УДК 004.891 к.т.н., доц. **Джулій В.М.** (ХмНУ)
Копачовець О.М. (ХмНУ)
Лєнков О.С. (ПФ «Арій»)

ЗАСТОСУВАННЯ ІНДЕКСІВ В РЕЛЯЦІЙНИХ ОПЕРАЦІЯХ НА ОСНОВІ РОЗПОДІЛУ ДОМЕННИХ ІНТЕРВАЛІВ

Стаття присвячена питанням декомпозиції реляційних операцій шляхом використання розподілених колоночних індексів з доменно-інтервальної фрагментацією. Така декомпозиція дозволяє організувати паралельне виконання ресурсоємких реляційних операций без обмінів даними між процесорними ядрами. Всі фрагменти колоночного індексу зберігаються в оперативній пам'яті в стислому вигляді. При паралельному виконанні реляційної операції упаковані фрагменти індексів вхідних відносин завантажуються на різні процесорні ядра, де відбуваються їх розпакування, виконання реляційної операції над фрагментами та упаковка часткового результату, що представляє собою набори ключів. Потім часткові результати об'єднуються в результуючий набір ключів, з використанням якого СУБД збирає результуюче відношення. Зазначений підхід дозволяє організувати ефективне паралельне виконання запитів до надвеликих баз даних на сучасних кластерних обчислювальних системах, оснащених багатоядерними прискорювачами.

Ключові слова: надвеликі бази даних, паралельна обробка запитів, колоночні індекси, домено-інтервальна фрагментація, декомпозиція реляційних операцій.

Вступ. В даний час наукова і практична діяльність людини висуває все нові масштабні завдання, що вимагають обробки надвеликих об'ємів даних. Згідно з прогнозами аналітичної компанії IDC до 2020р. кількість даних у світі досягне 40 зеттабайт. У зв'язку з появою задач, що вимагають обробки надвеликих баз даних, необхідні нові ефективні методи паралельної обробки та аналізу таких обсягів даних на багатопроцесорних обчислювальних системах. Фактично єдиним ефективним вирішенням проблеми зберігання і обробки надвеликих об'ємів даних є використання паралельних систем баз даних, що забезпечують паралельну обробку запитів на багатопроцесорних обчислювальних системах.

Фактично єдиним ефективним вирішенням проблеми зберігання і обробки надвеликих баз даних є використання паралельних систем баз даних, що забезпечують розподілену обробку запитів на багатопроцесорних обчислювальних системах з розподіленою пам'яттю [2].

Постановка задачі. Традиційним підходом до організації зберігання баз даних є строкове подання даних. Однак при виконанні типових аналітичних запитів до таблиць потрібно зчитувати тільки невелику частину полів у рядках цих таблиць, через це строкове представлення в цьому випадку виявляється неефективним. Причиною цього є зчитування з диска «зайвих» полів на додаток до тих полів, які необхідні в даному запиті [3]. Можливим вирішенням цієї проблеми може бути використання механізмів колоночного представлення даних, що дозволяють отримати значно кращу продуктивність при обробці аналітичних запитів [5]. Колоночне представлення даних полягає в тому, що дані зберігаються не по рядках, а по колонках. Це означає, що з точки зору SQL-клієнта дані представлені у вигляді таблиць, але фізично ці таблиці є сукупністю колонок, кожна з яких являє собою таблицю з одного поля. Додатковою перевагою колоночного представлення є можливість використання ефективних алгоритмів стиснення даних, оскільки в одній колонці таблиці містяться дані одного типу. Стиснення може призвести до значного підвищення продуктивності, оскільки менше часу займають операції вводу-виводу.

У відповідність з цим актуальною є задача розробки нових ефективних методів паралельної обробки баз даних в оперативній пам'яті на сучасних багатопроцесорних обчислювальних системах з багатоядерними прискорювачами, з використанням колоночного представлення і стиснення даних. Для вирішення цього завдання можна використати індексні структури спеціального виду, які називаються розподіленими колоночними індексами. Розподілені колоночні індекси дозволяють провести декомпозицію реляційних операцій, яка забезпечує їх ефективне паралельне виконання на кластерних обчислювальних системах з багатоядерними прискорювачами. У даній роботі розглянуті питання декомпозиції наступних реляційних операцій: перетину, природного з'єднання і тета-з'єднання.

Виклад основного матеріалу досліджень. Спочатку формалізуємо поняття колоночного індексу та домено-інтервальної фрагментації.

Під $R(A^*, B_1, ..., B_n)$ будемо розуміти відношення R з первинним ключем A та атрибутами $B_1, ..., B_n$, що являють собою множину кортежів довжини $n+1$ та виду $a, b_1, ..., b_n$, де $a \in Z_{\geq 0}$ та $\forall j \in \{1, ..., n\}(b_{j\in} D_{B_j})$. Тут D_{B_j} - домен атрибута B_J. Через $r.B_J$ будем позначати значення атрибуту B_J, Через $r.A$ - значення первинного ключа в кортежі $r = (r.A, r.B_1, ..., r.B_n)$. Первинний ключ відношення R має властивість $\forall r', r'' \in R(r' \neq r'' \Leftrightarrow r'.A \neq r''.A)$. Під адресом кортежу r ми будемо розуміти значення первинного ключа цього кортежу. Для отримання кортежу відношення R за його адресу будемо використовувати функцію розіменування $\&_r : \forall r \in R(\&_r\, (r.A) = r)$.

Нехай задано відношення $R(A^*, B, ...), T(R) = n$. Нехай на множині D_B задано відношення лінійного порядку. Колоночним індексом $I_{R.B}$ атрибута B відношення R називається впорядковане відношення $I_{R.B}(A^*, B)$, що задовольняє наступні вимоги:

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

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

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

Умова (1) значить, що множина значень первинних ключів (адрес) індексу та індексуємих відношень співпадають. Умова (2) значить, що елементи індексу впорядочені в порядку зростання значень атрибута B. Умова (3) значить, що атрибут A елемента індексу містить адресу кортежу відношення R, що має таке саме значення атрибута B, як і в даного елемента колоночного індексу.

Припустимо задано відношення $R(A^*, B, ...)$. Нехай для відношення R заданий колоночний індекс $I_{R.B}$. Тоді:

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

Іншими словами, колоночний індекс $I_{R.B}$ представляє всю множину значень атрибута B відношення R з урахуванням повторних значень.

Щоб довести (4) візьмемо довільне $b = \mathrm{D}_B$. Нехай $T(\sigma_{B-b}(I_{R.B})) = k$. Без обеження спільності ми можемо стверджувати, що $\forall r \in R(r.A < k \Leftrightarrow r.B = b)$. Тоді з (1) та (3) випливає, що $\forall x \in I_{R.B}(x.A < k \Leftrightarrow x.B = b)$. Звідки отримуємо $T(\sigma_{B-b}(I_{R.B})) = k$.

Таким чином колоночний індекс дозволяє використовувати один стовпчиковий індекс для індексування декількох атрибутів одного відношення.

Нехай на множині значень домену D_B задано відношення лінійного порядку. Нехай також задано розбиття множини D_B на $\mathrm{k} > 0$ непересічних інтервалів:

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

Функція $\varphi_{D_B} : D_B \to \{0,...,k-1\}$ називається інтервальною функцією фрагментації для домену D_B, якщо вона задовольняє наступній умові:

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

Нехай задано колоночний індекс $I_{R.B}$ для відношення $R(A^*, B, ...)$ з атрибутом B над доменом D_B та інтервальна функція фрагментації φ_{D_B}. Функція:

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

Називається домено-інтервальною функцією фрагментації для індексу $I_{R.B}$, якщо вона задовольняє наступну умову:

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

Визначимо *i*-тий фрагмент $(i = 0,...,k-1)$ індексу $I_{R.B}$ наступним чином:

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

Це означає, що в *i*-тий фрагмент попадають кортежі, у яких значення атрибута B належить *i*-тому доменому інтервалу. Будемо називати фрагментацію, побудовану таким чином, домено-інтервальною. Кількість фрагментів k будемо називати ступінь фрагментації.

Домено-інтервальна фрагментація володіє наступними фундаментальними властивостями, які витікають безпосередньо з її визначення:

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

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

Розглянемо декомпозиція операції перетину.

Нехай задані два відношення $R(A^*, B_1, ..., B_n)$ та $S(A^*, B_1, ..., B_n)$, які мають однаковий набір атрибутів. Ми припускаєм, що $\delta\left(\pi_{B_1,...,B_n}(R)\right) = \pi_{B_1,...,B_n}(R)$ та $\delta\left(\pi_{B_1,...,B_n}(S)\right) = \pi_{B_1,...,B_b}(S)$, тобто $\pi_{B_1,...,B_n}(R)$ і $\pi_{B_1,...,B_n}(S)$ не містять дублікатів. Нехай маємо два набора колоночних індексів по атрибутам $B_1,...,B_n$: $I_{R.B_1},...,I_{R.B_n}$ та $I_{S.B_1},...,I_{S.B_n}$.

Нехай для всіх цих індексів задана доменно-інтервальна фрагментація степені k:

$$I_{R.B_j} = \bigcup_{i=0}^{k-1} I^i_{R.B_j}; \tag{12}$$

$$I_{S.B_j} = \bigcup_{i=0}^{k-1} I^i_{S.B_j}. \tag{13}$$

Припустим

$$P^i_j = \pi_{I^i_{R.B_j}.A \to A_R,\, I^i_{S.B_j}.A \to A_S}\left(I^i_{R.B_j} \underset{\left(I^i_{R.B_j}.B_j = I^i_{S.B_j}.B_j\right)}{\rhd\lhd} I^i_{S.B_j}\right) \tag{14}$$

Для всіх $i = 0,...,k-1$ і $j = 1,...,n$. Визначимо

$$P_j = \bigcup_{i=0}^{k-1} P^i_j. \tag{15}$$

Припустим

$$P = \bigcap_{j=1}^{n} P_j. \tag{16}$$

Визначаємо

$$Q = \{r | r \in R \wedge r.A \in p_{A_R}(P)\}. \tag{17}$$

Доведемо наступне припущення: $\pi_{B_1,...,B_n}(Q) = \pi_{B_1,...,B_n}(R) \cap \pi_{B_1,...,B_n}(S)$.

Доведення. Спочатку доведемо, що

$$\pi_{B_1,\dots,B_n}(Q) \subset \pi_{B_1,\dots,B_n}(R) \cap \pi_{B_1,\dots,B_n}(S). \tag{18}$$

Нехай

$$(a, b_1, \dots, b_n) \in Q. \tag{19}$$

Із (27) слідує, що

$$(a, b_1, \dots, b_n) = r \in R \tag{20}$$

та

$$a = r.A \in \pi_{A_R}(P). \tag{21}$$

Звідси виходить, що $\exists p \in P(p.A_R = a \wedge p.A_S = a')$. З врахуванням (16) одержуємо, що $\forall j \in \{1,\dots,n\}(\exists p \in P_j(p.A_R = a \wedge p.A_S = a'))$. З врахуванням (17) звідси одержуємо $\forall j \in \{1,\dots,n\}(\exists i \in \{0,\dots,k-1\}(\exists p \in P_j^i(p.A_R = a \wedge p.A_S = a')))$. Звідси та з (12) – (14) слідує, що $\forall j \in \{1,\dots,n\}(\exists x \in I_{R.B_j}(\exists y \in I_{S.B_j}(x.A = a \wedge x.B_j = y.B_j \wedge y.A = a')))$.

За визначенням колоночного індексу звідси отримуємо

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

Оскільки A являється первинним ключем в R і S, з врахуванням (20) звідси слідує $(a', b_1, \dots, b_n) \in S$, тобто $(b_1, \dots, b_n) \in \pi_{B_1,\dots,B_n}(R) \cap \pi_{B_1,\dots,B_n}(S)$. Таким чином (18) має місце.

Тепер доведемо, що

$$\pi_{B_1,\dots,B_n}(R) \cap \pi_{B_1,\dots,B_n}(S) \subset \pi_{B_1,\dots,B_n}(Q). \tag{22}$$

Нехай

$$(a, b_1, \dots, b_n) = r \in R \tag{23}$$

та

$$(a', b_1, \dots, b_n) = s \in S. \tag{24}$$

Тоді за визначенням колоночного індексу з врахуванням (4) маємо

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

На основі (12) звідси отримуємо

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

З врахуванням (17) звідси випливає

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

Використовуючи (15) одержуємо

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

Враховуючи (16) звідси маємо $(a, a') \in P$. Разом з (17) і (23) це дає

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

Тобто (22) має місце та декомпозиція операції перетину можлива.

Тепер розглянемо декомпозицію операції природнього з'єднання.

Нехай задані два відношення $R(A^*, B_1, \dots, B_n, C_1, \dots, C_v)$ та $S(A^*, B_1, \dots, B_n, D_1, \dots, D_w)$.

Нехай маються два набори колоночних індексів по атрибутам $B_1, \dots, B_n$: $I_{R..B_1}, \dots, I_{R.B_n}$ та $I_{S.B_1}, \dots, I_{S.B_n}$.

Нехай для всіх цих індексів задана доменно-інтервальна фрагментація степені k:

$$I_{R.B_j} = \bigcup_{i=0}^{k-1} I_{R.B_j}^i; \tag{25}$$

$$I_{S.B_j} = \bigcup_{i=0}^{k-1} I_{S.B_j}^i. \tag{26}$$

Припустимо

$$P_j^i = \pi_{I^i_{R.B_j}.A\to A_R,\, I^i_{S.B_j}.A\to A_S}\left(I^i_{R.B_j} \underset{I^i_{R.B_j}.B_j = I^i_{S.B_j}.B_j}{\rhd\lhd} I^i_{S.Bj} \right) \qquad (27)$$

Для всіх $i = 0,...,k-1$ та $j = 1,...,n$. Визначимо

$$P_j = \bigcup_{i=0}^{k-1} P_j^i \ . \qquad (28)$$

Припустим

$$P = \bigcap_{j=1}^{u} P_j \ . \qquad (29)$$

Визначимо

$$Q = \{ r \circ (s.D_1,...,s.D_w) \mid r \in R \wedge s \in S \wedge (r.A, s.A) \in P \}. \qquad (30)$$

Доведемо наступне припущення: $\pi_{*\backslash A}(Q) = \pi_{*\backslash A}(R) \rhd\lhd \pi_{*\backslash A}(S)$.

Спочатку доведемо, що

$$\pi_{*\backslash A}(Q) \subset \pi_{*\backslash A}(R) \rhd\lhd \pi_{*\backslash A}(S). \qquad (31)$$

Нехай

$$(a, b_1,...,b_n, c_1,...,c_v, d_1,...,d_w) \in Q. \qquad (32)$$

Із (33) слідує, що існують кортежі r та s такі, що

$$(a, b_1,...,b_n, c_1,...,c_v) = r \in R\,, \qquad (33)$$

$$(a' b_1',...,b_n', d_1,...,d_w) = s \in S \qquad (34)$$

та

$$(r.A, s.A) \in P\,. \qquad (35)$$

Звідси випливає, що $\exists p \in P(p.A_R = a \wedge p.A_s = a')$. З врахуванням (29) отримуємо, що $\forall j \in \{1,...,n\}(\exists p \in P_j(p.A_R = a \wedge p.A_s = a'))$. З врахуванням (28) звідси отримуємо $\forall j \in \{1,...,n\}(\exists i \in \{0,...,k-1\}(\exists p \in P_j^i(p.A_R = a \wedge p.A_s = a')))$. Звідси та з (25) – (27) випливає, що

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

По визначенню колоночного індексу звідси отримуємо

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

Оскільки A являється первинним ключем в R та S, з врахуванням (33) та (34) маємо $(a', b_1,...,b_n, d_1,...,d_w) \in S$, тобто $(b_1,...,b_n, c_1,...,c_v, d_1,...,d_w) \in \pi_{*\backslash A}(R) \rhd\lhd \pi_{*\backslash A}(S)$. Таким чином (31) має місце.

Тепер доведемо, що

$$\pi_{*\backslash A}(R) \rhd\lhd \pi_{*\backslash A}(S) \subset \pi_{*\backslash A}(Q)\ . \qquad (36)$$

Нехай

$$(a, b_1,...,b_n, c_1,...,c_v) = r \in R \qquad (37)$$

та

$$(a', b_1,...,b_n, d_1,...,d_w) = s \in S\,. \qquad (38)$$

Тоді за визначенням колоночного індексу з врахуванням (4) маємо

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

На основі (12) звідси отримуємо

$$\forall j \in \{1,...,n\}(\exists i \in \{0,...,k-1\}(\exists x \in I^i_{R.B_j}(\exists y \in I^i_{S.B_j}(x.A = a \wedge x.B_j = y.B_j \wedge y.A = a'))))$$

з врахуванням (27) звідси випливає

$$\forall j \in \{1,...,n\}(\exists i \in \{0,...,k-1\}(\exists p \in P_j^i(p.A_R = a \wedge p.A_s = a'))).$$

Застосовуючи (28) отримуємо

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

Враховуючи (29) звідси маємо $(a,a') \in P$. Разом з (30), (37) і (38) це дає

$$(a,b_1,...,b_n,c_1,...,c_v,d_1,...,d_w) \in Q,$$

тобто (36) має місце, а це значить, що декомпозиція операції природнього з'єднання можлива.

Висновки. У статті була проведена формалізація понятть колоночного індексу та домено-інтервальної фрагментації. Також представлено формальний опис методів декомпозиції операцій перетину та природного з'єднання на основі домено-інтервальної фрагментації колоночних індексів. Для кожного методу було доведено його коректність. Практична цінність запропонованих методів декомпозиції полягає в тому, що ресурсомісткі обчислення можуть проводитись незалежно над відповідними фрагментами. Для операції перетину це – обчислення P_j^i за формулою 17, а для операції природнього з'єднання це – визначення P_j^i за формулою (30). Також планується розробка методів декомпозиції для інших реляційних операцій.

ЛІТЕРАТУРА:

1. П.С. Костенецкий, Моделирование иерархических многопроцессорных систем баз данных / П.С. Костенецкий, Л.Б. Соколинский // Программирование, 2013. – Т. 39 – № 1. – С. 3-22.
2. Соколинский Л.Б. Параллельные системы баз данных. – М.: Издательство Московского университета, 2013. – 179 с.
3. David J DeWitt, The Gamma Database Machine Project. / David J DeWitt, Shahram Ghandeharizadeh, Donovan A. Schneider, Allan Bricker, H-I Hsiao, and Rick Rasmussen // IEEE Trans. on Knowledge and Data Engineering, 1990. – №2. – С.44–62.
4. Andrew Lamb, The Vertica Analytic Database: C-store 7 Years Later. / Andrew Lamb, Matt Fuller, Ramakrishna Varadarajan, Nga Tran, Ben Vandiver, Lyric Doshi, and Chuck Bear // Proc. of the VLDB Endowment, 2012 – №12 – С.1790–1801.

Рецензент: **к.т.н., доц. Муляр І.В.** доцент кафедри комп'ютерних систем та мереж, Хмельницький національний університет

к.т.н., доц. Джулий В.Н., Копачовец А.М., Ленков А.С.

ПРИМЕНЕНИЕ ИНДЕКСОВ В РЕЛЯЦИОННЫХ ОПЕРАЦИЯХ НА ОСНОВЕ РАСПРЕДЕЛЕНИЯ ДОМЕННЫХ ИНТЕРВАЛОВ

В статье представлено разложение некоторых реляционных операций, основанное на распределенных индексов столбцов и фрагментации доменных интервалов. Это разложение допускает параллельное выполнение ресурсоемких реляционных операций без передачи самих данных. Все фрагменты столбцовых индексов хранятся в основной памяти в сжатом виде для экономии места. Во время параллельного выполнения реляционных операций, сжатые фрагменты индекса загружаются на разные процессорные ядра. Происходит распаковка та выполнение определенных реляционных операций и сжимание фрагментов частичного результата, который представляет собой набор ключей. Частичные результаты объединяются в результирующий набор ключей. СУБД использует результирующий набор ключей для построения результирующей таблицы. Описанный подход позволяет обеспечить эффективную параллельную обработку запросов для очень больших баз данных на современных кластерных вычислительных системах с многоядерными ускорителями.

Ключевые слова: параллельная обработка запросов, индексы столбцов, фрагментация данных, разложение реляционных операций, сверхбольшие базы данных.

Ph.D. Julie V.M., Kopachovets A.M, Lenkov A.S.

USING AN INDEX IN A RELATIONAL OPERATION BASED ON DOMAIN INTERVALS FRAGMENTATION

The article 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 uncompressed 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.