6.2.4. Характеристики СМО с приоритетами

Реальные системы не всегда используют дисциплину обслуживания *FIFO* (первая поступившая заявка обслуживается первой). Особенно часто встречающейся дисциплиной является приоритетная, в которой заявки, имеющие приоритет, обслуживаются вне очереди рис.(6.2).

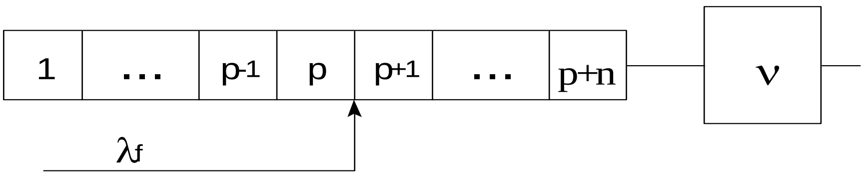

Рис. 6.2. СМО с приоритетным обслуживанием

Если заявка, имеющая приоритет, немедленно прерывает обслуживание другой заявки, то система работает с абсолютным приоритетом; если же заявка дожидается окончания обслуживания ранее поступившей, то система работает с относительным приоритетом.

Рассмотрим способ определения среднего времени ожидания для систем с относительным приоритетом.

Предположим, что поступающие заявки принадлежат одному из F различных приоритетных классов, обозначаемых через $f(f=1,2,\dots,F)$. Будем считать, что чем больше значение индекса класса, тем выше приоритет этого класса. Таким образом, заявкам из приоритетного класса f представляется преимущественное обслуживание по сравнению с заявками из приоритетного класса $f-1$. Если в СМО находится несколько заявок с одинаковым приоритетом, то их обслуживание производится в соответствии с дисциплиной *FIFO*.

Рассмотрим модель, в которой заявки из приоритетного класса f образуют пуассоновский поток с интенсивностью λ_f, а время обслуживания v_f распределено по экспоненциальному закону.

Определим параметры модели:

$$\lambda = \sum_{f=1}^{F} \lambda_f ;$$

$$v = \sum_{f=1}^{F} \frac{\lambda_f}{\lambda} v_f ;$$

$$p_f = \lambda_f v_f ;$$

$$\rho = \lambda v = \sum_{f=1}^{F} \rho_f \ .$$

Здесь под ρ_f подразумевается загрузка обслуживающего прибора заявки из приоритетного класса f; ρ - общая загрузка обслуживающего прибора, причём для обеспечения стационарного режима необходимо выполнение условия $\rho<1$.

Введём следующие обозначения: ω_p - среднее время ожидания заявки из приоритетного класса f; u_f - среднее время пребывания в СМО заявки, поступившей из приобретённого класса f.

Время ожидания разлагается на три составляющие:

1) время, связанное с тем, что в момент поступления данной заявки прибор был занят обслуживанием другой заявки;

2) время, связанное с тем, что в момент поступления данной заявки в очереди уже находятся заявки с приоритетом, равным или более высоким, чем у поступившей;

3) время, обусловленное тем, что заявки, поступившие позже, имеют более высокий приоритет.

Исследование приоритетных систем обычно начинают с поступившей вновь заявки из приоритетного класса *f*.

Будем называть эту заявку меченой. Обозначим через ω_0 среднюю задержку меченной заявки, связанную с наличием другой заявки на обслуживание; ω_0 будем определять через загрузку прибора ρ_i, которая представляет собой вероятность занятия обслуживающего прибора заявками из *i* -го класса. ρ_i также может быть интерпретирована как доля времени, в течение которого прибор занят заявками из *i* -го класса, тогда.

$$\omega_o = \sum_{i=1}^{F} \rho_i \, . \tag{6.22}$$

Рассмотрим теперь вторую составляющую времени ожидания, связанную с тем, что перед меченой заявкой обслуживаются заявки с равным или более высоким приоритетом, которые меченая заявка застала в очереди. Введём обозначение l_{if} - число заявок из класса *i*, которые застала в очереди меченая заявка из класса *f*. В соответствии с формулой Литтла $l_{if} = \lambda_i \omega_i$, где *i=f,f+1,…,F*. Отсюда средняя задержка меченой заявки составит

$$\sum_{i=f}^{F} v_i \lambda_i \omega_i \; .$$

Аналогично можно определить третью составляющую среднего времени ожидания (задержка меченой заявки за счёт того, что заявки, поступающие после неё, имеют более высокий приоритет). Пусть m_{if} - число заявок из класса *i*, поступающих в СМО, когда меченая заявка (из класса *f*) находится в очереди, и получающих обслуживание раньше меченой заявки.

Тогда $m_{if} = \lambda_{if} \omega_{if}$, где *i=f+1, f+2,…,F*, а средняя задержка составит

$$\sum_{i=f+1}^{P} \lambda_i \omega_f .$$

Таким образом, для СМО с относительным приоритетом

$$\omega_f = \omega_o + \sum_{i=f}^{F} v_i \lambda_i \omega_i + \sum_{i=f+1}^{F} v_i \lambda_i \omega_f . \qquad (6.23)$$

Так как $v_i \lambda_i = \rho_i$, то, преобразовав выражение относительно ω_f, получим

$$\omega_f = \frac{\omega_o + \sum_{i=f+1}^{F} v_i \lambda_i \omega_f}{1 - \sum_{i=f}^{F} \rho_i} \quad , \qquad (6.24)$$

где $f=1,2,\ldots,F$

Система уравнений (6.24) может быть легко решена рекуррентно, т.е. сначала находится ω_f, затем ω_{f-1} и т.д.

Например, имеется три класса приоритетов, причём наибольший приоритет имеет класс 3. Тогда среднее время задержки каждого класса определится так:

$$\omega_3 = \frac{\omega_o}{1 - \rho_3};$$

$$\omega_2 = \frac{\omega_o + \rho_3 \omega_3}{1 - \rho_2 - \rho_3};$$

$$\omega_1 = \frac{\omega_o + \rho_2 \omega_2 + \rho_3 \omega_3}{1 - \rho_1 - \rho_2 - \rho_3}.$$