

ДЕКОМПОЗИЦИЯ ОПЕРАЦИЙ ПЕРЕСЕЧЕНИЯ И СОЕДИНЕНИЯ НА ОСНОВЕ ДОМЕННО-ИНТЕРВАЛЬНОЙ ФРАГМЕНТАЦИИ КОЛОНОЧНЫХ ИНДЕКСОВ

Е.В. Иванова, Л.Б. Соколинский

Статья посвящена вопросам декомпозиции реляционных операций путем использования распределенных колоночных индексов с доменно-интервальной фрагментацией. Такая декомпозиция позволяет организовать параллельное выполнение ресурсоемких реляционных операций без обменов данными между процессорными ядрами. Все фрагменты колоночного индекса хранятся в оперативной памяти в сжатом виде. При параллельном выполнении реляционной операции упакованные фрагменты индексов входных отношений загружаются на различные процессорные ядра, где происходят их распаковка, выполнение реляционной операции над фрагментами и упаковка частичного результата, представляющего собой наборы ключей. Затем частичные результаты объединяются в результирующий набор ключей, с использованием которого СУБД собирает результирующее отношение. Указанный подход 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, декомпозиция реляционных операций.

Введение

В настоящее время научно-практическая деятельность человека выдвигает все новые масштабные задачи, требующие обработки сверхбольших баз данных. Согласно прогнозам аналитической компании IDC, количество данных в мире удваивается каждые два года и к 2020 г. достигнет 44 Зеттабайт, или 44 триллионов гигабайт [1]. При этом современные технологии баз данных не могут обеспечить обработку столь крупных объемов данных. По оценке IDC в 2013 г. из всего объема существующих данных потенциально полезны 22%, из которых менее 5% были подвергнуты анализу. К 2020 году процент потенциально полезных данных может вырасти до 35%, преимущественно за счет данных от встроенных систем.

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 [2–6].

Традиционным подходом к организации хранения баз данных является строковое представление данных. Однако при выполнении типичных аналитических запросов к таблицам требуется считывать только небольшую часть полей в строках этих таблиц, поэтому строковое представление в этом случае оказывается неэффективным. Причиной этого является считывание с диска «лишних» полей в дополнение к тем полям, которые необходимы в данном запросе [7]. Возможным решением этой проблемы может быть использование механизмов колоночного представления данных, позволяющих получить на порядок лучшую производительность при обработке аналитических запросов [8]. Колоночное представление данных заключается в том, что данные хранятся не по строкам, а

по колонкам. Это означает, что с точки зрения SQL-клиента данные представлены в виде таблиц, но физически эти таблицы являются совокупностью колонок, каждая из которых представляет собой таблицу из одного поля. Дополнительным преимуществом колоночного представления является возможность использования эффективных алгоритмов сжатия данных, поскольку в одной колонке таблицы содержатся данные одного типа. Сжатие может привести к повышению производительности на порядок, поскольку меньше времени занимают операции ввода-вывода.

В последние годы основным способом наращивания производительности процессоров является увеличение количества ядер, а не тактовой частоты, и эта тенденция, вероятно, сохранится [9]. 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 [10–12].

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, с использованием колоночного представления и сжатия данных. Для решения этой задачи в работах [13, 14] 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 рассмотрены вопросы декомпозиции следующих реляционных операций: пересечения, естественного соединения и тета-соединения. Для обозначения реляционных операций в статье используется нотация, заимствованная из монографии [15]. Символом \circ обозначается конкатенация кортежей.

Статья организована следующим образом. В разделе 1 приведено формальное описание колоночного индекса и доменно-интервальной фрагментации, а также доказываются две вспомогательные теоремы. Разделы 2–4 посвящены описаниям методов декомпозиции соответствующих реляционных операций на основе колоночных индексов с доменно-интервальной фрагментацией. Для каждой операции дается формальное описание метода декомпозиции и доказывается теорема, подтверждающая его корректность. В заключении суммируются полученные результаты, делаются итоговые выводы и даются направления дальнейших исследований.

1. Колоночный индекс и доменно-интервальная фрагментация

Под $R(A^*, B_1, \dots, B_u)$ будем понимать отношение R с первичным ключом A и атрибутами B_1, \dots, B_u , представляющее собой множество кортежей длины $u + 1$ вида (a, b_1, \dots, b_u) , где $a \in \mathbb{Z}_{\geq 0}$ и $\forall j \in \{1, \dots, u\} (b_j \in \mathcal{D}_{B_j})$. Здесь \mathcal{D}_{B_j} - домен атрибута B_j . Через $r.B_j$ будем обозначать значение атрибута B_j , через $r.A$ - значение *первичного ключа* в кортеже r : $r = (r.A, r.B_1, \dots, r.B_u)$. *Первичный ключ* отношения R обладает свойством

$\forall r', r'' \in R (r' \neq r'' \Leftrightarrow r'.A \neq r''.A)$. Под *адресом кортежа* r мы будем понимать значение первичного ключа этого кортежа. Для получения кортежа отношения R по его адресу будем использовать *функцию разыменования* $\&_R : \forall r \in R (\&_R(r.A) = r)$.

Определение 1. Пусть задано отношение $R(A^*, B, \dots)$, $T(R) = n$. Пусть на множестве \mathfrak{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:

$$T(I_{R,B}) = n \text{ и } \pi_A(I_{R,B}) = \pi_A(R); \quad (1)$$

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

$$\forall r \in R (\forall x \in I_{R,B} (r.A = x.A \Rightarrow r.B = x.B)). \quad (3)$$

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 $R(A^*, B, \dots)$. Пусть для отношения R задан колоночный индекс $I_{R,B}$. Тогда

$$\pi_B(I_{R,B}) = \pi_B(R). \quad (4)$$

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

Доказательство. Возьмем произвольное $b \in \mathfrak{D}_B$. Пусть $T(\sigma_{B=b}(R)) = k$. Без ограничения общности мы можем считать, что $\forall r \in R (r.A < k \Leftrightarrow r.B = b)$. Тогда из (1) и (3) следует, что $\forall x \in I_{R,B} (x.A < k \Leftrightarrow x.B = b)$. Откуда получаем $T(\sigma_{B=b}(I_{R,B})) = k$. Таким образом (4) имеет место. *Теорема доказана.*

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

$$\left. \begin{aligned} &V_0 = [v_0; v_1], V_1 = (v_1; v_2], \dots, V_{k-1} = (v_{k-1}; v_k]; \\ &v_0 < v_1 < \dots < v_k; \\ &\mathfrak{D}_B = \bigcup_{i=0}^{k-1} V_i. \end{aligned} \right\}. \quad (5)$$

Функция $\varphi_{\mathfrak{D}_B} : \mathfrak{D}_B \rightarrow \{0, \dots, k-1\}$ называется *интервальной функцией фрагментации* для домена \mathfrak{D}_B , если она удовлетворяет следующему условию:

$$\forall i \in \{0, \dots, k-1\} (\forall b \in \mathfrak{D}_B (\varphi_{\mathfrak{D}_B}(b) = i \Leftrightarrow b \in V_i)). \quad (6)$$

Определение 3. Пусть задан колоночный индекс $I_{R.B}$ для отношения $R(A^*, B, \dots)$ с атрибутом B над доменом \mathfrak{D}_B и интервальная функция фрагментации $\varphi_{\mathfrak{D}_B}$. Функция

$$\varphi_{I_{R.B}} : I_{R.B} \rightarrow \{0, \dots, k-1\} \quad (7)$$

называется *доменно-интервальной функцией фрагментации* [3] для индекса $I_{R.B}$, если она удовлетворяет следующему условию:

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

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

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

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:

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

$$\forall i, j \in \{0, \dots, k-1\} (i \neq j \Rightarrow I_{R.B}^i \cap I_{R.B}^j = \emptyset). \quad (11)$$

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

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

Доказательство. Сначала докажем, что

$$\forall i \in \{0, \dots, k-1\} \left(\forall x \in I_{R.B} (x \in I_{R.B}^i \Rightarrow x.B \in V_i) \right). \quad (13)$$

Пусть $x \in I_{R.B}^i$. Тогда из (9) следует $\varphi_{I_{R.B}}(x) = i$. С учетом (8) получаем $\varphi_{\mathfrak{D}_B}(x.B) = i$. Отсюда и из (6) следует $x.B \in V_i$, то есть (13) имеет место. Теперь докажем, что

$$\forall i \in \{0, \dots, k-1\} \left(\forall x \in I_{R.B} (x.B \in V_i \Rightarrow x \in I_{R.B}^i) \right). \quad (14)$$

Пусть $x \in I_{R.B}$ и $x.B \in V_i$. Тогда из (6) следует, что $\varphi_{\mathfrak{D}_B}(x.B) = i$. С учетом (8) получаем $\varphi_{\mathfrak{D}_B}(x.B) = \varphi_{I_{R.B}}(x) = i$. Отсюда и из (9) следует, что $x \in I_{R.B}^i$, то есть (14) имеет место. *Теорема доказана.*

2. Декомпозиция операции пересечения

Пусть заданы два отношения $R(A^*, B_1, \dots, B_u)$ и $S(A^*, B_1, \dots, B_u)$, имеющие одинаковый набор атрибутов. Мы предполагаем, что $\delta(\pi_{B_1, \dots, B_u}(R)) = \pi_{B_1, \dots, B_u}(R)$ и $\delta(\pi_{B_1, \dots, B_u}(S)) = \pi_{B_1, \dots, B_u}(S)$, то есть $\pi_{B_1, \dots, B_u}(R)$ и $\pi_{B_1, \dots, B_u}(S)$ не содержат дубликатов. Пусть имеется два набора колоночных индексов по атрибутам B_1, \dots, B_u :

$$I_{R.B_1}, \dots, I_{R.B_u};$$

$$I_{S.B_1}, \dots, I_{S.B_u}.$$

Пусть для всех этих индексов задана доменно-интервальная фрагментация степени k :

$$I_{R.B_j} = \bigcup_{i=0}^{k-1} I_{R.B_j}^i; \quad (15)$$

$$I_{S.B_j} = \bigcup_{i=0}^{k-1} I_{S.B_j}^i. \quad (16)$$

Положим

$$P_j^i = \pi_{I_{R.B_j}^i.A \rightarrow A_R, I_{S.B_j}^i.A \rightarrow A_S} \left(I_{R.B_j}^i \bowtie_{(I_{R.B_j}^i.B_j = I_{S.B_j}^i.B_j)} I_{S.B_j}^i \right) \quad (17)$$

для всех $i = 0, \dots, k-1$ и $j = 1, \dots, u$. Определим

$$P_j = \bigcup_{i=0}^{k-1} P_j^i. \quad (18)$$

Положим

$$P = \bigcap_{j=1}^u P_j. \quad (19)$$

Определим

$$Q = \{r \mid r \in R \wedge r.A \in \pi_{A_R}(P)\}. \quad (20)$$

Теорема 3. $\pi_{B_1, \dots, B_u}(Q) = \pi_{B_1, \dots, B_u}(R) \cap \pi_{B_1, \dots, B_u}(S)$.

Доказательство. Сначала докажем, что

$$\pi_{B_1, \dots, B_u}(Q) \subset \pi_{B_1, \dots, B_u}(R) \cap \pi_{B_1, \dots, B_u}(S). \quad (21)$$

Пусть

$$(a, b_1, \dots, b_u) \in Q. \quad (22)$$

Из (20) следует, что

$$(a, b_1, \dots, b_u) = r \in R \quad (23)$$

и

$$a = r.A \in \pi_{A_R}(P). \quad (24)$$

Отсюда следует, что $\exists p \in P(p.A_R = a \wedge p.A_S = a')$. С учетом (19) получаем, что $\forall j \in \{1, \dots, u\} (\exists p \in P_j(p.A_R = a \wedge p.A_S = a'))$. С учетом (18) отсюда получаем $\forall j \in \{1, \dots, u\} (\exists i \in \{0, \dots, k-1\} (\exists p \in P_j^i(p.A_R = a \wedge p.A_S = a')))$. Отсюда и из (15)-(17) следует, что

$$\forall j \in \{1, \dots, u\} (\exists x \in I_{R.B_j} (\exists y \in I_{S.B_j} (x.A = a \wedge x.B_j = y.B_j \wedge y.A = a'))).$$

По определению колоночного индекса отсюда получаем

$$\forall j \in \{1, \dots, u\} (\exists \tilde{r} \in R (\exists \tilde{s} \in S (\tilde{r}.A = a \wedge \tilde{r}.B_j = \tilde{s}.B_j \wedge \tilde{s}.A = a'))).$$

Поскольку A является первичным ключом в R и S , с учетом (23) отсюда следует $(a', b_1, \dots, b_u) \in S$, то есть $(b_1, \dots, b_u) \in \pi_{B_1, \dots, B_u}(R) \cap \pi_{B_1, \dots, B_u}(S)$. Таким образом (21) имеет место.

Теперь докажем, что

$$\pi_{B_1, \dots, B_u}(R) \cap \pi_{B_1, \dots, B_u}(S) \subset \pi_{B_1, \dots, B_u}(Q). \quad (25)$$

Пусть

$$(a, b_1, \dots, b_u) = r \in R \quad (26)$$

и

$$(a', b_1, \dots, b_u) = s \in S. \quad (27)$$

Тогда по определению колоночного индекса с учетом (4) имеем

$$\forall j \in \{1, \dots, u\} \left(\exists x \in I_{R.B_j} \left(\exists y \in I_{S.B_j} \left(x.A = a \wedge x.B_j = b_j = y.B_j \wedge y.A = a' \right) \right) \right).$$

На основе (12) отсюда получаем

$$\forall j \in \{1, \dots, u\} \left(\exists i \in \{0, \dots, k-1\} \left(\exists x \in I_{R.B_j}^i \left(\exists y \in I_{S.B_j}^i \left(x.A = a \wedge x.B_j = y.B_j \wedge y.A = a' \right) \right) \right) \right).$$

С учетом (17) отсюда следует

$$\forall j \in \{1, \dots, u\} \left(\exists i \in \{0, \dots, k-1\} \left(\exists p \in P_j^i \left(p.A_R = a \wedge p.A_S = a' \right) \right) \right).$$

Применяя (18) получаем

$$\forall j \in \{1, \dots, u\} \left(\exists p \in P_j \left(p.A_R = a \wedge p.A_S = a' \right) \right).$$

Учитывая (19) отсюда имеем $(a, a') \in P$. Вместе с (20) и (26) это дает

$$(a, b_1, \dots, b_u) \in Q,$$

то есть (25) имеет место. Теорема доказана.

3. Декомпозиция операции естественного соединения

Пусть заданы два отношения

$$R(A^*, B_1, \dots, B_u, C_1, \dots, C_v)$$

и

$$S(A^*, B_1, \dots, B_u, D_1, \dots, D_w).$$

Пусть имеется два набора колоночных индексов по атрибутам B_1, \dots, B_u :

$$I_{R.B_1}, \dots, I_{R.B_u};$$

$$I_{S.B_1}, \dots, I_{S.B_u}.$$

Пусть для всех этих индексов задана доменно-интервальная фрагментация степени k :

$$I_{R.B_j} = \bigcup_{i=0}^{k-1} I_{R.B_j}^i; \quad (28)$$

$$I_{S.B_j} = \bigcup_{i=0}^{k-1} I_{S.B_j}^i. \quad (29)$$

Положим

$$P_j^i = \pi_{I_{R.B_j}^i.A \rightarrow A_R, I_{S.B_j}^i.A \rightarrow A_S} \left(I_{R.B_j}^i \bowtie_{I_{R.B_j}^i.B_j = I_{S.B_j}^i.B_j} I_{S.B_j}^i \right) \quad (30)$$

для всех $i = 0, \dots, k-1$ и $j = 1, \dots, u$. Определим

$$P_j = \bigcup_{i=0}^{k-1} P_j^i. \quad (31)$$

Положим

$$P = \bigcap_{j=1}^u P_j. \quad (32)$$

Определим

$$Q = \{r \circ (s.D_1, \dots, s.D_w) \mid r \in R \wedge s \in S \wedge (r.A, s.A) \in P\}. \quad (33)$$

Теорема 4. $\pi_{*A}(Q) = \pi_{*A}(R) \bowtie \pi_{*A}(S)$.

Доказательство. Сначала докажем, что

$$\pi_{*A}(Q) \subset \pi_{*A}(R) \bowtie \pi_{*A}(S). \quad (34)$$

Пусть

$$(a, b_1, \dots, b_u, c_1, \dots, c_v, d_1, \dots, d_w) \in Q. \quad (35)$$

Из (33) следует, что существуют кортежи r и s такие, что

$$(a, b_1, \dots, b_u, c_1, \dots, c_v) = r \in R, \quad (36)$$

$$(a', b'_1, \dots, b'_u, d_1, \dots, d_w) = s \in S \quad (37)$$

и

$$(r.A, s.A) \in P. \quad (38)$$

Отсюда следует, что $\exists p \in P(p.A_R = a \wedge p.A_S = a')$. С учетом (32) получаем, что $\forall j \in \{1, \dots, u\} (\exists p \in P_j(p.A_R = a \wedge p.A_S = a'))$. С учетом (31) отсюда получаем $\forall j \in \{1, \dots, u\} (\exists i \in \{0, \dots, k-1\} (\exists p \in P_j^i(p.A_R = a \wedge p.A_S = a')))$. Отсюда и из (28)–(30) следует, что

$$\forall j \in \{1, \dots, u\} \left(\exists x \in I_{R.B_j} \left(\exists y \in I_{S.B_j} (x.A = a \wedge x.B_j = y.B_j \wedge y.A = a') \right) \right).$$

По определению колоночного индекса отсюда получаем

$$\forall j \in \{1, \dots, u\} \left(\exists \tilde{r} \in R \left(\exists \tilde{s} \in S (\tilde{r}.A = a \wedge \tilde{r}.B_j = \tilde{s}.B_j \wedge \tilde{s}.A = a') \right) \right).$$

Поскольку A является первичным ключом в R и S , с учетом (36) и (37) отсюда следует $(a', b'_1, \dots, b'_u, d_1, \dots, d_w) \in S$, то есть $(b_1, \dots, b_u, c_1, \dots, c_v, d_1, \dots, d_w) \in \pi_{*A}(R) \bowtie \pi_{*A}(S)$. Таким образом (34) имеет место.

Теперь докажем, что

$$\pi_{*A}(R) \bowtie \pi_{*A}(S) \subset \pi_{*A}(Q). \quad (39)$$

Пусть

$$(a, b_1, \dots, b_u, c_1, \dots, c_v) = r \in R \quad (40)$$

и

$$(a', b'_1, \dots, b'_u, d_1, \dots, d_w) = s \in S. \quad (41)$$

Тогда по определению колоночного индекса с учетом (4) имеем

$$\forall j \in \{1, \dots, u\} \left(\exists x \in I_{R.B_j} \left(\exists y \in I_{S.B_j} (x.A = a \wedge x.B_j = b_j = y.B_j \wedge y.A = a') \right) \right).$$

На основе (12) отсюда получаем

$$\forall j \in \{1, \dots, u\} \left(\exists i \in \{0, \dots, k-1\} \left(\exists x \in I_{R.B_j}^i \left(\exists y \in I_{S.B_j}^i (x.A = a \wedge x.B_j = y.B_j \wedge y.A = a') \right) \right) \right).$$

С учетом (30) отсюда следует

$$\forall j \in \{1, \dots, u\} \left(\exists i \in \{0, \dots, k-1\} (\exists p \in P_j^i(p.A_R = a \wedge p.A_S = a')) \right).$$

Применяя (31) получаем

$$\forall j \in \{1, \dots, u\} (\exists p \in P_j(p.A_R = a \wedge p.A_S = a')).$$

Учитывая (32) отсюда имеем $(a, a') \in P$. Вместе с (33), (40) и (41) это дает

$$(a, b_1, \dots, b_u, c_1, \dots, c_v, d_1, \dots, d_w) \in Q,$$

то есть (39) имеет место. Теорема доказана.

4. Декомпозиция операции тета-соединения

Пусть заданы два отношения $R(A^*, B_1, \dots, B_u)$ и $S(A^*, C_1, \dots, C_v)$. Пусть также заданы функции $f: \pi_{B_1, \dots, B_u}(R) \rightarrow \mathbb{R}$ и $g: \pi_{C_1, \dots, C_v}(S) \rightarrow \mathbb{R}$. Определим $\bar{R} = \pi_{A, f(r.B_1, \dots, r.B_u) \rightarrow F}(R)$ и $\bar{S} = \pi_{A, g(s.C_1, \dots, s.C_v) \rightarrow G}(S)$. Пусть имеются колоночные индексы $I_{\bar{R}.F}$ и $I_{\bar{S}.G}$. Пусть для этих индексов задана доменно-интервальная фрагментация степени k :

$$I_{\bar{R}.F} = \bigcup_{i=0}^{k-1} I_{\bar{R}.F}^i; \quad (42)$$

$$I_{\bar{S}.G} = \bigcup_{i=0}^{k-1} I_{\bar{S}.G}^i. \quad (43)$$

Положим

$$\tilde{P}_i = \pi_{I_{\bar{R}.F}^i.A \rightarrow A_R, I_{\bar{S}.G}^i.A \rightarrow A_S} \left(I_{\bar{R}.F}^i \bowtie_{F < G} I_{\bar{S}.G}^i \right). \quad (44)$$

Определим

$$M_i = \pi_{A \rightarrow A_R} \left(I_{\bar{R}.F}^i \right); \quad (45)$$

$$N_i = \pi_{A \rightarrow A_S} \left(I_{\bar{S}.G}^i \right) \quad (46)$$

для всех $i = 0, \dots, k-1$ и

$$\bar{P}_i = M_i \times \bigcup_{l=i+1}^{k-1} N_l. \quad (47)$$

Положим

$$P_i = \tilde{P}_i \cup \bar{P}_i; \quad (48)$$

$$P = \bigcup_{i=0}^{k-1} P_i. \quad (49)$$

Определим

$$Q = \{r \circ s \mid r \in R \wedge s \in S \wedge (r.A, s.A) \in P\}. \quad (50)$$

Теорема 5. $\pi_{*\setminus R, A, S, A}(Q) = \pi_{*\setminus R}(R) \bowtie_{f(r.B_1, \dots, r.B_u) < g(s.C_1, \dots, s.C_v)} \pi_{*\setminus A}(S)$.

Доказательство. Сначала докажем, что

$$\pi_{*\setminus A}(Q) \subset \pi_{*\setminus A}(R) \bowtie_{f(r.B_1, \dots, r.B_u) < g(s.C_1, \dots, s.C_v)} \pi_{*\setminus A}(S). \quad (51)$$

Пусть

$$(a, b_1, \dots, b_u, a', c_1, \dots, c_v) \in Q. \quad (52)$$

Из (50) следует, что существуют кортежи r и s такие, что

$$(a, b_1, \dots, b_u) = r \in R, \quad (53)$$

$$(a', c_1, \dots, c_v) = s \in S \quad (54)$$

и

$$(r.A, s.A) \in P. \quad (55)$$

Отсюда следует, что $\exists p \in P(p.A_R = a \wedge p.A_S = a')$. Отсюда и из (49) получаем, что существует $i \in \{0, \dots, k-1\}$, такой, что

$$p \in P_i \wedge p.A_R = a \wedge p.A_S = a'. \quad (56)$$

В силу (48) имеем $p \in \tilde{P}_i \vee p \in \bar{P}_i$. Рассмотрим сначала случай, когда $p \in \tilde{P}_i$. Тогда из (56) и (42)–(44) следует, что

$$\exists x \in I_{\bar{R},F} \left(\exists y \in I_{\bar{S},G} (x.A = a \wedge x.F < y.G \wedge y.A = a') \right).$$

По определению колоночного индекса отсюда получаем

$$\exists \tilde{r} \in R \left(\exists \tilde{s} \in S \left(\tilde{r}.A = a \wedge \tilde{s}.A = a' \wedge f(\tilde{r}.B_1, \dots, \tilde{r}.B_u) < g(\tilde{s}.C_1, \dots, \tilde{s}.C_v) \right) \right). \quad (57)$$

Поскольку A является первичным ключом в R и S , с учетом (53) и (54) отсюда следует, что $\tilde{r} = (a, b_1, \dots, b_u)$, $\tilde{s} = (a', c_1, \dots, c_v)$ и $(b_1, \dots, b_u, c_1, \dots, c_v) \in \pi_{*A}(R) \underset{f(r.B_1, \dots, r.B_u) < g(s.C_1, \dots, s.C_v)}{\bowtie} \pi_{*A}(S)$.

Таким образом, (51) имеет место.

Пусть теперь $p \in \bar{P}_i$. Тогда из (56), и (45)–(47) следует, что существуют $i \in \{0, \dots, k-1\}$ и $l \in \{i+1, \dots, k-1\}$, такие, что

$$\exists x \in I_{\bar{R},F}^i (x.A = a), \quad (58)$$

$$\exists y \in I_{\bar{S},G}^l (y.A = a'). \quad (59)$$

Так как $l > i$, то по определению колоночного индекса и доменно-интервальной фрагментации из (58) и (59) получаем, что

$$\exists x \in I_{\bar{R},F} \left(\exists y \in I_{\bar{S},G} (x.A = a \wedge y.A = a' \wedge x.F < y.G) \right),$$

откуда следует (57), и, таким образом, (51) также имеет место.

Теперь докажем, что

$$\pi_{*A}(R) \underset{f(r.B_1, \dots, r.B_u) < g(s.C_1, \dots, s.C_v)}{\bowtie} \pi_{*A}(S) \subset \pi_{*R,A,S,A}(Q). \quad (60)$$

Пусть

$$(a, b_1, \dots, b_u) = r \in R, \quad (61)$$

$$(a', c_1, \dots, c_v) = s \in S, \quad (62)$$

и

$$f(b_1, \dots, b_u) < g(c_1, \dots, c_v). \quad (63)$$

Тогда по определению колоночного индекса с учетом (4) имеем

$$\exists x \in I_{\bar{R},F} \left(\exists y \in I_{\bar{S},G} (x.A = a \wedge x.F = f(b_1, \dots, b_u) < g(c_1, \dots, c_v) = y.G \wedge y.A = a') \right).$$

На основе (5)–(9) и (12) отсюда следует, что возможны только два следующих случая:

$$1) \quad \exists i \in \{0, \dots, k-1\} \left(x \in I_{\bar{R},F}^i \wedge y \in I_{\bar{S},G}^i \right); \quad (64)$$

$$2) \quad \exists i \in \{0, \dots, k-1\} \left(\exists l \in \{i+1, \dots, k-1\} \left(x \in I_{\bar{R},F}^i \wedge y \in I_{\bar{S},G}^l \right) \right). \quad (65)$$

Допустим сначала, что имеет место первый случай. Тогда с учетом (44) из (64) следует

$$\exists i \in \{0, \dots, k-1\} \left(\exists p \in \tilde{P}_i (p.A_R = a \wedge p.A_S = a') \right).$$

Учитывая (48) и (49) отсюда имеем $(a, a') \in P$. Вместе с (50), (61) и (62) это дает

$$(a, b_1, \dots, b_u, a', c_1, \dots, c_v) \in Q,$$

то есть (60) имеет место.

Теперь допустим, что имеет место второй случай. Тогда с учетом (47) из (65) следует

$$\exists i \in \{0, \dots, k-1\} \left(\exists p \in \bar{P}_i (p.A_R = a \wedge p.A_S = a') \right).$$

Учитывая (48) и (49) отсюда имеем $(a, a') \in P$. Вместе с (50), (61) и (62) это дает

$$(a, b_1, \dots, b_u, a', c_1, \dots, c_v) \in Q,$$

то есть (60) имеет место. Теорема доказана.

Заключение

В статье были представлены формальные описания методов декомпозиции операций пересечения, естественного соединения и тета-соединения по неравенству на основе доменно-интервальной фрагментации колоночных индексов. Для каждого метода была доказана его корректность. Практическая ценность предложенных методов декомпозиции заключается в том, что ресурсоемкие вычисления могут производиться независимо над соответствующими фрагментами. Для операции пересечения это – вычисление P_j^i по формуле (17), для операции естественного это – вычисление P_j^i по формуле (30), для тета-соединения это – вычисление \tilde{P}_i по формуле (44). В дальнейшем мы планируем разработать методы декомпозиции для операций тета-соединения с конъюнкцией, дизъюнкцией и отрицанием, а также для операции группировки.

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).

Литература

1. Turner, V. The Digital Universe of Opportunities: Rich Data and the creasing Value of the Internet of Things. — White paper. — International Data Corporation. — 2014. / V. Turner, J.F. Gantz, D. Reinsel, et al. URL: <http://idcdocserv.com/1678> (дата обращения: 29.01.2015)
2. Соколинский, Л.Б. Параллельные машины баз данных / Л.Б. Соколинский // Природа. — 2001. — № 8. — С. 10–17.
3. Соколинский, Л.Б. Параллельные системы баз данных / Л.Б. Соколинский — Москва: Издательство Московского государственного университета, 2013. — 184 с.
4. 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.
5. 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.
6. Костенецкий, П.С. Моделирование иерархических многопроцессорных систем баз данных / П.С. Костенецкий, Л.Б. Соколинский // Программирование. — 2013. — Т. 39, № 1. — С. 3–22.
7. Plattner, H. In-Memory Data Management: An Inflection Point for Enterprise Applications / H. Plattner, A. Zeier — Springer, 2011. — 254 p.
8. Abadi, D.J. Column-Stores vs. Row-Stores: How Different Are They Really? / D.J. Abadi, S.R. Madden, N. Hachem // Proceedings of the 2008 ACM SIGMOD international conference on Management of data, June 9–12, 2008, Vancouver, BC, Canada. — ACM, 2008. — P. 967–980.
9. 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.
10. 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.
 11. 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.
 12. 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.
 13. Иванова, Е.В. Использование распределенных колоночных индексов для выполнения запросов к сверхбольшим базам данных / Е.В. Иванова, Л.Б. Соколинский // Параллельные вычислительные технологии (ПАВТ'2014). Труды международной научной конференции. — Челябинск: Издательский центр ЮУрГУ, 2014. — С. 270–275.
 14. Иванова, Е.В. Использование распределенных колоночных хеш-индексов для обработки запросов к сверхбольшим базам данных / Е.В. Иванова // Научный сервис в сети Интернет: многообразие суперкомпьютерных миров. Труды Международной суперкомпьютерной конференции. — М.: Изд-во МГУ, 2014. — С. 102–104.
 15. Гарсиа-Молина, Г. Системы баз данных. Полный курс. / Г. Гарсиа-Молина, Дж. Ульман, Дж. Уидом — М.: Издательский дом «Вильямс». — 2004. — 1088 с.

Иванова Елена Владимировна, программист отдела поддержки и обучения пользователей, Лаборатория суперкомпьютерного моделирования, Южно-Уральский государственный университет (Челябинск, Российская Федерация), ivanovaev@susu.ac.ru.

Соколинский Леонид Борисович, проректор по информатизации, Южно-Уральский государственный университет (Челябинск, Российская Федерация), Leonid.Sokolinsky@susu.ru.

Поступила в редакцию 4 февраля 2015 г.

DECOMPOSITION OF INTERSECTION AND JOIN OPERATIONS BASED ON THE DOMAIN-INTERVAL FRAGMENTED COLUMN INDICES

E. V. Ivanova, South Ural State University (Chelyabinsk, Russian Federation),

L. B. Sokolinsky, South Ural State University (Chelyabinsk, Russian Federation)

The paper presents decomposition of relational operations based on distributed column indices and domain-interval fragmentation. This decomposition admits parallel executing the resource-intensive relational operations without data transfers. All column index fragments are stored in main memory in compressed form to conserve space. During the parallel execution of relational operations, compressed index fragments are loaded on different processor cores. These cores uncompress fragments, perform relational operations and compress fragments of partial result, which is a set of keys. Partial results are merged in the resulting set of keys. DBMS use the resulting set of keys for building the resulting table. Described approach allows efficient parallel query processing for very large databases on modern computing cluster systems with many-core accelerators.

Keywords: very large databases, parallel query processing, column indices, domain-interval fragmentation, decomposition of relational operations.

References

1. Turner V., Gantz J.F, Reinsel D., et al. The Digital Universe of Opportunities: Rich Data and the creasing Value of the Internet of Things. White paper. International Data Corporation. 2014. URL: <http://idcdocserv.com/1678> (accessed: 29.01.2015)
2. Sokolinsky L.B. Parallelnye mashiny baz dannyh [Parallel Database Machines]. Priroda [Nature]. 2001. No 8. P. 10–17.
3. Sokolinsky L.B. Parallelnye systemy baz dannyh [Parallel Database Systems]. Moscow, Publishing of the Moscow State University, 2013. 184 p.
4. Sokolinsky L.B. Design and Evaluation of Database Multiprocessor Architecture with High Data Availability // Proceedings of the 12th International workshop on database and expert systems applications. IEEE Computer Society, 2001. P. 115–120.
5. Pan C.S., Zymbler M.L. Taming Elephants, or How to Embed Parallelism into PostgreSQL // Lecture Notes in Computer Science. 2013. Vol. 8055, Part 1. P. 153–164.
6. Kostenetsky P.S., Sokolinsky L.B. Modelirovanie ierarhicheskikh mnogoprocessornyh sistem baz dannyh [Simulation of Hierarchical Multiprocessor Database Systems]. Programirovanie [Programming]. 2013. Vol. 39, No 1. P. 3–22.
7. Plattner H., Zeier A. In-Memory Data Management: An Inflection Point for Enterprise Applications Springer, 2011. 254 p.
8. Abadi D.J., Madden S.R, Hachem N. Column-Stores vs. Row-Stores: How Different Are They Really? // Proceedings of the 2008 ACM SIGMOD international conference on Management of data, June 9–12, 2008, Vancouver, BC, Canada. ACM, 2008. P. 967–980.

9. Fang J., Varbanescu A.L., Sips H. Sesame: A User-Transparent Optimizing Framework for Many-Core Processors // 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.
10. Breß S., Rauhe H, et al. Efficient Co-Processor Utilization in Database Query Processing // Information Systems. 2013. Vol. 38, No. 8. P. 1084–1096.
11. Scherger M. Design of an In-Memory Database Engine Using Intel Xeon Phi Coprocessors // 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.
12. Besedin K.Y., Kostenetsky P.S. Modelirovanie obrabotki zaprosov na gibridnyh vychislitel'nyh sistemah s mnogojadernymi soproprocessorami i graficheskimi uskoritel'jami [Simulation of Query Processing on Hybrid Computing Systems with Multi-core Coprocessors and Graphic Accelerators]. Programmnye sistemy: teoriya i prilozheniya [Software Systems: Theory and Applications]. 2014. Vol. 5, No 1-1 (19). P. 91–110.
13. Ivanova E.V., Sokolinsky L.B. Ispol'zovanie raspredelennykh kolonochnykh indeksov dlya vypolneniya zaprosov k sverkhbol'shim bazam dannykh [Using Distributed Column Indices for Query Execution for Very Large Databases]. Parallel'nye vychislitel'nye tekhnologii (PaVT'2014). Trudy mezhdunarodnoy nauchnoy konferentsii [Proceedings of the International Conference Parallel Computational Technologies (PCT'2014)]. Chelyabinsk: SUSU publishing center. 2014. P. 270–275.
14. Ivanova E.V. Ispol'zovanie raspredelennykh kolonochnykh hesh-indeksov dlja obrabotki zaprosov k sverhbol'shim bazam dannykh [Using Distributed Column Hash Indices for the Query for Very Large Databases] // Nauchnyj servis v seti Internet: mnogoobrazie superkomp'yuternykh mirov. Trudy Mezhdunarodnoy nauchnoy konferentsii [Proceedings of the International Scientific Conference Scientific Service on the Internet: the Variety of Supercomputing Worlds]. Moscow. Bulletin of publishing house of the Moscow university. 2014. P. 102–104.
15. Garcia-Molina H., Ullman J., Widom J. Database Systems: The Complete Book. Prentice Hall Press, 2008. 1248 p.

Received February 4, 2015.