<td>(a &lt; r &lt; a + TP)</td>
<td>(b &gt; a + TP)</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>5</td>
<td>(i = m + 1)</td>
<td>(r &gt; a + TP)</td>
<td>--</td>
<td>TV</td>
<td></td>
</tr>
</tbody>
</table>

#### 3.2.2 Mean Completion Time

For the derivation of the mean completion time we use the same partition of the given service time \(-a-\) of our observed job as in 3.1. We now define \(TC_j(a)\) as the completion time of the \(j\)-th interval, consisting of the service time with overhead in the \(j\)-th interval \(TA^*_j(a)\) and the waiting time due to preemptions within the \(j\)-th interval \(TQ_j(a)\):

\[
TC_j(a) = TA^*_j(a) + TQ_j(a) \quad (12)
\]

\(TQ_j(a)\) again depends on the completion time \(TC_j(a)\):

\[
TQ_j(a) = TC_j(a)D(j-1) + P(j)[G(j) - H(j-1)] \quad (13)
\]

\[
TC_j(a) = \frac{TA^*_j(a) + P(j)[G(j) - H(j-1)]}{1 - D(j-1)} \quad (14)
\]

where \(D(j-1)\) stands for the mean prolongation factor for the \(j\)-th interval due to jobs with service times smaller than \((j-1)\cdot \Delta t\). In the SRPT/TV/TP strategy, however, we have to modify the above expression in the case of preemptions with probability \(P(j)\). First we have to add a term \(G(j)\) that accounts for jobs that arrived before the \(j\)-th interval, but their processing was delayed until the job was preempted in the \(j\)-th interval. These jobs have processing times between \((j-1)\cdot \Delta t - TP\) and \((j-1)\cdot \Delta t\). Accordingly we have to subtract a term \(H(j-1)\), which contains those jobs that arrived in the last unpartitioned part of the \(j\)-th interval with processing times in the same interval. We find the mean completion time of a job with given service time \(-a-\) as the limit of the following sum:

\[
\bar{TC}(a) = \lim_{m \to \infty} \sum_{j=1}^{m+1} TC_j(a) = G(a) + \sum_{j=1}^{m+1} TC_j(a) \quad (15)
\]

\[
[TV\cdot \exp\{\lambda(a)\cdot TV\} - \lambda \cdot \int_{a-TP}^{a} b(y)\,dy] / [1 - D(a^-)] +
\]

\[
\int_{0}^{x} \left[ 1 + \lambda(x)\cdot TV\cdot \exp\{\lambda(x)\cdot TV\} - \lambda(\lambda(x)TV+1) \int_{0}^{x-TP} b(y)\,dy \right] \frac{dx}{1 - D(x^-)}
\]

with \(\lambda(x)\) as defined above and

\[b(y) = f_A(y)\cdot TA^*(y)\]

\[D(x) = \int_{0}^{x} \lambda\cdot b(y)\,dy\]

\[G(a)\] accounts for that part of the limit which contains \(G(j)\). It is a rather complex formula with integration depth 4 eq. (A21, A22). The interval with index \(m+1\) stands for the processing of the first overhead time, which has to be evaluated separately.

### 4. METHOD OF COMPUTATION

The formulae derived for the mean delay \(\bar{TD}\) of the SRPT/TV/TP strategy contain nested integrals that cannot in general be reduced to elementary formulae. This also applies to the SPT and SRPT strategy without overhead. A numerical method had to be implemented for the comparison of strategies. There are two possible methods applicable here: numerical integration or discrete approximation of the continuous service time distribution and evaluation of the integrals as sums.

We chose the second approach, i.e. continuous distributions are approximated with a given degree of accuracy. For discrete service time distributions we obtain exact results.

An integral can always be evaluated as a sum, if the function to be integrated changes its value only a finite number of times within the integration interval. This applies to the integrals in our formulae implemented so far. There are two types of sums in our evaluation. If \(t_j, j=1..n\) denotes the discrete service time values, then we have to integrate functions that change their value at exactly these points, and we have functions that also change their value at \(t_j + TP\).

Our implementation contains the evaluation of the lower and upper bound of \(\bar{TD}\):

\[
\bar{TD}_{\text{min}} = \bar{TW}_{\text{min}} + \bar{TC}_{\text{min}} \quad (18)
\]

\[
\bar{TD}_{\text{max}} = \bar{TW}_{\text{max}} + \bar{TC}_{\text{max}} \quad (19)
\]
Although $\overline{TC}$ is given as an exact formula it is evaluated as $\overline{TC}_{\min}$ by neglecting $G(a)$, and as $\overline{TC}_{\max}$ by implementing a larger expression for $G(a)$ in eq. (A21). For relatively small values of $TP$ up to $TP<\overline{TA}/10$ we obtain good results with a relative difference of about $1-2\%$, which is in the order of the precision of the given diagrams. For the actual utilization factor $\rho^*$ we have implemented the exact formula for discrete distributions.

5. COMPARISON OF STRATEGIES

In the following comparison we use the hyperexponential service time process $H_2$ /7/, which is shown in table 2.

| $c_A = \sqrt{2}$, $\mu = 1$: $p = 0.79$, $\mu_1 = 1.58$, $\mu_2 = 0.42$ |
| $c_A = 5$, $\mu = 1$: $p = 0.98$, $\mu_1 = 1.96$, $\mu_2 = 0.04$ |

Table 2 Service time process $H_2$ for $c_A = 1$

To demonstrate the effect of overhead time for the SRPT/TV/TP strategy fig. 4 shows the mean delay $TD$ for this strategy as a function of the basic load $p = \lambda/\mu$ for the overhead time in the range of $0.20\%$ of the mean service time. The preemption distance is equal to the overhead time ($TP=TV$) for this $M/H_2/1$ system with $c_A=5$. This example displays the increase of mean delay with respect to overhead time and, as can be seen in the upper part of fig. 4, the increase of the utilization $\rho^*$, thus reducing the maximum load $\rho$ allowable for a stable system.

A first comparison of strategies with respect to mean delay time is given in fig. 5. The strategies compared are FIFO, SPT, Round Robin (RR), and SRPT/TV/TP for the $M/H_2/1$ system with $c_A=2$ and $p=0.8$. The time slice $TS$ for the Round Robin strategy was chosen as the optimal time slice, which depends on the overhead time /5/. The Round Robin strategy shows an improvement relative to the FIFO strategy and further improvement is achieved for those strategies that use information about the processing time, SPT and SRPT/TV/TP.

Comparisons for theoretical and empirical distributions /8/ have shown, that the difference of mean delay between the strategies FIFO, SPT, and SRPT increases with the variance of the service time distribution. A lower bound for the mean delay of the FIFO strategy is defined by the deterministic case with $c_A = 0$. It is interesting to note, that in some cases the mean delay for SPT and SRPT is even below this bound for FIFO.
The difference between strategies is demonstrated in fig. 6 for the M/H₂/1 system with cₐ=5 up to the point (p>0.85) where for the chosen preemption distance (TP=TV) the overhead time for preemptions outweighs the advantages of changing the order of jobs. To obtain results that are better or at least as good as the SPT strategy, we have to increase the preemption distance with the basic load p.

Fig. 7 shows an approximation of the influence of the preemption distance TP with respect to mean delay for the SRPT/TV/TP strategy. The mean delay for the SPT strategy, which does not depend on the preemption distance, is given as a point of reference in fig. 7. The approximation for the mean delay time is obtained by combining maximum waiting time and minimum completion time:

$$\mu_TD' = \mu_TW_{max} + \mu_TC_{min}$$

(20).

This approximation is obviously always within the interval defined by the lower and upper bound for the mean delay. Thus for small values of TP the approximation is valid because the bounds are very close. The approximation was checked against values obtained by simulation, which are represented in fig. 7 as intervals with symmetric absolute errors $\epsilon_u$, $\epsilon_l$. Fig. 7 clearly shows that choosing a preemption distance TP greater than TV can reduce the mean delay time. With this approximation it will be possible to evaluate the optimal preemption distance as a function of the basic load and overhead time.

6. CONCLUSION AND FUTURE RESEARCH

The analysis of the SRPT/TV/TP strategy has shown that the advantages of the SRPT strategy are fully preserved, if the overhead time can be kept within reasonable limits. However, the application of the SRPT/TV/TP strategy in today's data networks, packet or message switching systems, would require extensive technical changes, but it would certainly be possible to implement this strategy in important elementary data communication systems, e.g. a simple concentrator configuration for the direct connection of a large number of terminals to a central computer.

Future research will be dedicated to the evaluation of the optimal preemption distance TP and to the approximation of this strategy by simple formulae depending only on a few characteristic parameters of the service time distribution.
Appendix

The following definitions are used within the appendix:

\[ \mu_2(X) : \text{second moment of random variable } X \]

\[ \lambda(x) = \lambda(A_X(x)) \]

\[ r(x) = \exp(\lambda(x) \cdot TV) \]

\[ b(x) = f_A(x) \cdot T_A(x) \]

\[ I_1(a, b) = \int_a^b b(x) \, dx \]

\[ I_2(a, b) = \int_a^b \lambda(x) \, dx \]

\[ I_3(a, b) = \int_a^b \lambda(x) \cdot T(x) \, dx \]

\[ I_4(a, b) = \int_a^b \lambda(x) \cdot T^2 \cdot [2 \cdot T(x) - 1] \cdot r(x) \, dx \]

\[ D(x) = \lambda \cdot I_1(o, x) \]

A.1 Service Time with Overhead

\[ \mu_2(TA^*(a)) = I_4(o, a) + \left[ a \cdot TV + I_3(o, a) \right]^2 \]

\[ TA^* : \text{service time with overhead due to preemptions} \]

\[ \mu_2(TA^-(a)) = I_4(o, a) + \left[ a \cdot I_3(o, a) \right]^2 \]

A.2 Mean Initial Waiting Time

\[ TW(a) = \frac{\tau\lambda(a)}{[1-D(x^-)][1-D(x^+)]} \]

\[ \tau\lambda(a) = \sum_{j=1}^{5} \tau\lambda_j(a) \]

\[ \tau\lambda_1(x) = \frac{(\lambda/2) \int_0^{\mu_2(TA^*(y))} f_A(y) \, dy}{\int_0^{\infty} \left[ \mu_2(TA^*(y)) - \mu_2(TA^-(x+TP)) - y+TP \right] f_A(y) \, dy} \]

\[ \tau\lambda_2(x) = \frac{(\lambda/2) \cdot \mu_2(TA^*(x)) \cdot [1-F_A(x)]}{\int_0^{\infty} \left[ TA^*(y) - TA^-(x+TP) - y+TP \right] f_A(y) \, dy} \]

\[ \tau\lambda_3(x) = \frac{(\lambda/2) \cdot T(x)}{\int_0^{\infty} \left[ TA^*(y) - TA^-(x+TP) - y+TP \right] f_A(y) \, dy} \]

\[ \tau\lambda_4(x) = \int_0^{\infty} \left[ TA^*(y) - TA^-(x+TP) - y+TP \right] f_A(y) \, dy \]

\[ \tau\lambda_5(x) = \int_0^{\infty} \left[ TA^*(y) - TA^-(x+TP) - y+TP \right] f_A(y) \, dy \]

\[ \tau\lambda_1(a) \leq \tau\lambda_2(a) \leq \tau\lambda_3(a) \leq \tau\lambda_4(a) \leq \tau\lambda_5(a) \]

\[ TW_{\min}(a) = \frac{\tau\lambda_1(a) + \tau\lambda_2(a) + \tau\lambda_3(a)}{[1-D(x^-)][1-D(x^+)]} \]

\[ TW_{\max}(a) = \frac{\tau\lambda_1(a+TP) + \tau\lambda_2(a+TP) + \tau\lambda_3(a)}{[1-D(x^-)][1-D(x^+)]} \]

A.3 Mean Completion Time

\[ G(a) = \lambda \int_0^a \frac{f_A(x-TP)}{[1-D(x^-)][1-D(x^+)]} \int_{s=x}^{a} h(s) F_A(s-TP) ds + h(a) \, dx \]

\[ h(z) = \exp(-I_2(x,z)) \cdot \left[ \int_{y=x}^{z} I_1((y-TP)^+, x) \, dy + TV \cdot I_1((z-TP)^+, x) \right] \]

Acknowledgement

I would like to thank Prof. Dr. F. Schreiber for his continuous support of this project, Dipl.-Ing. H.-D. Suchofen for many valuable discussions, and Dipl.-Ing. M. Kochan for verifying these results by simulation.

References


ITC 10

Summary of Questions/Answers

Date: 13 June 1983
Session: 3.3
Paper: 7

Q.1 (Werner Bux)

Your results show that by the SRPT strategy, delays can be significantly reduced, if the overhead is not too big and the service times have a relatively high variance. Do you have an explanation why actual implementations do not attempt to benefit from this effect?

A.1 (C. Goerg)

The advantages of the SRPT strategy have been known for some time, but in most queueing environments the processing time is not known in advance, so that the strategy cannot be implemented, and for this reason was not further pursued.

In data networks however the processing time corresponds to the message. For an actual implementation it is important to evaluate the influence of overhead for preemptions. This was the main goal of this project and the results show that the advantages of the SRPT strategy are preserved. Estimations show that an overhead of 10% of the mean processing time is realistic. However, it would require considerable technical changes to implement this strategy in today's data networks. - Another aspect of this strategy is the use of these results as a standard of comparison for other strategies that make use of the knowledge of processing time, but are not as difficult to implement. This is a subject to be studied further.