«Южно-Уральский государственный университет»
(национальный исследовательский университет)
ФГБОУ ВПО «ЮУрГУ» (НИУ)

Оценка коммуникационных затрат при обработке фрагментированного отношения

докладчик Губин М.В.

Международная суперкомпьютерная конференция

Научный сервис в сети Интернет: все грани параллелизма

2013

Фрагментный параллелизм

Фрагментация отношения;

Функция фрагментации;

Функция пересылки.

Пересылки являются одной из самых затратных операций.

Необходимо оценивать количество коммуникаций.

Фрагментированное отношение $\boldsymbol{R}$.

Пусть $|R| = m$– количество кортежей в R.

Пусть отношение R разбито на k фрагментов с помощью ***функции фрагментации*** $\varphi\colon R \to \{0,1,2,\dots,k-1\}$.

Тогда фрагмент отношения R, хранящийся на j–том узле может быть определен следующим образом:

$$R_j = \{\, r \mid r \in R,\ \varphi(r) = j\}. \tag{1}$$

Лемма 1. Пусть для отношения R задана функция фрагментации $\varphi: R \to \{0,1,2,\dots,k-1\}$. Допустим, что существует взаимно однозначное отображение $\gamma: R \to \{0,1,2,\dots,m-1\}$ такое, что

$$\gamma(r) \bmod k = \varphi(r), \forall r \in R\ . \tag{2}$$

Тогда размеры фрагментов $R_0,\dots,R_{k\text{-}1}$ будут между собой ***примерно равны***, то есть:

$$\lim_{m\to\infty} \frac{|R_i|}{|R_j|} = 1, \forall i,j \in \{0,1,2,\dots,k-1\}. \tag{3}$$

Доказательство.

Определим

$$M_j = \{\, l \mid \gamma^{-1}(l) \in R_j, 0 \le l < m\}, j = 0,1, \dots, k-1, \qquad (4)$$

тогда

$$\left|M_j\right| = \left|R_j\right|. \qquad (5)$$

Определим

$$\varphi'(l) = l \bmod k, \qquad (6)$$

$$M'_j = \{\, l \mid \varphi'(l) = j, 0 \le l < m\},\ \ j = 0,1, \dots, k-1. \qquad (7)$$

Покажем, что

$$M_j = M'_j,\ \ \forall j = 0,1, \dots, k-1.$$

Пусть $l \in M_j$. (8)

Тогда из (4) следует $\gamma^{-1}(l) \in R_j$.

В силу (1), получаем $\varphi(\gamma^{-1}(l)) = j$.

Используя (2), получаем $\gamma(\gamma^{-1}(l)) \bmod k = j$,

что эквивалентно $l \bmod k = j$.

Сопоставляя это с (6) находим $\varphi'(l) = j$.

Учитывая (7), отсюда получаем $l \in M'_j$,

откуда в силу (8) следует

$$M_j \subseteq M'_j. \tag{9}$$

Пусть теперь $$l \in M'_j. \tag{10}$$

Учитывая (7), получаем $$\varphi'(l) = j.$$

Сопоставляя это с (6) находим $$l \bmod k = j,$$

что эквивалентно $$\gamma\left(\gamma^{-1}(l)\right) \bmod k = j.$$

Используя (2), получаем $$\varphi\left(\gamma^{-1}(l)\right) = j.$$

В силу (1), $$\gamma^{-1}(l) \in R_j.$$

Тогда из (4) следует $$l \in M_j.$$

Учитывая (10), получаем $$M'_j \subseteq M_j,$$

откуда в силу (9) следует $$M'_j = M_j.$$

В сочетании с (5) это дает $$\left|M'_j\right| = \left|R_j\right|,\ \forall j \in \{0,1,2,\dots,k-1\}. \tag{11}$$

Таким образом, утверждение (3) равносильно утверждению

$$\lim_{m\to\infty}\frac{|M_i'|}{|M_j'|}=1,\forall i,j\in\{0,1,2,\dots,k-1\}. \tag{12}$$

Докажем это утверждение. В силу (6) и (7) имеем

$$M_j'=\{\,l\mid l \bmod k=j, 0\le l<m\},\ \ j=0,1,\dots,k-1.$$

Отсюда следует $\left\lfloor\frac{m}{k}\right\rfloor\le\left|M_j'\right|\le\left\lceil\frac{m}{k}\right\rceil,\ \forall j\in\{0,1,2,\dots,k-1\}.$

Тогда $$\frac{\left\lfloor\frac{m}{k}\right\rfloor}{\left\lceil\frac{m}{k}\right\rceil}\le\frac{|M_i'|}{|M_j'|}\le\frac{\left\lceil\frac{m}{k}\right\rceil}{\left\lfloor\frac{m}{k}\right\rfloor},\forall i,j\in\{0,1,2,\dots,k-1\}. \tag{13}$$

Имеем $$\lim_{m\to\infty}\frac{\left\lfloor\frac{m}{k}\right\rfloor}{\left\lceil\frac{m}{k}\right\rceil}=\lim_{m\to\infty}\frac{\left\lceil\frac{m}{k}\right\rceil}{\left\lfloor\frac{m}{k}\right\rfloor}=1.$$

Вместе с (13) это доказывает справедливость утверждения (12), а вместе с ним и утверждения (3). Лемма доказана.

Определение 1. Пусть задана функция фрагментации

$$\varphi\colon R(A_1, \dots, A_n) \to \{0,1,2, \dots, k-1\}.$$

Будем говорить, что φ ***функционально зависит*** от атрибута A_i $(0 < i \leq n)$, если для любых $r, \acute{r} \in R$ таких, что $\pi_{A_i}(r) = \pi_{A_i}(\acute{r})$, имеем $\varphi(r) = \varphi(\acute{r})$.

Определение 2. Пусть имеется отношение $R(\dots, A, \dots)$. Пусть D_A – множество всех значений атрибута A, присутствующих в R. Обозначим через $T(R, A, a)$ количество кортежей в отношении R, у которых атрибут A принимает значение a. Будем говорить, что значения атрибута A ***равномерно распределены*** в R, если, для любых $a, \tilde{a} \in \mathrm{D}_A$ имеем

$$T(R, A, a) = T(R, A, \tilde{a}). \tag{14}$$

Лемма 2. Пусть значения атрибута A равномерно распределены в $R(\dots, A, \dots)$, $m = |R|$ кратно $V(R, A)$, а $V(R, A)$ кратно $k > 0$. Тогда существует функционально зависящая от A функция фрагментации φ, которая обеспечивает равномерное распределение кортежей по фрагментам:

$$\lim_{m\to\infty} \frac{|R_i|}{|R_j|} = 1, \forall i, j \in \{0,1,2, \dots, k-1\}. \tag{15}$$

Для доказательства леммы необходимо построить функцию фрагментации φ, которая обеспечивает равномерное распределение кортежей по фрагментам и функционально зависит от A.

Выполним сортировку кортежей R в порядке возрастания значений атрибута A:

$$R=\{r_0, r_1,\ldots, r_{m-1}\}, \tag{16}$$

где $\pi_A(r_i) \leq \pi_A(r_{i+1})$ для всех i=0,…,m-2.

Обозначим через $\eta(r)$ порядковый номер кортежа в последовательности (16). Определим функцию

$$\bar{\gamma}(i) = \left\lfloor \frac{i}{m/k} \right\rfloor + \left(i \bmod \frac{m}{k}\right) k, \quad i \in \mathbb{Z}_{\geq 0}, \tag{17}$$

Положим

$$\gamma(r) = \bar{\gamma}\big(\eta(r)\big).$$

Покажем, что $\bar{\gamma}$ является взаимно-однозначным отображением множества порядковых номеров кортежей R в множество $\{0, 1, ..., m-1\}$.

$$0 \le \bar{\gamma}(i) \le m - 1,\ \forall i \in \{0, \dots, m-1\} \qquad (18)$$

$$\text{и } \bar{\gamma}^{-1}\big(\bar{\gamma}(i)\big) = i.$$

Из (17) непосредственно следует $\min_{0\le i<m} \bar{\gamma}(i) = 0$. И

$$\max_{0\le i<m}\big(\bar{\gamma}(i)\big) = \max_{r\in R}\left\lfloor \frac{i}{m/k} \right\rfloor + \max_{r\in R}\left(\left(i \bmod \frac{m}{k}\right)k\right)$$

$$= \left\lfloor \frac{1}{m/k} \max_{r\in R} i \right\rfloor + k \max_{r\in R}\left(i \bmod \frac{m}{k}\right)$$

$$= \left\lfloor \frac{1}{m/k}(m-1) \right\rfloor + k\left(\frac{m}{k} - 1\right) = \left\lfloor \frac{k}{m}(m-1) \right\rfloor + k\left(\frac{m}{k} - 1\right)$$

$$= \left\lfloor k - \frac{1}{m} \right\rfloor + m - k = k - 1 + m - k = m - 1.$$

Для доказательства существования обратной функции $\bar{\gamma}^{-1}$

напишем её:

$$\bar{\gamma}^{-1}(j) = \left\lfloor \frac{j}{k} \right\rfloor + (j \; mod \; k)c. \tag{19}$$

Подставляя (19) в (17), получаем справедливость утверждения

$$\bar{\gamma}^{-1}\big(\bar{\gamma}(i)\big) = i.$$

Поскольку функция γ существует и является взаимно-однозначным отображением R в множество {0, 1, ..., m–1}, тогда, согласно лемме 1, функция φ обеспечивает равномерное распределение кортежей по фрагментам.

Определим $$\varphi(r) = \gamma(r) \bmod k. \quad (20)$$

Покажем, что φ функционально зависит от A, то есть, если

$$\pi_A(r) = \pi_A(\acute{r}), \text{ то } \varphi(r) = \varphi(\acute{r}) \ \ \forall r, \acute{r} \in R.$$

Поскольку мы упорядочили кортежи в R, то можем представить множество R в виде упорядоченных подмножеств

$$R=\{G_0, G_1, \ldots, G_{V(R,A)-1}\}$$

обладающих свойствами:

1) $\pi_A(r) = \pi_A(\acute{r}) \ \ \forall r, \acute{r} \in G_l$;

2) $\pi_A(r) \neq \pi_A(\acute{r}) \ \ \forall r \in G_l, \acute{r} \notin G_l$;

3) $|G_l| = \frac{m}{V(R,A)}$;

для любого и целого $l \in [0, V(R,A) - 1]$.

По условию леммы $V(R, A)$ кратно k, тогда любое подмножество G_l будет принадлежать только одному фрагменту множества R_j, т. е.

$$\text{если } G_l \subset R_j, \text{тогда } G_l \cap R_i = \emptyset$$

$$\text{для } \forall l \in \{0, \dots, V(R, A)\} \ \forall i, j \in \{0, \dots, k-1\}, i \neq j.$$

Из этого понятно, что если $\pi_A(r) = \pi_A(\acute{r})$, то $\varphi(r) = \varphi(\acute{r}) \ \ \forall r, \acute{r} \in R$.

И по определению 2, φ функционально зависит A.

Определение 3. Пусть имеется отношение $R(A, B, \dots)$. Пусть $D_A = \{a_0, \dots, a_{V(R,A)-1}\}$ – множество всех значений атрибута A, присутствующих в R. Обозначим $R_u^A = \sigma_{A=a_u}(R)$, $u \in \{0, \dots, V(R, A) - 1\}$. Пусть $D_B = \{b_0, \dots, b_{V(R,B)-1}\}$ – множество всех значений атрибута B, присутствующих в R. Будем говорить, что значения атрибута B ***распределены равномерно относительно атрибута*** A, если, для любого $b \in D_B$ имеем

$$T\left(R_0^A, B, b\right) = T\left(R_1^A, B, b\right) = \cdots = T\left(R_{V(R,A)-1}^A, B, b\right). \tag{21}$$

Теорема. Пусть имеется отношение $R(A, B, \ldots)$, в котором значения атрибута A распределены равномерно, а значения атрибута B, в свою очередь, распределены равномерно относительно атрибута A. Положим $m = |R|$. Пусть m кратно $V(R, A)$, а $V(R, A)$ кратно k, где k – количество фрагментов, на которое необходимо разбить R ($k > 0$). Тогда существует функция фрагментации φ, функционально зависящая от атрибута A, такая, что для любой функции пересылки ψ, функционально зависящей от атрибута B, количество пересылаемых кортежей равно $(1 - 1/k)m$.

Доказательство. Выберем в качестве функции фрагментации φ такую функцию, функционально зависящую от атрибута A, которая обеспечивает равномерное распределение кортежей по фрагментам. Такая функция существует в силу леммы 2. Обозначим через R_j j-тый фрагмент отношения R:

$$R_j = \{r|r \in R, \varphi(r) = j\}.$$

Пусть $D_A = \{a_0, \dots, a_{V(R,A)-1}\}$ – множество всех значений атрибута A, присутствующих в R. Обозначим $R_u^A = \sigma_{A=a_u}(R)$, $u = \{0, \dots, V(R,A) - 1\}$. Так как по условию теоремы значения атрибута A распределены равномерно в R, имеем

$$|R_u^A| = \frac{m}{V(R,A)} \qquad (u = 0, \dots, V(R,A) - 1). \tag{22}$$

Поскольку функция фрагментации φ функционально зависит от A, то

$$\forall u \in \{0, \dots, V(R,A) - 1\} \;\; \exists j \in \{0, \dots, \mathrm{k} - 1\} : \quad R_u^A \subset R_j.$$

В силу того, что φ обеспечивает равномерное распределение кортежей по фрагментам, отсюда следует, что мы можем перенумеровать элементы множества $\{R_u^A | u = 0, \dots, V(R,A) - 1\}$ таким образом, что

$$R_j = \bigcup_{u=\frac{V(R,A)}{k}j}^{\frac{V(R,A)}{k}(j+1)-1} R_u^A \qquad (j = 0, \dots, k-1). \tag{23}$$

Причем

$$\forall\, u \neq v : R_u^A \cap R_v^A = \emptyset, \tag{24}$$

и

$$\forall\, j \in \{0, \dots, k-1\}\ \forall\, u \in \{0, \dots, V(R,A) - 1\}: \quad \left|R_j\right| = \frac{V(R,A)}{k} |R_u^A|. \tag{25}$$

Так как по условию теоремы значения атрибута B распределены равномерно относительно атрибута A, то из определения 3 следует, что

$$\left|\sigma_{B=b}\left(R_0^A\right)\right| = \left|\sigma_{B=b}\left(R_1^A\right)\right| = \ldots = \left|\sigma_{B=b}\left(R_{V(R,A)-1}^A\right)\right| \ (\forall b \in \mathrm{D}_B).$$

Отсюда и из (23)-(25) получаем

$$|\sigma_{B=b}(R_0)| = |\sigma_{B=b}(R_1)| = \ldots = |\sigma_{B=b}(R_{k-1})| \ (\forall b \in \mathrm{D}_B). \qquad (26)$$

Пусть $D_B = \{b_0, \dots, b_{V(R,B)-1}\}$ – множество всех значений атрибута B, присутствующих в R. Так как по условию теоремы функция пересылки ψ функционально зависит от атрибута B, то существует функция $\bar{\psi}: D_B \to \{0, \dots, k-1\}$ такая, что

$$\forall r \in R:\ \psi(r) = \bar{\psi}(r.B). \tag{27}$$

Обозначим $R_v^B = \sigma_{B=b_v}(R)$, $v = \{0, \dots, V(R,B)-1\}$. Имеем

$$R = \bigcup_{0 \le v < V(R,B)} R_v^B, \tag{28}$$

причем

$$\forall v \ne \acute{v}:\ \ R_v^B \cap R_{\acute{v}}^B = \emptyset. \tag{29}$$

Имеем

$$\forall r \in R_v^B: \psi(r) = \bar{\psi}(b_v).$$

Тогда кортежи множества $\sigma_{B=b_v}\left(R_{\overline{\psi}(b_v)}\right)$ пересылаться не будут. Учитывая (26), заключаем, что из множества R_v^B пересылке не подвергнутся в точности $\left|\sigma_{B=b_v}(R_0)\right|$ кортежей. Используя (28) и (29) получаем, что количество кортежей из R, которые не подвергнутся пересылке, равно

$$\sum_{v=0}^{V(R,B)-1} \left|\sigma_{B=b_v}(R_0)\right| = |R_0|.$$

В силу того, что φ обеспечивает равномерное распределение кортежей по фрагментам, имеем $|R_0| = \frac{m}{k}$. Отсюда следует, что количество пересылаемых кортежей равно

$$m - \frac{m}{k} = \left(1 - \frac{1}{k}\right)m.$$

Теорема доказана.

Литература

1. Rahm E. Parallel Query Processing in Shared Disk Database Systems // ACM SIGMOD Record, 1993. Vol. 22, No. 4. P. 32-37.

2. Соколинский Л.Б. Обзор архитектур параллельных систем баз данных // Программирование. 2004. No. 6. C. 49-63.

3. Лепихов А.В., Соколинский Л.Б. Обработка запросов в СУБД для кластерных систем // Программирование. 2010. № 4. С. 25-39.

4*. Hasan W., Motwani R.* Coloring Away Communication in Parallel Query Optimization // VLDB'95, Proceedings of 21th International Conference on Very Large Data Bases, September 11–15, 1995, Zurich, Switzerland. Morgan Kaufmann. 1995. P. 239–250.