MTA Számitástechnikai és Automatizálási Kutató Intézete, Közlemények 18/1977.

A DETAILED, SYSTEM OVERHEAD INCLUDING, DESCRIPTION OF PRIORITY ALGORITHMS FOR REAL TIME TASK SCHEDULING

Adam Wolisz
Department of Complex Automation Systems of the Polish Academy of Sciences

1. Introdution

Author's point of view on the problem of creating an proper model describing real-time opertaing systems (0.S) has been presented in paper [1]. One can find there some characteristic features of O.S. as well as a classification of scheduling algorithms used in the real-time applications. The short overview of available theoretical results for some of these algorithms has also been provided, while a more detailed survey can be found in [2].

Up to now priority scheduling algorithms based on the minicomputers priority interrupt feature are implemented in the vast majority of O.S. Because of the hard constraints on reponse times imposed by real-time environment, a deep knowledge of their time characteristics is of essential importance. thus a proper queuing model has to be developed.

It can be easily demonstrated, that simple priority disciplinces like head of the line or preemptive resume disciplines, are not sufficient to obtain such a model. Under the head of the line discipline high-priority jobs can not receive desired response times. On the othet hand, mainly because of synchronization objectives and data integrity reasons, the use of strictily preemptive priority disciplines is not possible. During the execution of almost every programme there are some parts which should not be preempted, and which are therefore executed after disabling the priority interrupt system.

So the execution of a programme can be treated as service given in phases of either preemptive or nonreemptive type.
The existance of any given phase in the exacuted programme may also depend on data upon which the current exacution is fulfilled.

Implementation of various queuing disciplines is connected with some overhead which should be also included in the analysis. This overhead is usually caused by activities fulfilled in the privileged mode and as such can not be interrupted by any demand no matter of what priority.

The study of models reflecting the above mentioned features was started in [3] and then developed under more general assumptions in [4] and [5].
However all those investigations took place for the case when demands for programme's exacution belonged to a Poisson stream, thus open queuing systems had to be considered. This assumption is not justified, as in the majority of real-time systems one has to deal with demand's sources of finite dimension.
Corresponding to this situation closed queuing systems are difficult to analyze especially for the case of complex service disciplines, like multiphase priority service mentioned earlier.

Operational parameters of those systems, like waiting times for various demands, are often estimeted by the analysis of a corresponding system with either infinite-dimensional demand's sources (leading to an upper bound) or one dimensional demand's sources (which gives an lower bound).

The relative error caused by such an estimation in the first case has been investigated by Bazen and Goldberg [6], and according to their results may achieve even several hundred percents.

On the other hand, it can be proved that in the case when one-dimensional estimation is used the relative error may not exceed one hundred percent.

Besides, one dimensional sourced describe often the actual situation existing in technological plants.

Queuing systems with one dimensional sources create also a proper model for multiprogramming systems with a fixed number of nonhomogeneus jobs, having independent I/0 facilities, as it was pointed out in [7].

Such a systems have been investigated so far only in the case of simple priority algorithms and for the processor sharing model. In this paper the analysis of a generalized priority discipline corresponding to the real-time requirements is presented for th case of N independent one-dimensional sources of demands.

2. Description of the considered model

We shall consider a single server queuing system with N one-dimensional sources generating demands $z_1, z_2, \ldots, z_k, \ldots, z_N$ respectively. The time spent in the source by the demand z_k is a stochastic variable having an exponential distribution with parameter λ_k. The demand z_k (called herafter k-type) will have priority over the demand

$$z_\lambda \quad \text{if} \quad k < I.$$

The service of demand z_k is given in phases denoted F_k^i. Any phase F_k^i may be either of preemptive type (if $i \epsilon I_k$) or of nonpreemptive type (if $i \epsilon J_k$), where I_k, J_k are proper exclusive stes.

Everly phase has a given probability of it's execution r_k^i and an arbitrary conditional probability density function(pdf) of it's execution time $s_k^i(x)$.

Further we shall use frequently the total pdf $S_k^i(x)$ given by

$$(2.1.) \qquad S_k^i(x) = \delta(x)[1 - r_k^i] + s_k^i(x) \cdot r_k^i,$$

where $\delta(x)$ is the Dirac function.

The joint pdf of k-th programme execution $S_k(x)$ can be obtained basing on S_k^i.

If for every $i \neq j$, S_k^i and S_k^j are statistically independent, then*)

$$\bar{S}_k(s) = \prod_{i \in (I_k \vee I_k)} \bar{S}_k^i(s).$$

In other cases $\bar{S}_k(s)$ must be derived according to the form of dependence between S_k^i and $S_k^j (i,j \in \{I_k V J_k\},\ i \neq j)$.

For the preemptive phase F_k^i, $i \in I_k$ we shall assume the overhead connected with preemption and the overhead connected with resuming the preempted execution to be of nonpreemptive type and have the lenght given by arbitrary pdf's $Q_k^i(x)$ and $R_k^i(x)$ respectively. We shall also assume that if during R_k any demand of type 1, $1 < k$ will be generated, it shall be regarded as next, independent preemption as presented in fig. 1.

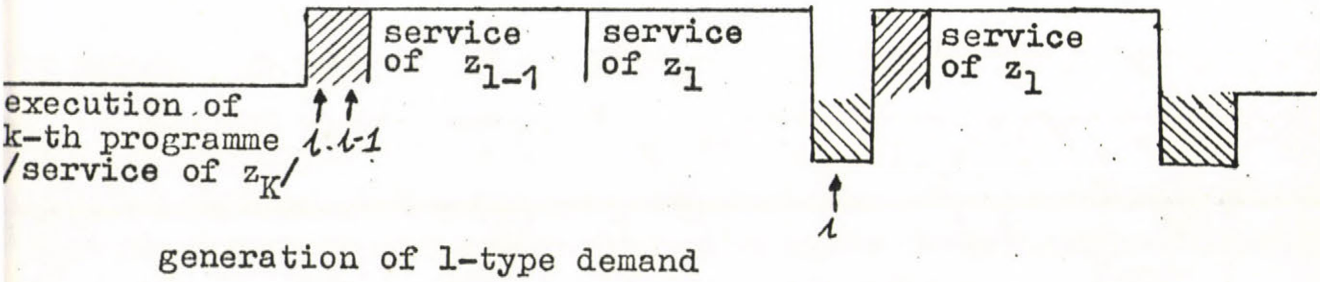

Fig.1 Illustration of the assumptions according the overhead during the preemptive phases

For the nonpreemptive phases F_k^i, $i \in J_k$ we shall assume, that if during the service of the phase F_k^i any demand of higher priority was generated, then after finishing F_k^i an overhead with arbitrary pdf $Q_k^i(x)$ should take place. Only after that the service of higher priority demands can be initiated.

This model will be investigated using methods similar to presented in [9] and all concepts which will not be defined here precisely are used in the meaning of [9]. In this paper the outline of the analisis and main results are given, while all additional detailes are described in [2].

* Let $f(t)$ be any given pdf. We shall denote by $f(s)$ the Laplace transform of $f(t)$

$$\bar{f}(s) = \int_0^\infty e^{-st} f(t) dt.$$

and by f any sample from the distribution $f(t)$.

Our considerations will consist of three steps. First we shall discuss the busy period for some basic schemes, where the server deals only with demands generated by a single source. Afterwards the busy period for all types of demands will be investigated.
Finally we shall describe the full process consisting of a sequence of busy and idle periods consecutively. We shall also investigate the waiting times for demands of various types.

3. **Basic service schemes**

Let us now discuss busy period parameters for four basic schemes of serving demands from one source with parameter λ which require service having a pdf $S(x)$.

a.) A scheme with regeneration

We shall assume, that the server works in a following mode. At time $t = 0$ a first service starts, and goes on for some time Z. A part V of this service time is devoted to an "incorrect" service in the sense, that after completion of the service, server has to be regenerated. This regeneration takes some time y depending on V in the way described be a conditional pdf $r(y/V)$. Both the times Z and V are given by a joint pdf $f_{ZV}(Z,V)$. During the regeneration a consecutive demand can be generated, which service can start only after completing of the regeneration.

Let us denote by $p_m(x,t,Z, \upsilon)dxd\upsilon dZ$ the pdf of the following event: at time t service is in progress with the elapsed service of the demand being equal to $x,^*$ $x < x^* < x + dx$, while the service will take time Z^* including time V^* of "incorrect" service, $Z < Z^* < Z + dZ$, $\upsilon < \upsilon^* < \upsilon + d\upsilon$ $q_m(m,y,t\ \upsilon)dyd\upsilon$ -the joint pdf of the following events: At time t regeneration of the server is in progress for a period equal to y^*, and the "incorrect" part of the last fulfilled service took time V^*, $y < y^* < y + dy$, $V < V^* < V + dV$, $m = 0{,}1$ denotes the number of demands waiting for service during the regeneratin.

$p_m(t)$- the pdf of the service starting at time t

Directly from this definitions yields

$$p_m(t) = \int_0^\infty \int_0^\infty \int_0^\infty p_m(x,t,Z.\ \upsilon)dxd\upsilon dZ$$

Operation of the considered system can be described by following partial differential equations:

$$(3.1)\qquad [\frac{\partial}{\partial t} + \frac{\partial}{\partial x}]p_m(x,t,Z,V) + p_m(x,t,Z,V)\eta(x) = 0,$$

$$(3.2)\qquad [\frac{\partial}{\partial t} + \frac{\partial}{\partial y} + \eta^{\diamond}(y/V)]q_m(1,y,t,V) = \lambda q_m(0,y,t,V),$$

$$(3.3)\qquad [\frac{\partial}{\partial t} + \frac{\partial}{\partial y} + \eta^{\diamond}(y/V) + \lambda]q_m(0,y,t,V) = 0,$$

where

$$(3.4) \qquad \eta(x) = \frac{f_z(x)}{1 - \int\limits_0^x f_2(u)du}$$

and

$$f_z(x) = \int\limits_0^\infty f_{ZV}(x,V)dV$$

$\eta(x)$ is the conditional pdf of finisching the service if the elapsed service time is equal to x Similarly

$$(3.5) \qquad \eta^\diamond(y/V) = \frac{r(y/V)}{1 - \int\limits_0^y r(u/V)du})$$

is the pdf of regeneration time completion after time y.

Equations (3.1) – (3.3) are to be solved with respect to boundary conditions:

$$(3.6) \qquad p_m(0,t,Z,V) = f_{ZV}(Z,V) \int\limits_0^\infty \int\limits_0^\infty q_m(1,y,t,V)\eta^\diamond(y/V)dydV,$$

$$(3.7) \qquad q_m(0,0,t,V) = \int\limits_0^\infty p_m(Z,t,Z,V)dZ,$$

$$(3.8) \qquad q_m(1,0,t,V) = 0,$$

and initial condition

$$(3.9) \qquad p_m(x,0,Z,V) = \delta(x)f_{ZV}(Z,V).$$

We shall also define the pdf of the length of busy period, which for this case can be found as

$$(3.10) \qquad b_m(t) = \int\limits_0^\infty \int\limits_0^\infty q_m(0,y,tV)\eta^\diamond(y/V)\,dydV.$$

Solving equations (3.1) – (3.3) one can obtain [2],

$$(3.11) \qquad \overline{p}_m(s) = \overline{a}(s) = \frac{1}{1 - \overline{c}(s) + \overline{H}(s+\lambda)},$$

$$(3.12) \qquad \overline{p}_m(x,s,Z,V) = e^{-sx}\, e^{-\int\limits_0^x \eta(u)\,du} \cdot \overline{a}(s) \cdot f_{ZV}(Z,V),$$

$$(3.12) \qquad \overline{b}_n(s) = \overline{a}(s) \cdot \overline{H}(s+\lambda) = \frac{\overline{H}(s+\lambda)}{1 - \overline{c}(s) + \overline{H}(s+\lambda)},$$

where

$$(3.13) \qquad \overline{H}(s+\lambda) \doteq \int\limits_0^\infty f_{ZV}(s,V)\overline{r}\,[(s+\lambda)/V]dV,$$

(3.14) $$\bar{c}(s) = \int_0^\infty \bar{f}_{ZV}(s,V)r(s,V)dV.$$

Let us now discuss the waiting time in such a service scheme. For calculating the waiting time distribution in open service schemes the following reasoning is frequently used:
Due to the Poisson stream of demands, the probability of consecutive call's generation is constant and equal to λdt (where λ is the parameter of the input stream), so one can consider with proper probabilities all possible situations occuring during the service, (eg. idle and busy periods, etc.).

Using the same methods with closed queuing systems, one can obtain so called virtual waiting time (of [8]) reflecting the period which would spend in the queue a fictitious demand, not influancing system's operation, and having the proper priority.

The real waiting time is different from the virtual waiting time, because in the case of one-dimensional sources there is no possibility of generation of the demand z_k if that demand is either in the queue or is being served.

In this paper we shall consider the real waiting time (called shortly "waiting time") exelusively.

Let us notice, that in the service scheme with regeneration the demand starting every busy period is served without waiting, while consecutive service can start only after some waiting time. The mean number of services given during one busy period is equal to

(3.15) $$\int_0^\infty p_m(t) = \bar{p}_m(0) = \frac{1}{\bar{H}(\lambda)}$$

and the mean number of services preluded by waiting in queue is equal to

$$\frac{1}{\bar{H}(\lambda)} - 1 = \frac{1-\bar{H}(\lambda)}{\bar{H}(\lambda)}$$

The queuing phenomenon is caused in this scheme only by existance of the regeneration. We shall calculate the waiting-time distribution, assuming that the conditional pdf $r(y/V)$ is of a specific form, namely

(3.16) $$r(y/V) = r^*(y/V)[1 - \psi(V)] + \delta(y)\psi(V),$$

where $\psi(v)$ is some given function. This assumption will be discussed later. The stochastic variable V has a pdf

$$f_V(V) = \int_0^\infty f_{ZV}(Z,V)dZ = \bar{f}_{Z,V}(0,V).$$

thus if regeneration takes place, than it's length has an unconditional pdf given by (3.17).

(3.17) $$r^{\diamond}(y) = \int_0^\infty r^*(y/V)\cdot f_V(V)dV.$$

The waiting time can be in this case determined as the result of substraction of two stochastic variables, describing the lenght of the regeneration and the generation process respectively, under the condition that a demand will be generated during the regeneration. So the pdf of waiting time is equal to

$$(3.18) \qquad \omega_m^{-I}(\Theta) = \lambda \frac{\overline{r}^{\diamond}(\Theta) - \overline{r}^{\diamond}(\lambda)}{\lambda - \Theta} \cdot \frac{1}{1 - \overline{r}^{\diamond}(\lambda)}$$

where $[1 - \overline{r}^{\diamond}(\lambda)]$ is the probability of demand's generation during the regeneration of a non-zero lenght.

As functions of the shape (3.18) will be often used in further considerations, we shall introduce

$$\overline{F_\lambda}(\varphi,\Theta) = \lambda \frac{\overline{\varphi}(\Theta) - \overline{\varphi}(\lambda)}{\lambda - \Theta} \cdot \frac{1}{1 - \overline{\varphi}(\lambda)}$$

describing the pdf of waiting time caused by some process having the pdf $\varphi(t)$ if the demands are generated according to the negative exponential distribution with parameter λ.

So

$$\overline{\omega}_m^I(\Theta) = \overline{F}_\lambda(\overline{r}^{\diamond}, \Theta).$$

Finally the total pdf of waiting time in the service scheme with regeneration is given by (3.19)

$$(3.19) \qquad \begin{aligned} \overline{\omega}_m(\Theta) &= \frac{\dfrac{1 - \overline{H}(\lambda)}{\overline{H}(\lambda)}}{\dfrac{1}{\overline{H}(\lambda)}} \cdot \lambda \frac{\overline{r}^{\diamond}(\Theta) - \overline{r}^{\diamond}(\lambda)}{\lambda - \Theta} \frac{1}{1 - \overline{r}^{\diamond}(\lambda)} + \frac{1}{\dfrac{1}{\overline{H}(\lambda)}} \cdot 1 = \\ &= [1 - \overline{H}(\lambda)] \cdot \lambda \frac{\overline{r}^{\diamond}(\Theta) - \overline{r}^{\diamond}(\lambda)}{(\lambda - \Theta)[1 - \overline{r}^{\diamond}(\lambda)]} + \overline{H}(\lambda) \cdot 1 = \\ &= [1 - \overline{H}(\lambda)] \cdot \overline{F}_\lambda(r^{\diamond}, \Theta) + \overline{H}(\lambda) \cdot 1. \end{aligned}$$

b.) A scheme with initial process and regeneration

In this scheme at $t = 0$ an initial period having a pdf $\Omega(x)$ starts. If during the Ω no demand is generated the busy period is supposed to complete (a short busy period). On the other hand if during Ω the demand is generated, then after completing the initial period it's service starts, (a long busy period).

As, similarly to the case a.) the service includes some "incorrect"part and causes perhaps a necessity of regeneration, the process may continue as in the previous scheme.

The busy period distribution $b_m^{\Omega}(t)$ can be found using the results from previous scheme,

$$b_m^{\Omega}(t) = e^{-\lambda t} \cdot \Omega(t) + (1 - e^{-\lambda t}) \cdot \Omega(t) * b_m(t),$$

and after transformation

$$(3.20) \quad \bar{b}_m^{\Omega}(s) = \bar{\Omega}(s+\lambda) + [\bar{\Omega}(s) - \bar{\Omega}(s+\lambda)] \cdot \bar{b}_m(s).$$

Similarly one can obtain $p_m^{\Omega}(t)$ (being the analogon of $p_m(t)$), observing, that

$$(3.21) \quad p_m^{\Omega}(t) = (1 - e^{-\lambda t}) \cdot \Omega(t) \cdot p_m(t),$$

which after transformation and utilizing (3.11) yields

$$(3.22) \quad \bar{p}_m^{\Omega}(s) = \frac{\bar{\Omega}(s) - \bar{\Omega}(s+\lambda)}{1 - \bar{c}(s) + \bar{H}(s+\lambda)} = \frac{\bar{\Omega}(s) - \bar{b}_m^{\Omega}(s)}{1 - \bar{c}(s)}$$

$\bar{H}(s+\lambda)$ and $\bar{c}(s)$ used in (3.22) are given by (4.13) and (4.14).

The mean number of services given during a busy period can be found utilizing $p_m^{\Omega}(t)$, as being equal to

$$(3.23) \quad \int_0^{\infty} p_m^{\Omega}(t) = \bar{p}_m^{\Omega}(0) = \frac{1 - \bar{\Omega}(\lambda)}{H} .$$

In the considered scheme every service is preluded by waiting in queue, however queuing before the first service has a different lenght than the consecutive ones. Speaking more precisely, $[1 - \bar{\Omega}(\lambda)]$ services (that means the first one if there is any service at all) come after the waiting time determined by the initial period, thus having the pdf equal to $F_{\lambda}(\Omega,\Theta)$. All remaining services can be fulfilled after waiting time identical as in scheme a.), having the pdf equal to $F_{\lambda}(r,^{\diamond} \Theta)$.

The pdf of total waiting time can be for this scheme experessed by (3.)

$$(3.24) \quad \bar{\omega}_m^{\Omega}(\Theta) = \frac{\dfrac{1 - \bar{\Omega}(\lambda)}{1 - \bar{\Omega}(\lambda)}}{\bar{H}(\lambda)} \bar{F}_{\lambda}(\Omega,0) + \frac{\dfrac{[1 - \bar{H}(\lambda)][1 - \bar{\Omega}(\lambda)]}{\bar{H}(\lambda) \cdot [1 - \bar{\Omega}(\lambda)]}}{\bar{H}(\lambda)} \cdot \bar{F}_{\lambda}(r,^{\diamond} \Theta) =$$

$$= \bar{H}(\lambda) \cdot \bar{F}_{\lambda}(\Omega,\Theta) + [1 - \bar{H}(\lambda)] \cdot \bar{F}_{\lambda}(\bar{r}^{\diamond}, \Theta).$$

Let us notice, that schemes a.) and b.) discussed here are a generalization of simple schemes given in [1].
Results presented there can be obtained from formulas obtained above, by following subtitutions:

$$\bar{r}(s/V) = 1 \qquad \text{or} \qquad f_{ZV}(Z,V) = f_Z(Z) \cdot \delta(V).$$

c.) A scheme with set-up time and regeneration.

In this scheme we shall assume, that first service in each busy period must be preceded by a set-up time having the pdf $\Omega(y)$ begining after the generation of the demand starting that busy period.

Every service can cause the regeneration, exatcly as it was in the previous schames.

It can easily proved that $b_m^{\Omega,1}$ and $p_m^{\Omega,1}$ (describing for that scheme the lenght of busy period, and the begining of a consecutive service are to derived from equations (3.25) and (3.26),

$$(3.26) \qquad \bar{b}_m^{\Omega,1}(s) = \bar{\Omega}(s) = \bar{\Omega}(s) \cdot \bar{b}_m(s)$$

$$(3.27) \qquad \bar{p}_m^{\Omega,1}(s) = \bar{\Omega}(s) \cdot \bar{p}_m(s).$$

where $\bar{p}_m(s)$ and $\bar{b}_m(s)$ are given by (3.11) and (3.12).

The mean number of services given during a single busy period is equal to $\dfrac{1}{\bar{\bar{H}}(\lambda)}$.

In this scheme the first service in every busy period is preluded by a waiting time equal to Ω, while all consective services are given after waiting time having the pdf equal to $\bar{F}_\lambda(r^\diamond,\Theta)$.

Of course $r^\diamond$ is given as previously by (3.17).

Thus, the total pdf of waiting time can be calculated from (3.27).

$$(3.27) \qquad \bar{\omega}_m^{\omega,1}(\Theta) = \frac{1}{\frac{1}{\bar{\bar{H}}(1)}} \cdot \bar{\Omega}(\Theta) + \frac{\frac{1-\bar{H}(\lambda)}{\bar{\bar{H}}(\lambda)}}{\frac{1}{\bar{\bar{H}}(\lambda)}} \cdot \bar{F}_\lambda(r^\diamond,\Theta) = \bar{H}(\lambda) \cdot \bar{\Omega}(\Theta) + [1-\bar{H}(\lambda)] \cdot \bar{F}_\lambda(\bar{r}^\diamond,\Theta)$$

d.) A scheme with initial period, set–up time and regeneration

In this scheme at $t = 0$ the initial period having a pdf $\Omega(y)$ starts and if during this period a demand is generated, the set-up time with pdf $\Omega(x)$ begins after completion of the initial period, and then the service takes place (a long busy period). If during the initial period no demand is generated, busy period will be finished (a short busy period). Service paremeters similar to the pervious ones can be obtained easily also for this scheme:

$$(3.28) \qquad b_m^{\Omega,\Omega}(t) = e^{-\lambda t}\Omega(t) + (1-e^{-\lambda t})\Omega(t) * b_m(t) * \Omega(t),$$

$$(3.29) \qquad p_m^{\Omega,\Omega}(t) = (1-e^{-\lambda t}) \cdot \Omega(t) * p_m(t) * \Omega(t),$$

where $b_m(t)$ and $p_m(t)$ are given by (3.12) and (3.11).

After transformation we obtain (utilizing (3.22)

$$(3.30) \qquad \bar{b}_m^{\Omega,\Omega}(s) = \bar{\Omega}(s+\lambda) + [\bar{\Omega}(s) - \bar{\Omega}(s+\lambda)] \cdot \bar{b}_m(s) \cdot \bar{\Omega}(s),$$

and

$$(3.31) \qquad \bar{p}_m^{\Omega,v}(s) = \bar{\Omega}(s) \cdot \frac{\bar{\Omega}(s) - \bar{\Omega}(s+\lambda)}{1 - \bar{c}(s) + \bar{H}(s+\lambda)} .$$

The mean number of services given during a single busy period is equal to

$$\int_0^\infty p_m^{\Omega,v}(t) = \frac{1 - \bar{\Omega}(\lambda)}{\bar{H}(\lambda)} ,$$

including $[1 - \bar{\Omega}(\lambda)]$ services preluded by the waiting time with pdf equal to $\bar{\Omega}(\Theta) \cdot \bar{F}_\lambda(\Omega,\Theta)$

and $\dfrac{[1 - \bar{H}(\lambda)][1 - \bar{\Omega}(\lambda)]}{\bar{H}(\lambda)}$ services preluded by waiting time with pdf equal to

$\bar{F}_\lambda(r^\diamond, \Theta)$. The justification is for this case identical as in the earlier schemes. Finally, the pdf of total waiting time is given by (3.32)

$$(3.32) \qquad \bar{w}_m^{\Omega,v}(\Theta) = \bar{H}(\lambda) \cdot \bar{\Omega}(\Theta) \cdot \bar{F}_\lambda(\Omega,\Theta) + [1 - \bar{H}(\lambda)] \cdot \bar{F}_\lambda(r^\diamond, \Theta)$$

In furter considerations we shall utilize the results obtained for the basic service schemes in the analysis of our priority system. It will be therefore necessary to specify the obtained formulas for a given type of demands, which will be denoted by an additional subscript.

For example $b_{m,3}^{\Omega}$ will denote the busy period in the scheme b.) after substituting λ_3, S_3, c_3, H. Of course for calculating H_3 and c_3 from formulas (3.1) and (3.1) one has to determine previously the functions $f_{ZV,3}(Z,V)$ and $r_3(y/V)$.

4. The joint busy period end server's utilization factor

After presenting the results describing basic service schemes we shall return to the model introduced in section 2.

We shall define completion time c_k^i of the phase F_k^i as time period between the beginning of execution of F_k^i until the servoce station becoms ready to start service of the same type demand's next phase. Quite similarly can be defined the completion time of the full demand c_k.

As it is well known in the priority queues theory, introducing of the completion time makes it possible to regard the existance of higher priority demands.

We shall assume, for the time being, the full knowledge of pdfs for c_K^i, $i\epsilon(I_K V J_K)$, $K = 1,2, \ldots, N$.
The detailed analysis of c_K^i will be given later in this section. For simplification of further analysis we shall assume, that the service of demand z_K consists of a finite number of phases, equal to k_K, and that the last phase is of nonpreemptive type,*) i.e. $k_K \epsilon J_K$.

Let γ_K denote the joint busy period for demands $z_1, z_2, \ldots, z_K$ γ_K can start in two possible ways:

- either because of the generation of z_K, which service starts immidiately, being however influanced by the higher priority demands. All the preemptive phases F_K^i, $i \epsilon I_K$ may bepreemptived, while the nonpreemtive phases F_K^i, $i \epsilon J_K$ may be followed by the service of higher priority demands. Thus, if the last phase in the service of z_K is nonpreemptive, the demand z_K can be generated once more during the service of higher-priority demands following this phase.
- or because of the generation of any one of demands $z_1, z_2, \ldots, z_{K-1}$ starting the busy period γ_{k-1}, which in turn creates initial period for the process described above.

According to these remarques γ_K can be derived through an application of the basic schemes a.) and b.) in which the completion times of all but last one of the demand's z_K service phases form jointly the "correct" part of the service, while tha last, nonpreemptive phases $F_K^{k_K}$ is considered as the "incorrect" service V.

This incorrect service may couse the necessity of regeneration, i.e. the necessity of serving some hihger priority demands generated during $F_K^{K_k}$, where according to the earlier introduced denotation, k_K is the number of the last, nonpremptive phase. Following this reasoning we can formulate:

$$(4.1) \qquad \bar{\gamma}_K(s) = \frac{\lambda_K}{\Lambda_K} \cdot \bar{b}_{m,K}(s) + \frac{\Lambda_{K-1}}{\Lambda_K} \cdot \bar{b}_m^{\gamma_{K-1}}(s),$$

where $\Lambda_K = \sum_{i=1}^{K} \lambda_i$.

For obtaing $b_{m,K}^{\gamma_{K-1}}$ it is necessary to determine $f_{ZV,K}$ and $r_K(y/V)$ which in this case are of the folowing form:

$$(4.2) \qquad f_{ZV,K}(Z,V) = C_K^*(Z) * \delta(Z, - V) \cdot S_K^{k_K}(V),$$

* Let us notice that if the last phase would not be of nonpreemptive type then we can always add one artifical phase with r_K^i and $s_K^i(x) = \delta(x)$, i being the number of the additional phase.

(4.3) $$\bar{r}_K(s/V) = e^{-\Lambda_{K-1}V} \cdot \delta(y) + \frac{1 - e^{-\Lambda_{K-1}V}}{1 - \Omega^*(\lambda_{K-1})} e^{sV} \cdot$$
$$\cdot [\bar{\gamma}^{\Omega^*}_{K-1}(s) - \bar{\Omega}^*(s + \Lambda_{K-1})],$$

where

(4.4) $$\bar{c}^*_K(s) = \prod_{i=1}^{k_{K-1}} \bar{c}^i_K(s),$$

(4.5) $$\bar{\Omega}^*(s) = e^{-sV} \cdot \bar{Q}_K^{k_K}(s).$$

We can observe that $\bar{r}_K(s/V)$ is of the form suggested in (3.16) where

$$\psi_K(V) = e^{-\Lambda_{K-1}V}$$

and

(4.6) $$\bar{r}^*_K(s/V) = \frac{1}{1 - \bar{\Omega}^*(\Lambda_{K-1})} e^{sV} [\bar{\gamma}^{\Omega^*}_{K-1}(s) - \bar{\Omega}^*(s + \Lambda_{K-1})].$$

We shall now consider completion times for phases of both types a.) completion time for preemptive phases

During the service time S^i_K of the phase F^i_K there can occure n preemptive with probability π_n

$$\pi_n = \frac{(\Lambda_{K-1} \cdot y)^n}{n!} e^{-\Lambda_{K-1}y}$$

We shall assume that every preemption lasts for time x, with pdf equal tp $\Gamma_{K-1}(x)$ Therefore

(4.7) $$c^i_k(x) = \int_0^x \sum_{n=0}^{\infty} e^{-\Lambda_{k-1}y} \frac{(\Lambda_{k-1}y)^{n^*}}{n!} \Gamma^n_{k-1}(x - y) \cdot S^i_k(y)dy,$$

where n^* denotes n time convolution.
After tr insformation

(4.8) $$\bar{c}^i_K(s) = \bar{S}^i_K[s + \Lambda_{K-1} \cdot E(\Gamma^i_{K-1})],$$

We shall now derive $\Gamma_{K-1}(s)$ basing on assumptions acording the overhaed time.

From Fig 1 one see that Γ_{K-1} can be divided into two parts:

– the first one, denotd by g^i_K from generation of the demand, until the overhead R begins for the first time, and

– the second one, denoted by σ^i_K from the instant when R begins for the first time, until it is finished succesfully (that mean without any demand of higher priority being generated during R).

As the two parts are independent, we can present $\overline{\Gamma}_{K-1}(s)$ in the following way

(4.9) $$\Gamma^i_{K-1}(s) = \overline{g}^i_K(s) \cdot \overline{\sigma}^i_K(s).$$

The first factor is equal to a following sum

(4.10) $$\overset{i}{\underset{K}{g}}(s) = \sum_{j=1}^{K-1} \frac{\lambda_j}{\Lambda_{K-1}} \overline{\Phi}^j_{K-1}(s)$$

where

$$\overline{\Phi}^j_l(s) = \begin{cases} \overline{b}_{m,l}^{\;\overline{\Phi}^j_{l-1}(s)}(s) & d/a \quad j < l \leqslant k-1 \\ \overline{b}_{m,j}(s) \cdot \overline{\gamma}_{j-1}^{\;Q^i_K}(s) & d/a \quad l = j \end{cases}$$

(4.11) $$\overline{\gamma}^{\Omega}_l(s) = \overline{b}_{m,l}^{\;\gamma^{\Omega}_{l-1}}(s)$$

Formula (4.10) reflects the order of service, in the case when F^i_K was interrupted by the demand z_j, which occures with probability $\dfrac{\lambda_j}{\Lambda_{K-1}}$

In that case the overhead Q takes place, during which some demands of priority higher than j can be generated, and only after completing Q the service of demands can start in the sequence determined by their priorities.

On the other hand $\overline{\sigma}_K(s)$ can be derived from equation (4.12)

(4.12) $$\overline{\sigma}^i_K(s) = \overline{R}^i_K(s + \Lambda_{K-1}) + [1 - \overline{R}^i_K(\Lambda_{K-1})] \cdot \frac{\overline{\gamma}^{\Omega^*}_{K-1}(s) - \overline{\Omega}^*(s + \Lambda_{K-1})}{1 - \overline{\Omega}(\Lambda_{K-1})} \cdot \sigma^i_K(s)$$

where

$$\overline{\Omega}^*(s) = \overline{R}^i_K(s) \cdot \overline{Q}^i_K(s)$$

Formula (4.12) describes the two possible situations:

either no demand of higher priority than K is generated during R^i_K (this case is given by the first component) or at least one of them is generated, starting a quite complicated process. According to the assumptions about the overhead, R^i_K is in this case followed by Q^i_K, and both these times of overhead create the initial period for the busy period γ_{K-1}, which is a "long" one.

After completing γ^{Ω}_{K-1} (which can be derived from the iterative formula (4.11)) the overhead R^i_K starts again, begining ones more the whole process.

b.) completion time for nonpreemptive phase

For obtaining the completion time c^i_K, $i \in J_K$ we can use the same reasoning as in the case of σ^i_K, thus obtaining

$$(3.13) \quad \underset{i \in J_K}{\bar{c}^i_K(s)} = \bar{S}^i_K(s + \Lambda_{K-1}) + [1 - \bar{S}^i_K(\Lambda_{K-1})] \frac{\bar{\gamma}^{\Omega}_{K-1}(s) - \bar{\Omega}(s + \Lambda_{K-1})}{1 - \bar{\Omega}(\Lambda_{K-1})},$$

Now we are in the position to derive $f_{ZV,K}(Z,V)$ and $r_K(y/V)$ according to formulas (4.2) and (4.3), because all c^i_K for $i \in (I_K V J_K)$, $K = 1,2, \ldots, N$ can be calculated from either (4.8) or (4.13).

We can also obtain $\bar{c}_K(s)$ directly from (3.14) as well as γ_K from (4.1).

So far we have considered the busy periods, exclusively. Th general service process consists of a sequence of busy periods and idle periods and the busy periods starting instant can be treated as regeneration points of a renewal process.

Assuming that at $t = 0$ a busy period starts, the transform of renewal density function $\bar{h}_K(s)$ can be calculated as

$$(4.14) \quad \bar{h}_K(s) = \frac{\Lambda_K \bar{\gamma}_{K(s)}}{s + \Lambda_K[1 - \bar{\gamma}_K(s)]}$$

Let $e_K(t)$ be the probability of system being idle at time t of the general process, with respect to demands $z_1, z_2, \ldots, z_K$.

Using the renewal argument it is easy to obtain

$$(4.15) \quad \bar{\hat{e}}_K(s) = \frac{1}{s + \Lambda_K[1 - \bar{\gamma}_K(s)]},$$

Let us notice that $[1 - \hat{e}_K(t)]$ determines the ulilization factor of the server.

For $t \to \infty$ the stationary state probability of the server being idle is

$$(4.16) \quad \hat{e}_K = \frac{1}{1 + \Lambda_K E(\gamma_K)}.$$

Basing on (4.16) one can easily find the fraction of server's time spend on dealing with demands z_K $(K = 1,2, \ldots, N)$, that means needed for serving this demand and for the overhead connected with it's preemption and resuming of the service.

This factor is equal to $\hat{e}_{K-1} - \hat{e}_K$ and is influanced by all the demand's parameters and also by the service discipline.

5. Waiting time distributions

The discussion of waiting time distribution will be presented for the stationary state, and propen formulas will be obtained by a detailed analysis of service processes occuring the busy period.

For estimating $W_K^{(N)}(\tau)$ the pdf of waiting time of the demand z_K in a system serving demands $z_1, z_2, \ldots, z_K, \ldots z_N$, we shall first investigate the number of services n_K given to demand z_K during a single busy period γ_N.
We shall also consider waiting times preluding respective services. $n_K^{(N)}$ will be derived using an iterative formula

$$(5.1) \qquad n_K^{(k)} = \frac{\Lambda_{k-1}}{\Lambda_k} \cdot n_k^{(k-1)} + M_k N_K^k \qquad \begin{matrix} k = 1,2,\ldots,N \\ K = 1,2,\ldots,k \end{matrix}$$

where M_k denotes the number of demands z_k served during a busy period γ_K, and can be found equation (5.2)

$$(5.2) \qquad n_k^{(k)} = M_k = \frac{\lambda_k}{\Lambda_k} \cdot \overline{p}_{m,k}(0) + \frac{\Lambda_{k-1}}{\Lambda_k} \overline{p}_{m,k}^{\gamma_{k-1}}(0),$$

after a reasoning similar to those which justified (4.1).

Using $N - K$ times (5.1) we obtain

$$(5.3) \qquad n_K^{(N)} = \frac{\Lambda_K}{\Lambda_N} n_K^{(k)} + \sum_{L=K+1}^{N} \frac{\Lambda_L}{\Lambda_N} M_L N_K^L$$

Both in (5.1) and (5.3) N_K^L denotes a mean number of demands z_K served during the completion time c_K, and has to be calculated with respect to the multiphase structure of demand's z_L service.
Thus

$$(54) \qquad N_K^L = \sum_{i \in (I_L \vee J_L)} r_L^i \cdot N_K^{L,i}$$

where $N_K^{L,L}$ denotes a mean number of demands z_K served during the completion time c_L^i.
Let us notice, that generation of demand z_K is possible only if that demand is neither served nor waiting in the queue.
Thus the generation of K type demand during c_L^i can happen in the instant when

- the service of phase F_L^i is in progress
- overhead Q_L^i or R_L^i is in progress
- the busy period γ_{K-1} is in progress
- the completion time of an type 1 demand $(l = K+1,\ K+2, \ldots L-1)$ is in progress

Denoting by $m_K^{L,i}$ the number of services given to demands z_K generated in the first three of obove listed situations leads to the following formula describing

$$(5.5)\qquad N_K^{L,i} = m_K^{L,i} + \sum_{l=K+1}^{L-1} m_l^{L,i}\ N_K^l$$

We shall now given the formulas for calculating $m_K^{L,i}$ seperately for preemptive and non-preemptive phases, providing them only with short comments, while the full eduction is given in [2].

a.) preemptive phases

For F_L^j, $j\epsilon L_L$ we obtain

$$m_K^{L,j} = \frac{1}{r_L^j}\,\nu_L^j\,\{\lambda_K \bar{p}_{m,K}(0) + \Lambda_{k-1}\bar{p}_{m,K}^{A}(0) +$$

$$(5.6)\qquad + (\Lambda_{L-1} - \Lambda_K)\bar{p}_{m,K}^{\gamma_{K-1}^{Q_L^j}}(0)\} + \frac{\nu_L^j}{r_L^j}\frac{\Lambda_{L-1}}{\bar{R}_L^j(\Lambda_{L-1})}\{[1 - \bar{R}_L^j(\Lambda_{K-1})]\bar{p}_{m,K}^{\Omega^{\mathrm{I}}}(0)$$

$$+ [\bar{R}_L^j(\Lambda_{K-1}) - \bar{R}_L^j(\Lambda_{L-1} - \lambda_K)]\bar{p}_{m,K}^{\Omega^{\mathrm{II}}}(0) +$$

$$+ \bar{R}_L^j(\Lambda_{K-1} - \lambda_K)\bar{p}_{m,K}^{\Omega^{\mathrm{III}}}(0)\}$$

where

$$A = \sum_{i=1}^{K-1}\frac{\lambda_i}{\Lambda_{K-1}}\cdot\bar{\Phi}_{K-1}^i(s), \qquad \nu_L^j = E(S_L^j)$$

$$\bar{\Phi}_l^i(s) = \begin{cases} \bar{b}_{m,l}^{\varphi_{L-1}^i}(s) & \text{for } i < l \leqslant K-1 \\ \bar{b}_{m,i}(s)\cdot\bar{\gamma}_{i-1}(s) & \text{for } l = i \end{cases}$$

$$\bar{\Omega}^{\mathrm{I}}(s) = \frac{\bar{\gamma}_{K-1}^{\Omega^*}(s) - \bar{\Omega}^*(s + \Lambda_{K-1})}{[1 - \bar{\Omega}^*(\Lambda_{K-1})]} \qquad \text{for } \bar{\Omega}^*(s) = \bar{R}_L^j(s)\cdot\bar{Q}_L^j(s)$$

$$\bar{\Omega}^{\mathrm{II}}(s) = \frac{\bar{R}_L^j(s + \Lambda_{K-1}) - \bar{R}_L^j(\Lambda_{L-1} - \lambda_K + s)}{\bar{R}_L^j(\Lambda_{K-1}) - \bar{R}_L^j(\Lambda_{L-1} - \lambda_K)}\cdot\bar{\gamma}_{K-1}(s)$$

$$\bar{\Omega}^{\mathrm{III}}(s) = \frac{\bar{R}_L^j(\Lambda_{L-1} - \lambda_K + s)}{\bar{R}_L^j(\Lambda_{L-1} - \lambda_K)}$$

The right side of (5.6) consitss of six components liked in two groups. In the first group there are considered the cases when z_K is generated during

- the service of F_L^j
- the busy period caused by generation of the demands $z_1, z_2, \ldots, z_K$
- the overhead Q_L^j caused by the generation of demands $z_{K+1}, z_{K+2}, \ldots, z_{L-1}$ or during the service of any demand having the priority higher than K which could be generated during Q_L^j.

The second group contains the possibilities of z_K generation during following processes, connected with tha part σ_L^j of preemption (cf section 4), namely the possibilities that:

- during R_L^j at least one demand of priority higher than K is generated, causing the overhead Q_L^j and a proper busy period
- during R_L^j no demand of priority higher than K, and at least one demand of priority between K and L is generated.
- during R_L^j no demand of priority different than K is generated.

For evaluating the number of services given to z_K in any of this situations the basic schemes from section 3 where used. In fact, for the first situation the basic scheme c.) is a proper one for the last situation we have to utilize the scheme d.), while all other can be described applying the basic scheme b.) with due substitutions.

Discussing the basic schemes we have pointed out, the way of obtaining mean number of services given in every scheme, as well as possible waiting time distributions.
Using the derived in section 3 formulas with substitutions adapting them to all listed above situations we can observe, that the waiting time can have one of the seven possible density functions, given in (5.7).

$$(5.7)\quad \begin{cases} \overline{U}^{L,j}_{K,1}(\Theta) = \overline{\gamma}^{Q^j_L}_{K-1}(\Theta) \\ \overline{U}^{L,j}_{K,2}(\Theta) = \overline{F}_{\lambda_K}(A,\Theta) \\ \overline{U}^{L,j}_{K,3}(\Theta) = \overline{F}_{\lambda_K}(\gamma^{Q^j_L}_{K-1},\Theta) \\ \overline{U}^{L,j}_{K,4}(\Theta) = \overline{F}_{\lambda_K}(\Omega^{I},\Theta) \\ \overline{U}^{L,j}_{K,5}(\Theta) = \overline{F}_{\lambda_K}(\Omega^{II},\Theta) \\ \overline{U}^{L,j}_{K,6}(\Theta) = \overline{\gamma}^{Q^j_L}_{K-1}(\Theta) \cdot \overline{F}_{\lambda_K}(\Omega^{III},\Theta) \\ \overline{U}^{L,j}_{K,7}(\Theta) = \overline{F}_{\lambda_K}(r^{\diamond},\Theta) \end{cases}$$

Let us notice that Ω^{I}, Ω^{II} and Ω^{III} where defined in (5.6) while $r^{\diamond}$ was given ba (3.17) and (4.6).

Functions $U^{L,j}_{K,i}$, $j\epsilon I_L$, $i\epsilon[1,2,\ldots,6]$ describe the waiting times connected with the first service in everyone of the six considered situations, and $U^{L,j}_{K,7}$ gives the waiting time preluding the services caused by regeneration

Let us denote the number of services preluded with waiting time having the equal to $U^{L,j}_{K,i}$ by $[m^{L,j}_{K}]_i$.

Introducing adequate substitutions to the formulas given in section 3 it is easy to show, that

$$
(5.8)\quad
\begin{cases}
[m_K^{L,j}]_1 = \dfrac{\nu_L^j}{r_L^j}\,\lambda_K \\[2ex]
[m_K^{L,j}]_2 = \dfrac{\nu_L^j}{r_L^j}\,\Lambda_{K-1}[1-\overline{A}(\lambda_K)] \\[2ex]
[m_K^{L,j}]_3 = \dfrac{\nu_L^j}{r_L^j}\,(\Lambda_{L-1}-\Lambda_K)[1-\overline{\gamma}_{K-1}^{Q_L^j}(\lambda_K)] \\[2ex]
[m_K^{L,j}]_4 = \dfrac{\Lambda_{L-1}\nu_L^j}{r_L^j\cdot \overline{R}_L^j(\Lambda_{L-1})}\,[1-\overline{R}_L^j(\Lambda_{K-1})][1-\overline{\Omega}^{\mathrm{I}}(\lambda_K)] \\[2ex]
[m_K^{L,j}]_5 = \dfrac{\Lambda_{L-1}\nu_L^j}{r_L^j\,\overline{R}_L^j(\Lambda_{L-1})}\,\{[\overline{R}_L^j(\Lambda_{K-1})]-[\overline{R}_L^j(\Lambda_{K-1})]\}[1-\overline{\Omega}^{\mathrm{II}}(\lambda_K)] \\[2ex]
[m_K^{L,j}]_6 = \dfrac{\Lambda_{L-1}\nu_L^j}{r_L^j\,\overline{R}_L^j(\Lambda_{L-1})}\,\overline{R}_L(\Lambda_{K-1}-\lambda_K)[1-\overline{\Omega}^{\mathrm{III}}(\lambda_K)] \\[2ex]
[m_K^{L,j}]_7 = \dfrac{\nu_L^j}{r_L^j}\,\dfrac{1-\overline{H}_K(\lambda_K)}{\overline{H}_K(\lambda_K)}\,\{\lambda_K+\Lambda_{K-1}[1-\overline{A}(\lambda_K)]+ \\
\qquad +(\Lambda_{L-1}-\Lambda_K)[1-\overline{\gamma}_{K-1}^{Q_L^j}(\lambda_K)]\}+ \\
\qquad +\dfrac{\Lambda_{L-1}\cdot\nu_L^j}{r_L^j\cdot\overline{R}_L^j(\Lambda_{L-1})}\cdot\dfrac{1-\overline{H}_K(\lambda_K)}{\overline{H}_K(\lambda_K)}\,\{[1-\overline{R}_L^j(\Lambda_{K-1})][1-\overline{\Omega}^{\mathrm{I}}(\lambda_K)]+ \\
\qquad +[\overline{R}_L^j(\Lambda_{K-1})-\overline{R}_L^j(\Lambda_{L-1}-\lambda_K)][1-\overline{\Omega}^{\mathrm{II}}(\lambda_K)]+ \\
\qquad +\overline{R}_L^j(\Lambda_{K-1}-\lambda_K)[1-\overline{\Omega}^{\mathrm{III}}(\lambda_K)]\}
\end{cases}
$$

Obviously

$$m_K^{L,j} = \sum_{i=1}^{7} [m_K^{L,j}]_i \qquad j\epsilon I_L$$

b.) nonpremptive phases

The reasoning for the case of nonpreemptive phases F_K^j, $j\epsilon J_L$ is quite similar to that applied to the previous case, in the part concerning resuming the preempted service, i. e. the part σ_K^j of the preemption. Thus introducing S_k^j instead of R_L^j in the last three situations investigated during creation of (5.6) we obtain

$$(5.9) \qquad m_K^{L,j} = [1 - \bar{s}_L^j(\Lambda_{K-1})]\bar{p}_{m,k}^{\Omega^{IV}}(0) + [\bar{s}_L^j(\Lambda_{K-1}) - \bar{s}_L^j(\Lambda_{L-1} - \lambda_K)]\bar{p}_{m,K}^{\Omega^{IV}}(0) + \\ + \bar{s}_L^j(\Lambda_{K-1} - \lambda_k)\,\bar{p}_{m,K}^{\Omega^{III}}(0).$$

where

$$\bar{\Omega}^{IV}(s) = \frac{\bar{\gamma}_{K-1}^{\Omega^*}(s) - \bar{\Omega}^*(s + \Lambda_{K-1})}{1 - \bar{\Omega}^*(\Lambda_{K-1})} \qquad \text{for } \bar{\Omega}^*(s) = \bar{s}_L^j(s) \cdot \bar{Q}_L^j(s)$$

$$\bar{\Omega}^{V}(s) = \frac{\bar{s}_L^j(s + \Lambda_{K-1}) - \bar{s}_L^j(s + \Lambda_{L-1} - \lambda_K)}{\bar{s}_L^j(\Lambda_{K-1}) - \bar{s}_L^j(\Lambda_{L-1} - \lambda_K)} \quad \bar{\gamma}_{K-1}^{q_j^L}(s)$$

$$\bar{\Omega}^{VI}(s) = \frac{\bar{s}_L^j(\Lambda_{L-1} - \lambda_K + s)}{\bar{s}_L^j(\Lambda_{L-1} - \lambda_K)}$$

In the similar way we can list also the pdf describing all possible waiting times, $U_{K,i}^{L,j}$ and mean numbers of services preluded by the waiting time with respective

pdf $- [m_k^{L,j}]_i$ $\quad i\epsilon \{1,2,3,4\}$, $j\epsilon J_k$, $K = 1,2,\ldots,N$.

$$(5.10) \quad \begin{cases} \bar{U}_{K,1}^{L,j}(\Theta) = \bar{F}_{\lambda_K}(\Omega^{IV},\Theta) \\ \bar{U}_{K,2}^{L,j}(\Theta) = \bar{F}_{\lambda_K}(\Omega^{V}.\Theta) \\ \bar{U}_{K,3}^{L,j}(\Theta) = \bar{F}_{\lambda_K}(\Omega^{VI}\Theta) \cdot \gamma_{K-1}^{\Theta_L^j}(\Theta) \\ \bar{U}_{K,4}^{L,j}(\Theta) = \bar{F}_{\lambda_K}(r_k^{\diamond},\Theta) \end{cases}$$

$$(5.11)\quad \begin{cases} [m_K^{L,j}]_1 = [1 - \bar{s}_i^{\,j} \cdot (\Lambda_{k-1})][1 - \bar{\Omega}^{IV}(\lambda_K)] \\ [m_K^{L,j}]_2 = [\bar{s}_L^{\,j}(\Lambda_{K-1}) - \bar{s}_L^{\,j}(\Lambda_{L-1} - \lambda_k)][1 - \bar{\Omega}^{V}(\lambda_K)] \\ [m_K^{L,j}]_3 = \bar{s}_L^{\,j}(\Lambda_{K-1} - \lambda_K)[1 - \bar{\Omega}^{VI}(\lambda_k)] \\ [m_K^{L,j}]_4 = \dfrac{1 - \bar{H}_K(\lambda_K)}{\bar{H}_K(\lambda_K)} \{[m_K^{L,j}]_1 + [m_K^{L,j}]_2 + [m_K^{L,j}]_2 + [m_K^{L,j}]_3\} \end{cases}$$

Surely enough

$$m_K^{L,j} = \sum_{i=1}^{4} [m_K^{L,j}]_i \qquad j \epsilon J_L$$

After obtaining formulas 5.6. and 5.9 we are in position to calculate n_K^N – the mean number of demands z_K served during a busy period γ_N.

We have also discussed the waiting times preluding services of demand z_K given during the completion time of demand z_L, $L = K + 1, K + 2, \ldots, N - 1$. For obtaining the total pdf of waiting time we have to investigate additionally waiting times preluding the service of z_K during the busy period γ_K, that means preluding the $n_K^{(K)}$ services during in (5.2). One can observe that the considered process can be described either by the basic scheme a.) with probability $\frac{\lambda_K}{\Lambda_K}$ or by the basic scheme b.) with probability $\frac{\Lambda_{K-1}}{\Lambda_K}$.

Using formulas derived for those basic scheme it is easy to find that the waiting times may have following pdfs:

$\bar{U}_{K,1}^{(K)}(\Theta) = 1$	– if the demand z_K started the busy period γ_K
$\bar{U}_{K,2}^{(K)}(\Theta) = \bar{F}_{\lambda_K}(\gamma_{K-1}, \Theta)$	– if γ_K has been started by the demand of priority higher than K, and demand z_K is generated during the so initiated γ_K
$\bar{U}_{K,3}^{(K)}(\Theta) = \bar{F}_{\lambda_K}(r_K^{\diamond}, \Theta)$	– in the case of second and further generations of demand z_K during the busy period γ_K (regenerations)

Denoting by $[n_K^{(K)}]_i$ the number of services preluded by waiting time with pdf equal to $U_{K,i}^{(K)}$ we obtain

$$[n_K^{(K)}]_1 = \frac{\lambda_K}{\Lambda_K}$$

$$[n_K^{(K)}]_2 = \frac{\Lambda_{K-1}}{\Lambda_k}\,[1 - \bar{\gamma}_{K-1}(\lambda_K)]$$

$$[n_K^{(K)}]_3 = \frac{1 - H_K(\lambda_K)}{\bar{H}_K(\lambda_K)} \cdot \frac{\lambda_K}{\Lambda_K} + \frac{\Lambda_{K-1}}{\Lambda_K}[1 - \bar{\gamma}_{K-1}(\lambda_K)]$$

Obviously

$$\eta_K^{(K)} = \sum_{i=1}^{3} [\eta_K^{(K)}]_i$$

Finally we can derive the full formula for waiting time distribution considering with respective probabilities all discussed so for possible waiting time distributions.
Thus the pdf of waiting time $W_K^{(N)}(\tau)$ has a Laplace transform equal to

$$(5.13) \quad \overline{W}_K^{(N)}(\Theta) = \frac{1}{\eta_K^{(N)}} \; \frac{\Lambda_K}{\Lambda_N} \cdot \sum_{j=1}^{3} [\eta_K^{(K)}]_j \cdot U_{K,j}^K + \sum_{L=K+1}^{N} \frac{\Lambda_L}{\Lambda_N} M_L \cdot \omega_K^L$$

$$\omega_K^L = \sum_{i\epsilon(I_L V J_L)} r_L^i \cdot \omega_K^{L,i}$$

$$\omega_K^{L,i} = \nu_K^{L,i} + \sum_{l=K+1}^{L-1} \eta_L^{L,i}\, \omega_K^l$$

$$\nu_K^{L,i} = \sum_{j=1}^{7} [m_K^{L,i}]_j \cdot U_{K,j}^{L,i} \qquad d/a \quad i\epsilon I_l$$

$$\nu_K^{L,i} = \sum_{j=1}^{4} [m_K^{L,i}]_j \cdot U_{K,j}^{L,i} \qquad d/a \quad i\epsilon J_L$$

6. Final remaremarks

In this paper the analysis of a complex priority queuing system with one dimensional sources of demands has been presented. Using the reported results it is possible to investigate a wide range of priority algorithms as the considered model includes as special cases not only such simple disciplines as preepmtive resume and head of the line discipline (with or without overhead).

Basing on this model one can also determine easily theservice parameters of the discretionary priority discipline, as well as investigate priority systems where service is given in quanta os a predetermined value. Such priority systems with time slicing (and properly chosen quanta for every priority level) make it often possible to quarantee desired response times for demands of various types with overhead smaller than obtained in the case of other disciplines.

The special cases of the considered model can be obtained by proper defining of

$$r^i_K,\ S^i_K,\ I_K \text{ and } J_K \text{ for } K = 1,2,\ldots,N.$$

The comparison of various scheduling disciplines for the case of one dimensional sources, basing on the considered model will be presented in a seperate paper, as will as the losses of computer's throughput due to overhead.

Finally let us notice that the analisy was done under quite general assupmtions, allowing arbitrary distributions of service times, and overhead. The results were obtained in the form of Laplace transforms of probability distributions and not only in terms of mean values.

References

[1] A. Wolisz, "Real Time Operating Systems Probabilistic Models" Proceedings of the First Winter School on Mathematical Problems of Operating Systems, Közlemények 15/1975, 55-71.

[2] "Foundations of selection and analysis of plant-oriented operating systems", Vol. I. Single processor plant-oriented operating systems. A collective study. Research report, Department of Complex Automation systems, Polisch Academy of Sciences, Gliwice, December 1975 (in Polish).

[3] L. Schrage "A mixed – priority queue with applications to the analysis of real – time systems" Operations Research, Vol 17 No 4 pp. 728-741.

[4] A. Wolisz, "Modeling real–time operating systems as a single server multiphase queuing systems "Podstawy Sterowania Vol VI., No 2 (in Polish)

[5] A. Wolisz "A single server priority queuing model of real – time operating systems," Preprints of the International Conference "Informatica 75", paper No 1.14

[6] J.P.Buzen, P.S. Goldberg "Guidelines for the use of infinite source queuing models in the analysis of computer sytem performance", SJCC 1974, AFIPS Vol 43 pp 771-774.

[7] J. Tomkó "Some queueing models in the mathematical destription of computer systems" Közlemények 15/1975 pp 55-71

[8] N.K. Jaiswal "Priority Queues", Academic Press New York 1968.

Összefoglaló

"Real time" prioritásos rendszerek vizsgálata

Adam Wolisz

A dolgozat a prioritásos rendszer valószinüség elméleti vizsgálatával foglalkozik.

Р Е З Ю М Е

Пробный анализ приоритетных алгоритмов диспетчеризации задач при работе в реальном масштабе времени.

Воллиш А.

В статье предлагается некоторая одноканальная система приоритетного обслуживания с постоянными приоритетами как модель операционной системы вычислительной машины.

В этой модели обслуживание любой заявки происходит в фазах, которые могут подвергаться дисциплине с абсолютным или относительным приоритетом.

Предлагается, что заявки на обслуживание генерированные одномерными источниками заявок.

Исследуется время потраченное на прерывание текущей программы,а также на возврат в прерванную программу.

Принимая, что любую фазу описывает вероятность ее выполнения и любое расположение времени выполнениея, получаем параметры очередей в случае пуассоновских потоков заявок.