УДК 519.68

Г.Г.Цегелик, О.Б.Семенська

До вибору оптимального числа процесорів багатопроцесорних ЕОМ при пошуку інформації в послідовних файлах

В [1] вперше розглядається задача вибору оптимального числа процесорів багатопроцесорних ЕОМ при пошуку інформації в послідовних файлах для різних законів розподілу ймовірностей звертання до записів, досліджується картина залежності числа процесорів від зміни закону розподілу ймовірностей. При цьому вважається, що файл розбитий на n блоків по m записів в кожному і при пошуку запису використовується метод m-паралельного послідовного перегляду блоків записів в основній пам'яті. В даній статті розглядається більш загальна задача, коли блоки записів містять по ml записів, де l – натуральне число.

Припустимо, що послідовний упорядкований файл міститься в зовнішній пам'яті ЕОМ, до складу якої входить m процесорів, які паралельно працюють і мають спільне поле пам'яті. Будемо вважати, що записи файла умовно розбиті на n блоків по ml записів в кожному. Нехай N – число записів файла, $a_0 = b_0 + d_0 ml$ – час читання блоку записів в основну пам'ять, де b_0, d_0 – деякі константи, t_0 – час виконання операції m-паралельного послідовного перегляду записів в основній пам'яті, p_i - імовірність звертання до i-го запису файла, E – математичне сподівання загального часу, необхідного для пошуку запису в файлі. Будемо також вважати, що для пошуку запису відбувається послідовне зчитування блоків записів в основну пам'ять і їх m-паралельний послідовний перегляд. Тоді E виразиться формулою:

$$E = \sum_{k=1}^{n} \sum_{i=1}^{l} \sum_{j=1}^{m} (ka_0 + ((k-1)l + i)t_0) p_{(k-1)ml+(i-1)m+j} .$$

Знайдемо явний вираз для E у випадку різних законів розподілу ймовірностей звертання до записів і визначимо значення параметра m при фіксованому значенні l, при якому E досягає мінімуму.

1. Якщо розподіл імовірностей звертання до записів є рівномірний, то для E одержуємо вираз

$$E=\frac{1}{2}\left((n+1)a_0+(nl+1)t_0\right),$$

або

$$E=\frac{1}{2}\left[\left(\frac{N}{ml}+1\right)(b_0+d_0ml)+\left(\frac{N}{m}+1\right)t_0\right].$$

Функція E досягає мінімуму при

$$m=\frac{1}{l}\sqrt{N\left(\frac{b_0}{d_0}+l\frac{t_0}{d_0}\right)}.$$

2. Нехай імовірності звертання до записів задовольняють "бінарний" розподіл [1]. Тоді аналогічно як в [2] одержимо:

$$E=\frac{2^{ml}}{2^{ml}-1}\left(1-2^{-N}\right)a_0+\left[\frac{2^m}{2^m-1}\left(1-2^{-N}\right)+\frac{n-l}{2^N}\right]t_0.$$

Нехтуючи величиною 2^{-N}, з достатньо високою точністю можна прийняти

$$E=\frac{2^{ml}}{2^{ml}-1}(b_0+d_0ml)+\frac{2^m}{2^m-1}t_0.$$

Оскільки

$$\frac{dE}{dm}=-\frac{l2^{ml}\ln 2}{(2^{ml}-1)^2}(b_0+d_0ml)+\frac{2^{ml}}{2^{ml}-1}d_0l-\frac{2^m\ln 2}{(2^m-1)^2}t_0,$$

то для визначення значення параметра m, при якому функція E досягає мінімуму, одержуємо рівняння

$$(2^m-1)^2 2^{m(l-1)}l\left(2^{ml}-1-\ln 2\left(\frac{b_0}{d_0}+ml\right)\right)=(2^{ml}-1)^2\ln 2\frac{t_0}{d_0}.$$

3. Припустимо, що ймовірності звертання до записів задовольняють закон Зіпфа [1]. В цьому випадку аналогічно як в [2] одержимо

$$E=\frac{1}{H_N}\Big[\left((n+1)H_N-S_{ml}(n)\right)(a_0+t_0l)+$$
$$+\left((1-l)H_N+lS_{ml}(n)-S_m(nl)\right)t_0\Big],$$

де

$$S_{ml}(n)=\sum_{k=1}^{n}H_{kml},\qquad S_m(nl)=\sum_{k=1}^{nl}H_{km}.$$

Використовуючи апроксимацію $S_{ml}(n)$ і $S_m(nl)$ відповідно виразами [2,3]

$$\overline{S}_{ml}(n) = n(H_N - 1) + \frac{1}{2}\ln n + C_1,$$

$$\overline{S}_m(nl) = nl(H_N - 1) + \frac{1}{2}\ln nl + C_1,$$

де $C_1 = \frac{1}{2}\ln 2\pi$, з достатньо високою точністю можемо прийняти

$$E = \frac{1}{H_N}\left[\left(H_N + n - \frac{1}{2}\ln n - C_1\right)a_0 + \right.$$
$$\left. + \left(H_N + nl - \frac{1}{2}\ln nl - C_1\right)t_0\right],$$

або

$$E = \frac{1}{H_N}\left[\left(H_N + \frac{N}{ml} - \frac{1}{2}\ln\frac{N}{ml} - C_1\right)(b_0 + d_0 ml) + \right.$$
$$\left. + \left(H_N + \frac{N}{m} - \frac{1}{2}\ln\frac{N}{m} - C_1\right)t_0\right].$$

Оскільки

$$\frac{dE}{dm} = \frac{1}{H_N}\left[\left(H_N + \frac{N}{ml} - \frac{1}{2}\ln\frac{N}{ml} - C_1\right)d_0 l + \right.$$
$$\left. + (b_0 + d_0 ml)\left(\frac{1}{2m} - \frac{N}{m^2 l}\right) + \left(\frac{1}{2m} - \frac{N}{m^2}\right)t_0\right],$$

то для знаходження значення параметра m, при якому функція E досягає мінімуму, одержуємо рівняння

$$m^2 l^2 (2H_N - 2C_1 + 1 - \ln N + \ln ml) + ml\left(\frac{b_0}{d_0} + \frac{t_0}{d_0}\right) = 2N\left(\frac{b_0}{d_0} + l\frac{t_0}{d_0}\right).$$

4. Нехай імовірності звертання до записів задовольняють узагальнений закон розподілу [1], де $0 < c < 2, c \neq 1$. Тоді аналогічно як в [2] для E маємо вираз

$$E = \frac{1}{H_N^{(c)}}\left[\left((n+1)H_N^{(c)} - S_{ml}^{(c)}(n)\right)(a_0 + t_0 l) + \right.$$
$$\left. + \left((l-1)H_N^{(c)} + lS_{ml}^{(c)}(n) - S_m^{(c)}(nl)\right)t_0\right],$$

де

$$S_{ml}^{(c)}(n)=\sum_{k=1}^{n}H_{kml}^{(c)}, \quad S_{m}^{(c)}(nl)=\sum_{k=1}^{nl}H_{km}^{(c)}.$$

Використовуючи апроксимацію $S_{ml}^{(c)}(n)$ і $S_{m}^{(c)}(nl)$ відповідно виразами [2,3]

$$\bar{S}_{ml}^{(c)}(n)=nH_N^{(c)}+\frac{N^{1-c}}{1-c}\left(\frac{c-1}{2-c}n+\frac{\alpha^{(c)}(n)}{n^{1-c}}\right),$$

$$\bar{S}_{m}^{(c)}(nl)=nlH_N^{(c)}+\frac{N^{1-c}}{1-c}\left(\frac{c-1}{2-c}nl+\frac{\alpha^{(c)}(nl)}{(nl)^{1-c}}\right),$$

де

$$\alpha^{(c)}(k)=H_k^{(c-1)}-\frac{1}{2-c}k^{2-c},$$

з достатньо високою точністю можемо прийняти

$$E=\frac{1}{H_N^{(c)}}\left[\left(H_N^{(c)}-\frac{N^{1-c}}{1-c}\left(\frac{c-1}{2-c}n+\frac{\alpha^{(c)}(n)}{n^{1-c}}\right)\right)a_0+\right.$$
$$\left.+\left(H_N^{(c)}+\frac{N^{1-c}}{1-c}\left(\frac{c-1}{2-c}nl-\frac{\alpha^{(c)}(nl)}{(nl)^{1-c}}\right)\right)t_0\right],$$

або

$$E=\frac{1}{H_N^{(c)}}\left[\left(H_N^{(c)}-\frac{N^{1-c}}{1-c}\left(\frac{c-1}{2-c}n+\frac{\alpha^{(c)}(n)}{n^{1-c}}\right)\right)\left(b_0+\frac{d_0N}{n}\right)+\right.$$
$$\left.+\left(H_N^{(c)}+\frac{N^{1-c}}{1-c}\left(\frac{c-1}{2-c}nl-\frac{\alpha^{(c)}(nl)}{(nl)^{1-c}}\right)\right)t_0\right].$$

Знайдемо $\frac{dE}{dn}$, замінюючи похідну від функції $\alpha^{(c)}(n)$ різницею $\alpha^{(c)}(n+1)-\alpha^{(c)}(n)$. Отримаємо

$$\frac{dE}{dn}\approx\frac{1}{H_N^{(c)}}\left[\frac{N^{1-c}}{1-c}\left(\frac{1-c}{2-c}-\frac{\alpha^{(c)}(n+1)-\alpha^{(c)}(n)}{n^{1-c}}+(1-c)\frac{\alpha^{(c)}(n)}{n^{2-c}}\right)\times\right.$$
$$\times\left(b_0+\frac{d_0N}{n}\right)+\left(\frac{N^{1-c}}{1-c}\left(\frac{c-1}{2-c}n-\frac{\alpha^{(c)}(n)}{n^{1-c}}\right)-H_N^{(c)}\right)\frac{N}{n^2}d_0+$$

$$+\frac{N^{1-c}}{1-c}\left(\frac{1-c}{2-c}l-\frac{\alpha^{(c)}(nl+1)-\alpha^{(c)}(nl)}{(nl)^{1-c}}+(1-c)l\frac{\alpha^{(c)}(nl)}{(nl)^{2-c}}\right)t_0\Bigg].$$

Для наближеного визначення значення параметра n, при якому функція E досягає мінімуму, одержуємо таке рівняння

$$l^{1-c}\left(\frac{n^{3-c}}{2-c}-n^2\frac{\alpha^{(c)}(n+1)-\alpha^{(c)}(n)}{1-c}+n\alpha^{(c)}(n)\right)\frac{b_0}{d_0}+$$

$$+\left(\frac{n^{3-c}}{2-c}l^{2-c}-n^2\frac{\alpha^{(c)}(nl+1)-\alpha^{(c)}(nl)}{1-c}+n\alpha^{(c)}(nl)\right)\frac{t_0}{d_0}=$$

$$=l^{1-c}\left(H_N^{(c)}N^{(c)}n^{1-c}+\frac{N}{1-c}\Big(n\big(\alpha^{(c)}(n+1)-\alpha^{(c)}(n)\big)-(2-c)\alpha^{(c)}(n)\Big)\right).$$

1. Семенська О.Б., Цегелик Г.Г. Задача вибору раціонального чи оптимального числа процесорів багатопроцесорних ЕОМ при пошуку інформації в послідовних файла: Препринт №2/97. – Львів: Львівський державний університет ім. І.Франка, 1997. – 52 с. 2. Цегелик Г.Г. Организация и поиск информации в базах данных. Львов: Вища школа, 1987. – 176 с. 3. Цегелик Г.Г. Системы распределенных баз данных. Львов: Свит, 1990. – 168 с.

Стаття надійшла до редколегії 29.12.97