УДК004.657

ДЕКОМПОЗИЦИЯ ОПЕРАЦИИ ГРУППИРОВКИ НА БАЗЕ РАСПРЕДЕЛЕННЫХ КОЛОНОЧНЫХ ИНДЕКСОВ

Е.В. Иванова, Л.Б. Соколинский

Статья посвящена вопросу декомпозиции реляционной операции группировки на основе распределенных колоночных индексов с доменно-интервальной и транзитивной фрагментацией. Такая декомпозиция позволяет организовать эффективное параллельное выполнение запросов к сверхбольшим базам данных на современных кластерных вычислительных системах, оснащенных многоядерными ускорителями.

Ключевые слова: сверхбольшие базы данных, параллельная обработка запросов, колоночные индексы, доменно-интервальная фрагментация, транзитивная фрагментация, декомпозиция реляционных операций, операция группировки.

Введение

В настоящее время фактически единственным эффективным решением проблемы хранения и обработки сверхбольших баз данных является использование параллельных систем баз данных, обеспечивающих распределенную обработку запросов на многопроцессорных вычислительных системах с распределенной памятью [1–5].

В последние годы основным способом наращивания производительности процессоров является увеличение количества ядер, а не тактовой частоты, и эта тенденция, вероятно, сохранится [6]. Сегодня GPU (Graphic Processing Units) и Intel MIC (Many Integrated Cores) значительно опережают традиционные процессоры в производительности по арифметическим операциям и пропускной способности памяти, позволяя использовать сотни процессорных ядер для выполнения десятков тысяч потоков. Последние исследования показывают, что многоядерные ускорители могут эффективно использоваться для обработки запросов к базам данных в оперативной памяти [7–9].

В соответствие с этим актуальной является задача разработки новых эффективных методов параллельной обработки баз данных в оперативной памяти на современных многопроцессорных вычислительных системах с многоядерными ускорителями. Для решения этой задачи в работах [10, 11] были предложены индексные структуры специального вида, которые называются распределенными колоночными индексами. Распределенные колоночные индексы позволяют провести декомпозицию реляционных операций, допускающую их эффективное параллельное выполнение на кластерных вычислительных системах с многоядерными ускорителями. В данной работе рассмотрен вопрос декомпозиции реляционной операции группировки. Для обозначения реляционных операций в статье используется нотация, заимствованная из монографии [12].

1. Колоночный индекс и доменно-интервальная фрагментация

Под $R(A, B_1, \ldots, B_u)$ будем понимать отношение R с *виртуальным ключом* A (виртуальным идентификатором целочисленного типа, однозначно определяющим кортеж) и атрибутами $B_1, \ldots, B_u$, представляющее собой множество кортежей длины $u+1$ вида $(a, b_1, \ldots, b_u)$, где a – целое неотрицательное число, и $\forall j \in \{1, \ldots, u\}\left(b_j \in \mathcal{D}_{B_j}\right)$. Здесь $\mathcal{D}_{B_j}$ - домен атрибута B_j. Через $r.B_j$ будем обозначать значение атрибута B_j, через $r.A$ – значение виртуального ключа в кортеже r: $r = (r.A, r.B_1, \ldots, r.B_u)$. *Виртуальный ключ* отношения R обладает свойством $\forall r', r'' \in R\left(r' \neq r'' \Leftrightarrow r'.A \neq r''.A\right)$. Под *адресом кортежа* r мы будем понимать значение первичного ключа этого кортежа. Для получения кортежа отношения R по его адресу будем использовать *функцию разыменования* $\&_R$: $\forall r \in R\left(\&_R(r.A) = r\right)$.

Определение 1. Пусть задано отношение $R(A, B, \ldots)$, $T(R) = n$. Пусть на множестве $\mathcal{D}_B$ задано отношение линейного порядка. *Колоночным индексом* $I_{R.B}$ атрибута B отношения R называется упорядоченное отношение $I_{R.B}(A, B)$, удовлетворяющее следующим требованиям:

$$T(I_{R.B}) = n \text{ и } \pi_A\left(I_{R.B}\right) = \pi_A\left(R\right); \tag{1}$$

$$\forall x_1, x_2 \in I_{R.B}\left(x_1 \leq x_2 \Leftrightarrow x_1.B \leq x_2.B\right); \tag{2}$$

$$\forall r \in R\left(\forall x \in I_{R.B}\left(r.A = x.A \Rightarrow r.B = x.B\right)\right). \tag{3}$$

Условие (1) означает, что множества значений первичных ключей (адресов) индекса и индексируемого отношения совпадают. Условие (2) означает, что элементы индекса упорядочены в порядке возрастания значений атрибута B. Условие (3) означает, что атрибут A элемента индекса содержит адрес кортежа отношения R, имеющего такое же значение атрибута B, как и у данного элемента колоночного индекса.

Теорема 1. Пусть задано отношение $R(A, B, \ldots)$. Пусть для отношения R задан колоночный индекс $I_{R.B}$. Тогда

$$\pi_B\left(I_{R.B}\right) = \pi_B\left(R\right). \tag{4}$$

Другими словами, колоночный индекс $I_{R.B}$ представляет все множество значений атрибута B отношения R с учетом повторяющихся значений.

Определение 2. Пусть на множестве значений домена $\mathcal{D}_B$ задано отношение линейного порядка. Пусть также задано разбиение множества $\mathcal{D}_B$ на $k>0$ непересекающихся интервалов:

$$\left.\begin{array}{l} V_0=\left[v_0;v_1\right],V_1=\left(v_1;v_2\right],\ldots,V_{k-1}=\left(v_{k-1};v_k\right]; \\ v_0<v_1<\ldots<v_k; \\ \mathcal{D}_B=\bigcup\limits_{i=0}^{k-1}V_i. \end{array}\right\}. \tag{5}$$

Функция $\varphi_{\mathcal{D}_B}:\mathcal{D}_B\to\left\{0,\ldots,k-1\right\}$ называется *интервальной функцией фрагментации* для домена $\mathcal{D}_B$, если она удовлетворяет следующему условию:

$$\forall i\in\{0,\ldots,k-1\}\left(\forall b\in\mathcal{D}_B\left(\varphi_{\mathcal{D}_B}(b)=i\Leftrightarrow b\in V_i\right)\right). \tag{6}$$

Пусть задан колоночный индекс $I_{R.B}$ для отношения $R\left(A,B,\ldots\right)$ с атрибутом B над доменом $\mathcal{D}_B$ и интервальная функция фрагментации $\varphi_{\mathcal{D}_B}$. Функция

$$\varphi_{I_{R.B}}:I_{R.B}\to\left\{0,\ldots,k-1\right\} \tag{7}$$

называется *доменно-интервальной функцией фрагментации* [2] для индекса $I_{R.B}$, если она удовлетворяет следующему условию:

$$\forall x\in I_{R.B}\left(\varphi_{I_{R.B}}(x)=\varphi_{\mathcal{D}_B}(x.B)\right). \tag{8}$$

Определим *i*-тый фрагмент $\left(i=0,\ldots,k-1\right)$ индекса $I_{R.B}$ следующим образом:

$$I_{R.B}^i=\left\{x\mid x\in I_{R.B};\ \varphi_{I_{R.B}}(x)=i\right\}. \tag{9}$$

Это означает, что в *i*-тый фрагмент попадают кортежи, у которых значение атрибута B принадлежит *i*-тому доменному интервалу. Будем называть фрагментацию, построенную таким образом, *доменно-интервальной*. Количество фрагментов k будем называть *степенью фрагментации*.

Доменно-интервальная фрагментация обладает следующими фундаментальными свойствами, вытекающими непосредственно из ее определения:

$$I_{R.B}=\bigcup_{i=0}^{k-1}I_{R.B}^i; \tag{10}$$

$$\forall i,j\in\{0,\ldots,\mathrm{k}-1\}\left(i\neq j\Rightarrow I_{R.B}^i\cap I_{R.B}^j=\varnothing\right). \tag{11}$$

Теорема 2. Пусть для колоночного индекса $I_{R.B}$ отношения $R(A,B,\ldots)$ задана доменно-интервальная фрагментация степени k. Тогда

$$\forall i \in \{0,\ldots,\mathrm{k}-1\}\left(\forall x \in I_{R.B}\left(x \in I_{R.B}^{i} \Leftrightarrow x.B \in V_i\right)\right). \tag{12}$$

Определение 3. Пусть для отношения $R(A,B,C,\ldots)$ заданы колоночные индексы $I_{R.B}$ и $I_{R.C}$. *Транзитивной фрагментацией* индекса $I_{R.C}$ относительно индекса $I_{R.B}$ называется фрагментация, задаваемая функцией $\varphi_{I_{R.C}}^{I_{R.B}} : I_{R.C} \to \{0,\ldots,k-1\}$, удовлетворяющей условию

$$\forall x \in I_{R.C}\left(\varphi_{I_{R.C}}^{I_{R.B}}(x) = \varphi_{I_{R.B}}\left(\sigma_{A=x.A}\left(I_{R.B}\right)\right)\right). \tag{13}$$

2. Декомпозиция операции группировки

Пусть задано отношение $R(A,B,C_1,\ldots,C_u,D_1,\ldots,D_w,\ldots)$ с виртуальным ключом *A*. Пусть для атрибутов $D_1,\ldots,D_w$ задана агрегирующая функция **agrf**. Пусть имеется колоночный индекс $I_{R.B}$. Пусть также имеются колоночные индексы:

$$I_{R.C_1},\ldots,I_{R.C_u};$$
$$I_{R.D_1},\ldots,I_{R.D_w}.$$

Пусть для индекса $I_{R.B}$ задана доменно-интервальная фрагментация степени k:

$$I_{R.B} = \bigcup_{i=0}^{k-1} I_{R.B}^{i}. \tag{14}$$

Пусть для индексов $I_{R.C_1},\ldots,I_{R.C_u}$ и $I_{R.D_1},\ldots,I_{R.D_w}$ задана транзитивная относительно $I_{R.B}$ фрагментация:

$$\forall j \in \{1,\ldots,u\}\left(I_{R.C_j} = \bigcup_{i=0}^{k-1} I_{R.C_j}^{i}\right); \tag{15}$$

$$\forall j \in \{1,\ldots,w\}\left(I_{R.D_j} = \bigcup_{i=0}^{k-1} I_{R.D_j}^{i}\right). \tag{16}$$

Положим

$$P_i = \pi_{A,F}\left(\gamma_{\min(A)\to A,B,C_1,\ldots,C_u,\mathrm{agrf}(D_1,\ldots,D_w)\to F}\left(I_{R.B}^{i} \bowtie I_{R.C_1}^{i} \bowtie \cdots \bowtie I_{R.C_u}^{i} \bowtie I_{R.D_1}^{i} \bowtie \cdots \bowtie I_{R.D_w}^{i}\right)\right) \tag{17}$$

для всех $i = 0,\ldots,k-1$. Определим

$$P = \bigcup_{i=0}^{k-1} P_i \quad . \tag{18}$$

Построим отношение $Q(B, C_1, \ldots, C_u, F)$ следующим образом:

$$Q = \left\{ \left(\&_R(p.A).B, \&_R(p.A).C_1, \ldots, \&_R(p.A).C_u, p.F \right) \mid p \in P \right\}. \tag{19}$$

Теорема 3. $Q = \gamma_{B,C_1,\ldots,C_u,\mathrm{agrf}(D_1,\ldots,D_w)\to F}(R)$.

Доказательство. Сначала докажем, что

$$\pi_{*\backslash F}\left(Q\right) = \pi_{*\backslash F}\left(\gamma_{B,C_1,\ldots,C_u,\mathrm{agrf}(D_1,\ldots,D_w)\to F}\left(R\right)\right). \tag{20}$$

Для этого нам достаточно доказать справедливость следующих двух утверждений:

$$\pi_{*\backslash F}\left(Q\right) \subset \pi_{*\backslash F}\left(\gamma_{B,C_1,\ldots,C_u,\mathrm{agrf}(D_1,\ldots,D_w)\to F}\left(R\right)\right); \tag{21}$$

и

$$\pi_{*\backslash F}\left(Q\right) \supset \pi_{*\backslash F}\left(\gamma_{B,C_1,\ldots,C_u,\mathrm{agrf}(D_1,\ldots,D_w)\to F}\left(R\right)\right). \tag{22}$$

Справедливость утверждения (21) непосредственно следует из (19) и (1). Покажем справедливость утверждения (22). Пусть

$$(b, c_1, \ldots, c_u) \in \pi_{*\backslash F}\left(\gamma_{B,C_1,\ldots,C_u,\mathrm{agrf}(D_1,\ldots,D_w)\to F}\left(R\right)\right).$$

Это означает, что

$$T\left(\sigma_{B=b\wedge C_1=c_1\wedge\ldots\wedge C_u=c_u}\left(\gamma_{\min(A)\to A,B,C_1,\ldots,C_u}\left(R\right)\right)\right) = 1. \tag{23}$$

Положим

$$(a) \in \pi_A\left(\sigma_{B=b\wedge C_1=c_1\wedge\ldots\wedge C_u=c_u}\left(\gamma_{\min(A)\to A,B,C_1,\ldots,C_u}\left(R\right)\right)\right). \tag{24}$$

С учетом (4) отсюда получаем, что существуют

$$x_B \in I_{R.B},\ x_{C_1} \in I_{R.C_1}, \ldots,\ x_{C_u} \in I_{R.C_u}$$

такие, что

$$(x_B.A = a \wedge x_B.B = b) \wedge (x_{C_1}.A = a \wedge x_{C_1}.C_1 = c_1) \cdots \wedge (x_{C_u}.A = a \wedge x_{C_u}.C_u = c_u). \tag{25}$$

Пусть в контексте разбиения (5) $b \in V_l$ $(l \in \{0, \ldots, k-1\})$. Тогда из (12) и из (25) следует, что

$$x_B \in I_{R.B}^l. \tag{26}$$

По определению 3 транзитивной фрагментации из (25) и (26) получаем

$$x_{C_1} \in I_{R.C_1}^l, \cdots, x_{C_u} \in I_{R.C_u}^l. \tag{27}$$

Сопоставляя (17), (24), (25) и (27), получаем, что $(a) \in \pi_A(P_l)$. С учетом (18) отсюда следует $(a) \in \pi_A(P)$. Вместе с (19) это дает

$$\left(\&_R(a).B, \&_R(a).C_1, \ldots, \&_R(a).C_u\right) \in \pi_{*\backslash F}\left(Q\right),$$

откуда с учетом (24) получаем $(b, c_1, \ldots, c_u) \in \pi_{*\backslash F}\left(Q\right)$. Таким образом (22), а вместе с ним и (20) имеет место.

Для завершения доказательства теоремы нам достаточно показать, что для любых $q \in Q$ и $g \in \gamma_{B,C_1,\ldots,C_u,\mathrm{agrf}(D_1,\ldots,D_w)\to F}(R)$ справедливо:

$$\left(q.B = g.B \wedge q.C_1 = g.C_1 \wedge \ldots \wedge q.C_1 = g.C_1\right) \Rightarrow \left(q.F = g.F\right). \qquad (28)$$

Докажем (28). Пусть

$$q = (b, c_1, \ldots, c_u, f) \in Q \qquad (29)$$

и

$$g = (b, c_1, \ldots, c_u, f') \in \gamma_{B,C_1,\ldots,C_u,\mathrm{agrf}(D_1,\ldots,D_w)\to F}(R). \qquad (30)$$

Такие q и g существуют в силу (20). Для доказательства (28) достаточно показать, что $f = f'$.

В соответствии с (19) существует

$$p = (a, f) \in P \qquad (31)$$

такой, что

$$(a, b, c_1, \ldots, c_u, \ldots) \in R. \qquad (32)$$

Из (31) и (18) следует, что существует l, такое что

$$p = (a, f) \in P^l. \qquad (33)$$

Отсюда, с учетом (17) получаем

$$(a, b', c_1', \ldots, c_u') \in I^l_{R.B} \bowtie I^l_{R.C_1} \bowtie \cdots \bowtie I^l_{R.C_u}.$$

В силу свойств транзитивной фрагментации и колоночных индексов с учетом (32) отсюда следует, что

$$(a, b, c_1, \ldots, c_u) \in I^l_{R.B} \bowtie I^l_{R.C_1} \bowtie \cdots \bowtie I^l_{R.C_u}. \qquad (34)$$

Вместе с (33) и (17) это дает

$$f = \pi_F\left(\gamma_{B,C_1,\ldots,C_u,\mathrm{agrf}(D_1,\ldots,D_w)\to F}\left(\sigma_{B=b\wedge C_1=c_1\wedge\ldots\wedge C_u=c_u}\left(I^l_{R.B} \bowtie I^l_{R.C_1} \bowtie \cdots \bowtie I^l_{R.C_u} \bowtie I^l_{R.D_1} \bowtie \cdots \bowtie I^l_{R.D_w}\right)\right)\right). \qquad (35)$$

С другой стороны, из (30) непосредственно следует, что

$$f' = \pi_F\left(\gamma_{B,C_1,\ldots,C_u,\mathrm{agrf}(D_1,\ldots,D_w)\to F}\left(\sigma_{B=b\wedge C_1=c_1\wedge\ldots\wedge C_u=c_u}(R)\right)\right). \qquad (36)$$

Исходя из формул (35) и (36), условие $f = f'$ равносильно условию

$$\sigma_{B=b\wedge C_1=c_1\wedge\ldots\wedge C_u=c_u}\left(I^l_{R.B} \bowtie I^l_{R.C_1} \bowtie \cdots \bowtie I^l_{R.C_u} \bowtie I^l_{R.D_1} \bowtie \cdots \bowtie I^l_{R.D_w}\right) = \\ = \sigma_{B=b\wedge C_1=c_1\wedge\ldots\wedge C_u=c_u}\left(\pi_{B,C_1,\ldots,C_u,D_1,\ldots,D_w}(R)\right). \qquad (37)$$

Пусть

$$(b, c_1, \ldots, c_u, d_1, \ldots, d_w) \in \sigma_{B=b\wedge C_1=c_1\wedge\ldots\wedge C_u=c_u}\left(I^l_{R.B} \bowtie I^l_{R.C_1} \bowtie \cdots \bowtie I^l_{R.C_u} \bowtie I^l_{R.D_1} \bowtie \cdots \bowtie I^l_{R.D_w}\right).$$

Это означает, что существует a' такой, что

$$(a', b) \in I^l_{R.B} \subset I_{R.B};$$

$$(a', b) \in I^l_{R.C_1} \subset I_{R.C_1};$$

$$\ldots$$

$$(a',b) \in I^l_{R.C_u} \subset I_{R.C_u};$$
$$(a',b) \in I^l_{R.D_1} \subset I_{R.D_1};$$
$$\ldots$$
$$(a',b) \in I^l_{R.D_w} \subset I_{R.D_w}.$$
Отсюда в силу свойств колоночного индекса
$$(a',b,c_1,\ldots,c_u,d_1,\ldots,d_w) \in \pi_{A,B,C_1,\ldots,C_u,D_1,\ldots,D_w}(R),$$
И следовательно
$$(b,c_1,\ldots,c_u,d_1,\ldots,d_w) \in \sigma_{B=b \wedge C_1=c_1 \wedge \ldots \wedge C_u=c_u}\left(\pi_{B,C_1,\ldots,C_u,D_1,\ldots,D_w}(R)\right).$$
Пусть теперь
$$(b,c_1,\ldots,c_u,d'_1,\ldots,d'_w) \in \sigma_{B=b \wedge C_1=c_1 \wedge \ldots \wedge C_u=c_u}\left(\pi_{B,C_1,\ldots,C_u,D_1,\ldots,D_w}(R)\right). \qquad (38)$$
В силу свойств транзитивной фрагментации и колоночных индексов, принимая во внимание (34), получаем

$$\left(b,c_1,\ldots,c_u,d'_1,\ldots,d'_w\right) \in \left(I^l_{R.B} \bowtie I^l_{R.C_1} \bowtie \cdots \bowtie I^l_{R.C_u} \bowtie I^l_{R.D_1} \bowtie \cdots \bowtie I^l_{R.D_w}\right),$$
откуда немедленно следует
$$\left(b,c_1,\ldots,c_u,d'_1,\ldots,d'_w\right) \in \sigma_{B=b \wedge C_1=c_1 \wedge \ldots \wedge C_u=c_u}\left(I^l_{R.B} \bowtie I^l_{R.C_1} \bowtie \cdots \bowtie I^l_{R.C_u} \bowtie I^l_{R.D_1} \bowtie \cdots \bowtie I^l_{R.D_w}\right).$$
Таким образом (37), а, значит, и (28) имеют место. *Теорема доказана*.

Заключение

В статье было представлено формальное описание метода декомпозиции операции группировки на основе доменно-интервальной и транзитивной фрагментации колоночных индексов. Доказана корректность метода. Практическая ценность предложенного метода декомпозиции заключается в том, что ресурсоемкие вычисления могут производиться независимо над соответствующими фрагментами. Для операции группировки это – вычисление P_i по формуле (17).

Работа выполнена при финансовой поддержке Минобрнауки РФ в рамках ФЦП «Исследования и разработки по приоритетным направлениям развития научно-технологического комплекса России на 2014—2020 годы» (Госконтракт № 14.574.21.0035).

Библиографический список

1. Соколинский, Л.Б. Параллельные машины баз данных / Л.Б. Соколинский // Природа. — 2001. — № 8. — С. 10–17.
2. Соколинский, Л.Б. Параллельные системы баз данных / Л.Б. Соколинский

— Москва: Издательство Московского государственного университета, 2013. — 184 с.

3. Sokolinsky, L.B. Design and Evaluation of Database Multiprocessor Architecture with High Data Availability / L.B. Sokolinsky // Proceedings of the 12th International workshop on database and expert systems applications. — IEEE Computer Society, 2001. — P. 115–120.

4. Pan, C.S. Taming Elephants, or How to Embed Parallelism into PostgreSQL / C.S. Pan, M.L. Zymbler // Lecture Notes in Computer Science. — 2013. — Vol. 8055, Part 1. — P. 153–164.

5. Костенецкий, П.С. Моделирование иерархических многопроцессорных систем баз данных / П.С. Костенецкий, Л.Б. Соколинский // Программирование. — 2013. — Т. 39, № 1. — С. 3–22.

6. Fang, J. Sesame: A User-Transparent Optimizing Framework for Many-Core Processors / J. Fang, A.L. Varbanescu, H. Sips // Proceedings of the 13th IEEE/ACM International Symposium on Cluster, Cloud and Grid Computing (CCGrid2013), May 13–16, 2013, Delft, Netherlands. — IEEE, 2013. — P. 70–73.

7. Breß, S. Efficient Co-Processor Utilization in Database Query Processing / S. Breß, F. Beier, H. Rauhe, et al. // Information Systems. — 2013. — Vol. 38, No. 8. — P. 1084–1096.

8. Scherger, M. Design of an In-Memory Database Engine Using Intel Xeon Phi Coprocessors / M. Scherger // Proceedings of the International Conference on Parallel and Distributed Processing Techniques and Applications (PDPTA’14), July 21–24, 2014, Las Vegas, USA. — CSREA Press, 2014. — P. 21–27.

9. Беседин, К.Ю. Моделирование обработки запросов на гибридных вычислительных системах с многоядерными сопроцессорами и графическими ускорителями / К.Ю. Беседин, П.С. Костенецкий // Программные системы: теория и приложения. — 2014. — Т. 5, № 1-1 (19). — С. 91–110.

10. Иванова, Е.В. Исследование эффективности использования фрагментированных колоночных индексов при выполнении операции естественного соединения с использованием многоядерных ускорителей / Е.В. Иванова // Параллельные вычислительные технологии (ПаВТ'2015). Труды международной научной конференции. — Челябинск: Издательский центр ЮУрГУ, 2015. — С. 393-398.

11. Иванова, Е.В. Декомпозиция операций пересечения и соединения на основе доменно-интервальной фрагментации колоночных индексов / Е.В. Иванова, Л.Б. Соколинский // Вестник Южно-Уральского государственного университета. Серия: Вычислительная математика и информатика. — 2015. — Т. 4, № 1. — С. 44-56.

12. Гарсиа-Молина, Г. Системы баз данных. Полный курс. / Г. Гарсиа-Молина, Дж. Ульман, Дж. Уидом — М.: Издательский дом «Вильямс». — 2004. — 1088 с.