УДК 519.872; 519.876.5

РАСЧЕТ МНОГОУРОВНЕВОЙ СИСТЕМЫ ОЧЕРЕДЕЙ

Ю. И. Рыжиков,
доктор техн. наук, профессор
Военно-космическая академия им. А. Ф. Можайского

Излагаются соображения в пользу коллективного использования вычислительных ресурсов. Предложен основанный на законе сохранения объема работы метод расчета среднего времени пребывания заявок в одноканальной многоуровневой системе с квантованным обслуживанием в зависимости от заявленного времени счета. Длительности квантов и величина системных потерь предполагаются фиксированными и переменными по уровням.

The arguments are discussed for collective using of computer resources. An algorithm is proposed to compute mean sojourn time depending on requested computing time. The system is supposed to be one-channel and multi-level with quantified servicing, quanta and system losses being fixed and varying by levels.

Одной из ведущих тенденций в современной информатике является коллективное использование имеющихся вычислительных ресурсов. В сравнении с монопольным доступом такой подход дает следующие преимущества:

1. Клиент системы платит лишь за фактически использованные ресурсы.

2. В связи с относительным удешевлением единицы продукции на более мощных установках эти ресурсы обходятся дешевле (согласно закону Гроша стоимость вычислительной установки растет пропорционально квадратному корню из ее производительности).

3. Улучшаются показатели обслуживания за счет масштабного эффекта (n-кратное увеличение интенсивности потока заявок и обслуживания во столько же раз уменьшает среднее время ожидания заявки – этот неожиданный факт элементарно доказывается для системы M/M/1, но имеет место и в более общих ситуациях).

4. Уменьшается доля необходимых страховых ресурсов мощностей.

5. На более производительных установках поднимается потолок пиковых потребностей и появляется возможность оказывать качественно новые виды услуг.

Идея коллективного использования ресурсов по фактической потребности при сохранении психологических преимуществ индивидуального доступа еще в 1960-х гг. нашла свое отражение в операционных системах мощных стационарных ЭВМ (mainframes). В наши дни те же проблемы возникают вновь на более высоком уровне при использовании суперкомпьютеров (М-1000 в ВЦ РАН); при организации распределенных (кластерных) вычислений; при работе в локальных сетях с «тонкими» клиентами, переадресующими все серьезные вычисления серверу; при реализации оперативных задач на фоне трудоемких массовых вычислений (например, в вузовских ВЦ); при работе сетевых маршрутизаторов и т. д.

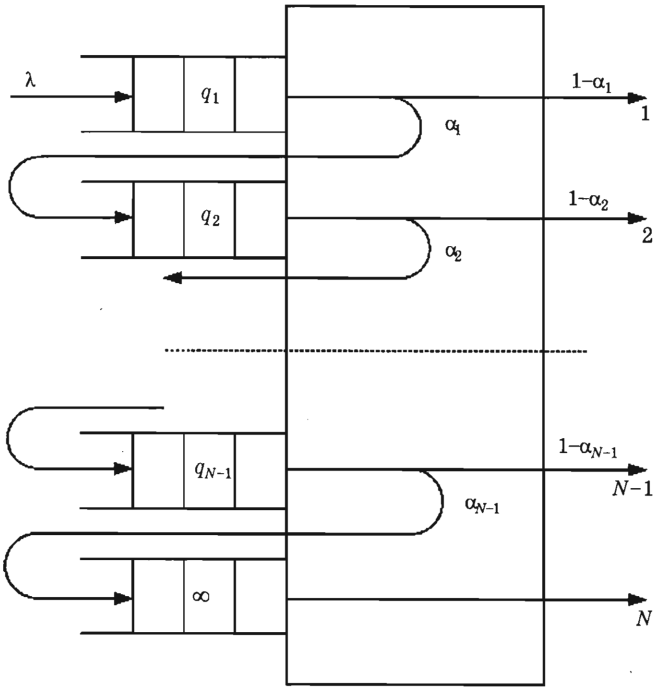
Ключевым элементом технологии коллективного использования процессора является *квантованное* обслуживание, при котором каждая из находящихся в системе активных задач поочередно получает «квант» времени процессора. В зависимости от порядка его предоставления рассматриваются различные дисциплины обслуживания:

циклические RR (Round Robin – «карусель»);

многоуровневые с N очередями разного приоритета FB_N (foreground-background, т. е. с передним и задним планом);

уравнительное разделение процессора EPS (egalitarian processor sharing).

Важнейшим показателем эффективности математической эксплуатации вычислительной системы коллективного доступа является количество задач, решенных на ней в единицу времени. Этот показатель будет максимален, если предоставить статический приоритет коротким заявкам. Однако в типичных условиях случайной трудоемкости эта стратегия физически не реализуема, поскольку длины заявок априорно не известны. Все перечисленные выше дисциплины *автоматически* обеспечивают приоритет корот-



■ *Расчет многоуровневой системы*

ким заявкам и к тому же по фактической, а не по ожидаемой длительности. Естественно, что делается это за счет длинных заявок и ценой дополнительных системных издержек на прерывания и обмены. Поэтому при анализе упомянутых дисциплин показатели обслуживания определяются в зависимости от длины заявки t. Основным показателем является среднее время ν_t ее пребывания в системе или среднее время ν_k пребывания в системе заявки, завершающей обслуживание на k-м кванте.

Затраты τ времени работы процессора на каждое переключение увеличивают общую загрузку системы циклического обслуживания и особенно сильно сказываются на прохождении «длинных», много раз прерываемых задач. Многоуровневая система [1–4, 6, 7] типа показанной на рисунке дает наибольшие преимущества коротким заявкам и к тому же уменьшает частоту переключений. Эта дисциплина в работе [7] называется Shortest Elapsed Time (SPT). В такой системе заявка, не завершившая обслуживание на i-м уровне в течение кванта q_i, переходит на следующий уровень и получает на нем квант $q_{i+1} > q_i$. Нижний уровень N предоставляет обслуживание до его завершения. Обслуживание заявок из разных очередей производится по схеме относительного приоритета. Переход к каждой новой заявке требует времени процессора τ независимо от уровня обслуживания.

Во всех перечисленных источниках поставленная задача решается при существенных ограничениях. В работах [1, 3] общая длительность обслуживания предполагается показательно распределенной, в работе [3] то же допущение делается относительно длительности кванта; в работах [1, 7] кванты считаются постоянными и не зависящими от номера уровня, а в исследовании [2] – бесконечно малыми. Описываемые методы, как правило, неконструктивны: к примеру, в работе [2] предлагается решать функциональное уравнение в преобразованиях Лапласа, а в работе [6] бесконечную систему линейных уравнений, причем среднее время ожидания на j-м уровне зависит от таковых на последующих уровнях (это явная ошибка). Нигде не учитываются играющие принципиальную роль потери на переключения и не приводится верификация расчетных методик.

Опишем свободный от указанных недостатков метод расчета среднего времени ν_k пребывания в одноканальной системе заявки, обслуживание которой завершается на k-м уровне, $k=1,\ N$. Для всех уровней введем понятия:

$Q_k=\sum_{i=1}^{k} q_{i,1}$ – интегральный квант;

$L_k=\sum_{i=1}^{k} \tau_{i,1}$ – средние интегральные потери до k-го уровня включительно;

$\alpha_k=\overline{B}(Q_k)$ – вероятность перехода заявки на следующий уровень ($\alpha_0=1$);

$b_{k,m}^{(p)}=\int_{Q_{k-1}}^{Q_k} t^m b(t)dt$ – частичный момент трудоемкости порядка m;

$b_{k,m}^{(o)}=\int_{Q_{k-1}}^{Q_k} (t-Q_{k-1})^m b(t)dt$ – момент длительности неполного кванта;

$b_{k,m}^{(W)}=\alpha_k q_{k,m}+b_{k,m}^{(o)}$ – «взвешенный» момент длительности k-обслуживания, учитывающий вероятную потребность в полном кванте;

$b_k^{(L)}=b_k^{(W)}+\tau_k$ – «нагрузочные» моменты, получаемые сверткой взвешенных моментов и системных потерь (распределение последних зависит от номера уровня, если информация по соответствующим заявкам размещается в запоминающих устройствах разных типов);

$\lambda_k=\lambda\alpha_{k-1}$ – интенсивность потока на k-м уровне;

w_k – среднее время ожидания обслуживания на k-м уровне;

W_k – среднее время ожидания в очередях до k-го уровня включительно.

Интересующий нас конечный результат

$$v_k=W_k+Q_{k-1}+L_k+b_{k,1}^{(o)}.$$

Ключевым элементом этой технологии является расчет средних времен $\{w_k\}$ ожидания меченой заявкой очередного кванта на k-м уровне для $k \ge 2$. Выделим моменты времени:

X – прибытия меченой заявки в систему,

Y – начала ее обслуживания на $(k-1)$-м уровне,

Z – ее перехода на уровень k.

Назовем заявку, обслуживаемую в момент X (эта заявка с вероятностью ρ_i выбирается из i-й очереди), A-заявкой.

Обслуживание меченой заявки на $(k-1)$-м уровне может начаться лишь при отсутствии заявок на всех вышележащих уровнях. К этому моменту (Y) все заявки, которые находились в системе в момент X, окажутся уже на уровне k, но в связи с наличием более приоритетной меченой обслуживаться пока не будут (здесь и далее имеется в виду та часть заявок, которая добралась до соответствующего уровня). Они создадут для меченой среднюю задержку T_1. Им будет предшествовать A-заявка (задержка T_2).

Все заявки, пришедшие после меченой, окажутся позади нее на уровне $(k-1)$ и после момента Z меченая заявка будет ждать прохождения ими этого уровня (задержка T_3). Наконец, за время обслуживания меченой заявки на $(k-1)$-м уровне в систему придут новые заявки, и меченая будет ждать прохождения ими уровней $i=\overline{1,\ k-1}$ – задержка T_4.

Запишем формулы подсчета средних значений упомянутых задержек:

$$T_1 = r_{k,\,k}\left[\overline{B}(Q_{k-1})\sum_{i=1}^{k-1}\lambda_i w_i + \lambda_k w_k\right];$$

$$T_2 = r_{k,\,k}\overline{B}(Q_{k-1})\sum_{i=1}^{k-1}\rho_i;$$

$$T_3 = \lambda r_{k-1,\,k-1}\overline{B}(Q_{k-2})\left[\sum_{i=1}^{k-2}\left(w_i+\tau_i+q_i\right)+w_{k-1}\right];$$

$$T_4 = \lambda r_{1,\,k-1}\left(\tau_{k-1}+q_{k-1}\right).$$

После приравнивания w_k их суммы и элементарных преобразований получаем для всех $k \ge 2$ формулу

$$w_k = E_k / F_k,$$

где

$$E_k = \overline{B}(Q_{k-1})r_{k,\,k}\left[\lambda\sum_{i=1}^{k-1}\overline{B}(Q_{i-1})w_i + \sum_{i=1}^{k-1}\rho_i\right] +$$

$$+\lambda\Big\{\left(\tau_{k-1}+q_{k-1}\right)r_{1,k-1}+\overline{B}(Q_{k-2})\times$$

$$\times\left[\lambda\sum_{i=1}^{k-2}\left(w_i+\tau_i+q_i\right)+w_{k-1}\right]r_{k-1,\,k-1}\Big\};$$

$$F_k = 1-\lambda\left[r_{1,\,k-1}+\overline{B}(Q_{k-1})r_{k,\,k}\right].$$

Для первого уровня расчет выполняется как для обычных систем с относительным приоритетом.

Трудоемкости $\{r_{i,j}\}$ обработки заявки на уровнях $i=\overline{1,\ N}$ включительно с учетом отсева полностью обслуженных вычисляются рекуррентно:

$$r_{N,N} = \tau_N + \frac{1}{\overline{B}(Q_{N-1})}\int_{Q_{N-1}}^{\infty}(t-Q_{N-1})dB(t) =$$

$$= \tau_N + \frac{1}{\overline{B}(Q_{N-1})}\left[b_1 - \int_{Q_{N-1}}^{\infty} t\,dB(t) - Q_{N-1}\overline{B}(Q_{N-1})\right];$$

$$r_{i,N} = \tau_i + \frac{1}{\overline{B}(Q_{i-1})}\int_{Q_{i-1}}^{Q_i}(t-Q_{i-1})dB(t) + B(Q_i)\left[q_i + r_{i+1,N}\right],$$

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

Для ярусов $j=\overline{1,\ N-1}$ аналогичные формулы имеют вид

■ ***Средние времена ожидания по уровням***

Уровни	$\lambda = 0{,}5$		$\lambda = 0{,}6$	
	Имитация	Расчет	Имитация	Расчет
1	0,222	0,222	0,275	0,276
2	0,289	0,287	0,429	0,427
3	1,382	1,328	2,506	2,411
4	3,669	3,384	8,541	7,922
5	4,757	4,623	11,459	11,489

$$r_{j,j} = \tau_j + \frac{1}{\overline{B}(Q_{j-1})}\left[\int\limits_{Q_{j-1}}^{Q_j}\left(t - Q_{j-1}\right)dB(t) + \overline{B}(Q_j)q_j\right];$$

$$r_{i,j} = \tau_i + \frac{1}{\overline{B}(Q_{i-1})}\times$$

$$\times\left\{\int\limits_{Q_{i-1}}^{Q_i}\left(t - Q_{i-1}\right)dB(t) + \overline{B}(Q_i)\left[q_i + r_{i+1,j}\right]\right\},$$

$$i = j-1,\ j-2,\ \ldots,\ 1.$$

Для завершения описания методики остается обсудить технику расчета частичных моментов. Эти моменты, если соответствующие интегралы явно не берутся, удобно считать численным интегрированием по Симпсону с последовательным делением шага пополам и вычислением только дополнительных ординат [4].

Возможно также (после аппроксимации $b(t)$ гамма-плотностью с поправочным многочленом) использование известных степенных рядов и рекуррентных соотношений для неполной гамма-функции [5]. К несложным формулам приводит и H_2 -аппроксимация распределения времени обслуживания.

Для последнего уровня $Q_N = \infty$, и частичные моменты получаются вычитанием ранее полученных из полных:

$$b_{N,m}^{(p)} = b_m^{(p)} - \sum_{i=1}^{N-1} b_{i,m}^{(p)}.$$

Описанная методика была запрограммирована и реализована применительно к пятиуровневой системе с начальным квантом $q_1 = 0{,}2$, удвоением кванта на уровнях до 4-го включительно, при постоянных потерях на переключение $\tau = 0{,}1$ и длительности обслуживания по закону Эрланга 2-го порядка с единичным средним. Результаты представлены в таблице.

Таким образом, расчет вполне согласуется с имитационным экспериментом.

Многоуровневая система несколько ухудшает оперативность обслуживания коротких заявок, которые могут застать ЦП занятым выдачей длинного кванта. Поэтому имеет смысл наиболее длинным заявкам выдавать требуемый квант по частям, в интервалах между которыми возможно обслуживание коротких заявок. Соответственно последний уровень на рисунке разбивается на несколько подуровней. Такая схема может быть применена для анализа квантованного приоритетного обслуживания, реализуемого на фоне вычислений большой трудоемкости.

При малых «дробных» квантах она фактически предоставляет коротким заявкам приоритет с прерыванием, хотя и несколько задержанным.

Укажем еще несколько возможных обобщений многоуровневой схемы:

прямой доступ заявок из внешнего источника на промежуточные уровни;

обслуживание части верхних уровней с абсолютным приоритетом (в этом случае к системе в целом применяется алгоритм расчета модели со смешанными приоритетами);

выдача «дробных» квантов случайной длительности с объединением подуровней в одну RR-систему (этот вариант хорошо моделирует длительный счет в фоновом разделе, прерываемый в моменты промежуточных выдач);

распределение недообслуженных заявок с определенными вероятностями между несколькими нижележащими уровнями.

Комбинация двух последних вариантов особенно удачно описывает реальные процессы обработки неоднородных заявок (на первом уровне вновь прибывшие заявки классифицируются по системным очередям). Разумеется, все перечисленные в начале статьи сферы коллективного использования вычислительных ресурсов имеют заметную специфику, но описанная схема может служить источником полезных аналогий и технологических рецептов.

В заключение отметим ее педагогическую ценность при изучении дисциплин типа «Компьютерное моделирование», поскольку составление уравнений баланса объемов работы на каждом из уровней далеко не тривиально и способствует лучшему уяснению законов сохранения теории очередей [4].

Литература

1. **Балыбердин В. А.** Методы анализа мультипрограммных систем. – М.: Радио и связь, 1992. – 152 с.
2. **Клейнрок Л.** Вычислительные системы с очередями / Пер. с англ. – М.: Мир, 1979. – 600 с.
3. Основы теории вычислительных систем /Под ред. С. А. Майорова. – М.: Высшая школа, 1978. – 408 с.
4. **Рыжиков Ю. И.** Машинные методы расчета систем массового обслуживания. – Л.: ВИКИ им. А. Ф. Можайского, 1979. – 177 с.
5. **Янке Е., Эмде Ф., Леш Ф.** Специальные функции. – М.: Наука, 1968. – 344 с.
6. **Яшков С. Ф.** Анализ очередей в ЭВМ. – М.: Радио и связь, 1989. – 216 с.
7. **Coffman E. G., Denning P. J.** Operating Systems Theory. – Englewood Cliffs, NJ: Prentice-Hall, 1973. – 331 p.