Использование распределенных колоночных индексов для выполнения запросов к сверхбольшим базам данных*

Е.В. Иванова, Л.Б. Соколинский

ФГБОУ ВПО «ЮУрГУ» (НИУ)

В связи с появлением задач, требующих обработки сверхбольших баз данных, необходимы новые эффективные методы параллельной обработки запросов на современных кластерных вычислительных системах. В работе рассматриваются индексные структуры специального вида, названные распределенными колоночными индексами. Данные структуры позволяют использовать фрагментный параллелизм для ускорения обработки запросов к сверхбольшим базам данных.

1. Введение

В настоящее время научная и практическая деятельность человека выдвигает все новые масштабные задачи, требующие обработки сверхбольших баз данных. Согласно прогнозам аналитической компании IDC к 2020 г. количество данных в мире достигнет 40 Зеттабайт [1]. Современные технологии баз данных не могут обеспечить обработку столь крупных объемов данных. По оценке IDC из всего объема потенциально полезных данных в 2012 г. всего лишь 3% данных были проиндексированы и только 0.5% были подвергнуты анализу.

Фактически единственным эффективным решением проблемы хранения и обработки сверхбольших баз данных является использование параллельных систем баз данных, обеспечивающих параллельную обработку запросов на многопроцессорных вычислительных системах [2].

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 [3]. Возможным решением этой проблемы может быть использование механизмов колоночного представления данных, позволяющих получить на порядок лучшую производительность при обработке аналитических запросов [4]. 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 [4].

В соответствие с этим актуальной является задача разработки новых эффективных методов параллельной обработки и анализа сверхбольших объемов структурированных данных на многопроцессорных вычислительных системах, с использованием колоночного представления и сжатия данных. Для решения этой задачи нами предлагаются индексные структуры специального вида, которые называются *распределенными колоночными индексами*.

* Исследование выполнено при финансовой поддержке РФФИ в рамках научного проекта № 12-07-00443 а.

2. Колоночный индекс

Пусть имеется отношение R, состоящее из n кортежей. Пусть на множестве кортежей $R=\{r_0,r_1,\ldots,r_{n-1}\}$ задано отношение линейного порядка:

$$r_0 < r_1 < \ldots < r_{n-1} . \tag{1}$$

Адресной функцией будем называть целочисленную функцию

$$@ : R \to \{0,\ldots,n-1\} ,$$

определяющую для кортежа $r \in R$ его порядковый номер $@(r)$ в упорядоченной последовательности (1). Для функции $@$ мы будем также использовать операторную форму ее применения к кортежу $r : @r$. Значения адресной функции мы будем называть адресами кортежей или просто *адресами*.

Функцией разыменования для отношения R будем называть функцию

$$* : \{0,\ldots,n-1\} \to R ;$$

являющуюся обратной по отношению к адресной функции $@$:

$$\forall r \in R\big(*(@(r)) = r\big).$$

Для функции $*$ мы будем также использовать операторную форму ее применения к адресу $a : *a$.

Пусть задано отношение $R(B,\ldots)$, $T(R)=n$. Обозначим через $\mathfrak{D}_B$ домен атрибута B. $\mathfrak{D}_B$ – множество без повторяющихся элементов. Пусть на множестве $\mathfrak{D}_B$ задано отношение линейного порядка. *Колоночным индексом* $I_{R.B}$ атрибута B отношения R будем называть упорядоченное отношение $I_{R.B}(A,B) \in (\mathbb{Z}_{\geq 0},\mathfrak{D}_B)$, удовлетворяющее следующим требованиям:

$$T(I_{R.B}) = n \text{ и } \pi_A(I_{R.B}) = \{0,\ldots,n-1\} ; \tag{2}$$

$$\forall x_1,x_2 \in I_{R.B}\left(x_1 \leq x_2 \Leftrightarrow x_1.B \leq x_2.B\right); \tag{3}$$

$$\forall r \in R\big(\forall x \in I_{R.B}\left(@r = x.A \Rightarrow r.B = x.B\right)\big). \tag{4}$$

Условие (2) означает, что количество элементов (кортежей) индекса совпадает с количеством кортежей индексируемого отношения и атрибут *A* представляет без повторений множество всех возможных адресов кортежей отношения R. Условие (3) означает, что элементы индекса упорядочены в порядке возрастания значений атрибута B. Условие (4) означает, что атрибут A элемента индекса содержит адрес кортежа отношения R, имеющего такое же значение атрибута B, как и у данного элемента колоночного индекса.

С содержательной точки зрения колоночный индекс $I_{R.B}$ представляет собой таблицу из двух колонок с именами A и B (см. рис. 1). Количество строк в колоночном индексе совпадает с количеством строк в индексируемой таблице. Колонка B индекса $I_{R.B}$ включает в себя все значения колонки B таблицы R (с учетом повторяющихся значений), упорядоченных в порядке возрастания. Каждая строка x индекса $I_{R.B}$ содержит в колонке A порядковый номер (адрес) строки r в таблице R, имеющей такое же значение в колонке B, что и строка x.

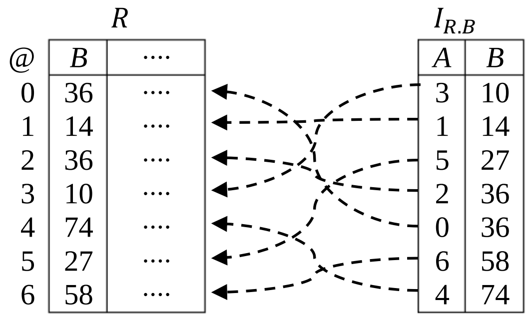

Рис. 1. Колоночный индекс

Утверждение 1. Пусть задана схема отношения $R(B,\ldots)$. Пусть для отношения R задан колоночный индекс $I_{R.B}$. Тогда

$$\pi_B\left(I_{R.B}\right)=\pi_B\left(R\right). \tag{5}$$

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 $T\left(\sigma_{B=b}\left(R\right)\right)=k$. Без ограничения общности мы можем считать, что $\forall r \in R(@r < k \Leftrightarrow r.B = b)$. Тогда из (2) и (4) следует, что $\forall x \in I_{R.B}\left(x.A < k \Leftrightarrow x.B = b\right)$. Откуда получаем $T\left(\sigma_{B=b}\left(I_{R.B}\right)\right)=k$. Таким образом (5) имеет место.

3. Доменно-интервальная фрагментация

Пусть на множестве значений домена $\mathfrak{D}_B$ задано отношение линейного порядка. Разобьем множество $\mathfrak{D}_B$ на $k>0$ непересекающихся интервалов:

$$V_0=\left[v_0;v_1\right],V_1=\left(v_1;v_2\right],\ldots,V_{k-1}=\left(v_{k-1};v_k\right];$$

$$v_0 < v_1 < \ldots < v_k;$$

$$\mathfrak{D}_B=\bigcup_{i=0}^{k-1}V_i.$$

Отметим, что в случае $\mathfrak{D}_B=\mathbb{R}$ будем иметь $v_0=-\infty$ и $v_k=+\infty$. Зададим *доменную функцию фрагментации* $\varphi_{\mathfrak{D}_B}:\mathfrak{D}_B \to \{0,\ldots,k-1\}$ для домена $\mathfrak{D}_B$ следующим образом:

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

Другими словами, доменная функция фрагментации сопоставляет значению b – номер интервала, которому это значение принадлежит. Данное определение корректно, так как интервалы $V_0,\ldots,V_{k-1}$ попарно не пересекаются и вместе составляют все множество $\mathfrak{D}_B$.

Пусть задан колоночный индекс $I_{R.B}$ для отношения $R(B,\ldots)$ с атрибутом B над доменом $\mathfrak{D}_B$. Определим для индекса $I_{R.B}$ *функцию фрагментации*

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

следующим образом

$$\forall x \in I_{R.B}\left(\varphi_{I_{R.B}}(x) = \varphi_{\mathfrak{D}_B}(x.B)\right). \tag{7}$$

Другими словами, функция фрагментации $\varphi_{I_{R.B}}$ сопоставляет каждому кортежу x из $I_{R.B}$ номер доменного интервала, которому принадлежит значение $x.B$.

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

$$I^i_{R.B} = \left\{x \mid x \in I_{R.B};\ \varphi_{I_{R.B}}(x) = i\right\}. \tag{8}$$

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:

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

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

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

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

Доказательство. Сначала докажем, что

$$\forall i \in \{0,\ldots,\mathrm{k}-1\}\left(\forall x \in I_{R.B}\left(x \in I^i_{R.B} \Rightarrow x.B \in V_i\right)\right). \tag{9}$$

Пусть $x \in I^i_{R.B}$. Тогда из (8) следует $\varphi_{I_{R.B}}(x) = i$. С учетом (7) получаем $\varphi_{\mathfrak{D}_B}(x.B) = i$. Отсюда и из (6) следует $x.B \in V_i$, то есть (9) имеет место. Теперь докажем, что

$$\forall i \in \{0,\ldots,\mathrm{k}-1\}\left(\forall x \in I_{R.B}\left(x.B \in V_i \Rightarrow x \in I^i_{R.B}\right)\right). \tag{10}$$

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 (7) получаем $\varphi_{\mathfrak{D}_B}(x.B) = \varphi_{I_{R.B}}(x) = i$. Отсюда и из (8) следует, что $x \in I^i_{R.B}$, то есть (10) имеет место. Утверждение доказано.

4. Пример выполнения реляционной операции с использованием распределенных колоночных индексов

Рассмотрим пример использования распределенных колоночных индексов для операции пересечения двух отношений.

Пусть заданы два отношения $R(B_1,\ldots,B_u)$ и $S(B_1,\ldots,B_u)$, имеющие одинаковый набор атрибутов. Мы предполагаем, что R и S не содержат дубликатов. Пусть имеется два набора колоночных индексов по каждому атрибуту:

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

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

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

$$I_{R.B_j} = \bigcup_{i=0}^{k-1} I_{R.B_j}^i \ ; \tag{11}$$

$$I_{S.B_j} = \bigcup_{i=0}^{k-1} I_{S.B_j}^i \ . \tag{12}$$

Положим

$$P_j^i = \pi_{I_{R.B_j}^i.A \to A_R,\ I_{S.B_j}^i.A \to A_S} \left(I_{R.B_j}^i \underset{R.B_j = S.B_j}{\bowtie} I_{S.B_j}^i \right) \tag{13}$$

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

$$P_j = \bigcup_{i=0}^{k-1} P_j^i \ . \tag{14}$$

Положим

$$P = \bigcap_{j=1}^{u} P_j \ . \tag{15}$$

Определим

$$Q = \left\{ r \mid r \in R \wedge @r \in \pi_{A_R}\left(P\right) \right\}. \tag{16}$$

Утверждение 3. $Q = R \cap S$.

Доказательство. Сначала докажем, что

$$Q \subset R \cap S .$$

Пусть

$$(b_1,\dots,b_u) = q \in Q .$$

Из (16) следует, что

$$(b_1,\dots,b_u) = r \in R . \tag{17}$$

Обозначим $a_R = @r$. Согласно (16) $\exists a_S \in \mathbb{Z}_{\geq 0}\left((a_R, a_S) \in P\right)$. Отсюда и из (15) следует, что $\forall j \in \{1,\dots,u\}\left((a_R, a_S) \in P_j\right)$. С учетом (14) получаем $\forall j \in \{1,\dots,u\}\left(\exists i \in \{0,\dots,k-1\}\left((a_R, a_S) \in P_j^i\right)\right)$. Отсюда и из (11)-(13) следует, что $\forall j \in \{1,\dots,u\}\left(\exists x \in I_{S.B_j}\left(x = (a_S, b_j)\right)\right)$. По определению колоночного индекса отсюда получаем, что $\exists s \in S\left(s = (b_1,\dots,b_u)\right)$. Вместе с (17) это дает

$$Q = R \cap S .$$

Теперь докажем, что $R \cap S \subset Q$. Пусть $(b_1,\dots,b_u) = r \in R$ и $(b_1,\dots,b_u) = s \in S$. Положим $a_R = @r$, $a_S = @s$. По определению колоночного индекса тогда имеем

$$\forall j \in \{1,\dots,u\}\left(\exists x_R \in I_{R.B_j}\left(x = (a_R, b_j)\right)\right);$$

$$\forall j \in \{1,\dots,u\}\left(\exists x_S \in I_{S.B_j}\left(x = (a_S, b_j)\right)\right).$$

Отсюда с учетом (11)-(16) получаем

$$\exists r' \in Q\left(@r' = a_R\right).$$

Поскольку $a_R = @r$, отсюда следует, что $r' = (b_1,\dots,b_u)$, а значит $R \cap S \subset Q$.

5. Заключение

В статье приведено формальное описание колоночных индексов и описана их структура. Введено понятие доменно-интервальной фрагментации колоночных индексов, являющейся основой применения фрагментного параллелизма для ускорения выполнения запросов с помощью распределенных колоночных индексов. Рассмотрен пример использования распределенных колоночных индексов для выполнения операции пересечения двух отношений. Колоночные индексы в сочетании с доменно-интервальной фрагментацией позволяют существенно ускорить обработку запросов к сверхбольшим базам данных. В качестве направления дальнейших исследований авторы предполагают разработать эффективные методы использования распределенных колоночных индексов в кластерных вычислительных системах, узлы которых оснащены многоядерными ускорителями.

Литература

1. *Gantz J., Reinsel D.* IDC. The Digital Universe in 2020: Big Data, Bigger Digital Shadows, and Biggest Growth in the Far East. Report, 2012. URL: http://www.emc.com/collateral/analyst-reports/idc-the-digital-universe-in-2020.pdf (дата обращения: 05.03.2014)
2. Соколинский Л.Б. Параллельные системы баз данных. М.: Издательство Московского университета, 2013. 179 с.
3. Plattner H., Zeier A. In-Memory Data Management: An Inflection Point for Enterprise Applications. Springer, 2011. 254 p.
4. 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.