



ся увеличения производительности при осуществлении векторных операций с данными, расположенными в памяти через равные промежутки.

ЛИТЕРАТУРА

1. *Воеводин В.В.* Математические модели и методы в параллельных процессах. М.: Наука, 1986.
2. *Воеводин В.В.* Параллельные структуры алгоритмов и программ. М., 1987.
3. *Корнеев В.В.* Параллельные вычислительные системы. М., 1999.
4. *Zabrodin A.V., Levin V.K., Korneev V.V.* The massively parallel computer system MBC-100. Lecture Notes in Computer Science. № 964 // Parallel Computing Technologies. Third International Conference, PaCT-95. St.Petersburg: Springer. 1995. P.341-355.
5. *Whaley R.C., Petitet A., Dongarra J.J.* Automated Empirical Optimization of Software and the ATLAS project. WWW.netlib.org/atlas.

УДК 519.68

© 2001 г. **А.В. Бурдаков,**
Ю.А. Григорьев, д-р техн. наук
(Московский государственный технический университет им. Н.Э. Баумана),
А.Д. Плутенко, канд. техн. наук
(Амурский государственный университет, Благовещенск)

ОЦЕНКА ВРЕМЕНИ ВЫПОЛНЕНИЯ ЗАПРОСОВ К ОБЪЕКТНО-ОРИЕНТИРОВАННЫМ БАЗАМ ДАННЫХ

Предложен математический метод оценки времени выполнения запросов к объектно-ориентированным базам данных. Рассмотрены алгоритмы выполнения запросов Forward Join и Reverse Join.

Введение

Разработка распределенных систем обработки данных (РСОД) сопряжена с оценкой различных функциональных и нефункциональных показателей предлагаемых решений. Одними из важных характеристик являются индексы производительности РСОД. Для оценки этих показателей на ранних этапах разработки системы необходимо применять специальные

методы расчета, т.к. натурные эксперименты либо невозможны, либо слишком трудоемки.

В настоящее время предложен ряд методов оценки времени выполнения запросов к объектно-ориентированным системам управления базами данных (ООСУБД), однако все они отличаются определенными недостатками, не позволяющими применять их на ранних стадиях разработки РСОД.

Так, в работе [11] предложена аналитическая модель расчета времени (стоимости) выполнения запросов к ООСУБД. Здесь рассмотрены четыре модели хранения объектов на диске и предложен метод оценки влияния кэш-буферов на выполнение запросов. Однако в этой модели не сделана оценка стоимости соединения как подобъектов, так и различных таблиц, а также не учтено влияние индексов. Кроме того, модель кэш-памяти предполагает, что ее объем сравним с объемом БД, что в общем случае неверно. Все расчеты выполняются на уровне средних величин. В работе [7] также предложена аналитическая модель оценки времени выполнения запросов к ООСУБД. Для нахождения коэффициентов аналитических выражений предлагается использовать калибрующую модель, представляющую собой определенную БД и набор запросов, а также программно-аппаратный комплекс, на котором выполняется калибрующий эксперимент. Основным недостатком данной модели является необходимость выполнения калибровки для каждой программно-аппаратной платформы, что по трудоемкости приближает ее к экспериментальному исследованию.

В работе [5] представлены и исследованы имитационные модели для нескольких вариантов архитектуры клиент-сервер (с учетом механизмов распараллеливания дискового ввода-вывода и кэширования). Однако эта модель не позволяет вычислить время обработки произвольных запросов к СУБД. Кроме того, модель является имитационной, а следовательно, позволяет моделировать лишь небольшие системы с малым числом клиентов и типов запросов. Работы [8], [9] и [12] предлагают аналитические модели оценки времени выполнения запросов к реляционным СУБД (РСУБД) по алгоритмам Hash, Sort-Merge и Nested-Loop Join, однако специфика объектно-ориентированных СУБД не рассматривается.

Отмеченные выше недостатки существующих методов анализа не позволяют применять их на ранних стадиях жизненного цикла разработки автоматизированных систем, построенных на основе ООСУБД. Это привело к необходимости создания нового математического метода, представленного в данной статье.

В статье рассматривается модель данных и алгоритмы выполнения запросов к ООСУБД, приводятся предпосылки анализа, а также разрабатывается математический метод анализа двух алгоритмов выполнения запросов к ООСУБД.

Алгоритмы выполнения запросов к ООСУБД

Специфика оценки времени выполнения запросов к объектно-ориентированным и объектно-реляционным СУБД (ОО/ОРСУБД) [1] вызвана рядом их отличий от реляционных СУБД: в ОО/ОРСУБД используются другие модели данных и алгоритмы выполнения запросов.

Модель данных (объектов) ОО/ОРСУБД [1] основывается на стандартах ODMG 2.0 и SQL-1999 [1]. Объектно-ориентированная модель отличается от реляционной наличием явно установленных ассоциативных связей, иерархических связей наследования, композиционной структурой объектов и наличием уникальных идентификаторов объектов (Object Identifier – OID) [1]. Основное отличие ООСУБД при выполнении запросов состоит в явном наличии и использовании ассоциативных связей между объектами. В реляционной модели такие связи отсутствуют, они устанавливаются динамически в процессе выполнения запросов.

В ОР/ООСУБД ассоциативные связи между объектами обычно представляются в виде вложенных наборов ссылок на объекты [3]. Эти наборы позволяют моделировать связи типа 1:1, 1:М и N:М (здесь при использовании последнего типа не требуется создавать дополнительную таблицу связи в отличие от ER-модели). В качестве примера ассоциативных связей на рис. 1 представлены экстенды классов R и S, при этом объекты R ссылаются на объекты S с помощью специального множественного атрибута REF.

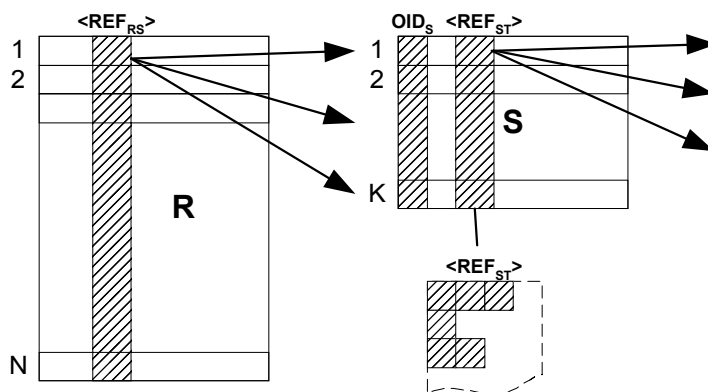


Рис. 1. Реализация ассоциативных связей в СУБД

Набор ссылок $\langle REF_{RS} \rangle$ объектов из R на объекты связанного класса S хранится непосредственно в объектах R (см. рис. 1). Каждая ссылка набора $\langle REF_{RS} \rangle$ указывает на объектный идентификатор OID объекта из S. Набор может иметь не только фиксированную, но и переменную длину.

Ссылки в ООСУБД могут быть как однонаправленными, так и двунаправленными. В случае установления двунаправленных ссылок во втором экстенде класса S добавляется дополнительный множественный атрибут, указывающий на OID объекта из R.

Одним из основных отличительных свойств декларативного языка запросов OQL [10] к ООСУБД является «выражение пути» (path expression):

$$R[F_R]S[F_S]T[F_T]...$$

где R, S, T, \dots – отношения (классы), ... (1)

F_R, F_S, F_T, \dots – предикаты селекции.

«Выражение пути» описывает ассоциативную связь между объектами и соответствует ветви ассоциативной связи между классами. В реляционном языке запросов SQL аналогами «выражения пути» являются внешние соединения (outer join) таблиц по ключевым атрибутам [4, 12].

Для выполнения запросов с «выражениями пути» в ООСУБД используются два основных метода: Forward Join (прямое соединение) и Reverse Join (обратное соединение).

Алгоритм Forward Join (FJ) применяется при наличии прямой связи между классами, например, при выполнении запроса (1) на схеме, представленной на рис. 1. При выполнении запроса по алгоритму FJ ООСУБД отбирает из экстенста класса S те объекты, OID которых указаны в атрибуте-ссылке объектов экстенста R , проверяя при этом соответствие объектов из S наложенному на них подусловию из (1). Для отобранных объектов из S СУБД повторяет поиск объектов в экстенсте T и т. д.

Алгоритм Reverse Join (RJ) применяется при наличии обратной логической связи между классами. В этом случае СУБД выполняет запрос в обратном порядке, т.е. сначала отбирает OID объектов из S (см. рис. 1), а затем выполняет поиск объектов из R , ссылающихся на отобранные объекты по их идентификаторам OID . Таким образом, соединение выполняется в обратном порядке.

Для описанных алгоритмов выполнения запросов Forward Join и Reverse Join ниже предложен математический метод оценки времени их выполнения.

Обозначения и предпосылки

Для анализа времени выполнения запросов достаточно рассмотреть частный случай соединения двух экстенстов классов R и S , представленных на рис. 1, так как соединение произвольного числа экстенстов сводится к рассматриваемым бинарным операциям.

Рассмотрим два экстенста классов, далее представленных в виде отношений R и S , связанных ассоциативной связью и соединяемых с помощью алгоритмов Forward Join и Reverse Join.

Число объектов в отношении R обозначим через N , а число объектов в отношении S – через K . Каждый объект из S имеет уникальный объектный идентификатор OID , а каждый объект из R может иметь произвольное

число ссылок на объекты S по их OID . Домен значений ссылок из R на S обозначим как $D_{REF_{RS}}$, а домен значений объектных идентификаторов из S как D_{OID_S} , или в сокращенной форме D_{REF} и D_{OID} соответственно.

Так как в модели данных ОО/ОРСУБД обязательна ссылочная целостность в ассоциативных связях, то домен значений ссылок является подмножеством домена объектных идентификаторов:

$$D_{REF} \subseteq D_{OID}. \quad (2)$$

Мощность доменов обозначим следующим образом: $|D_{REF}|, |D_{OID}|$.

Обозначим вероятность ссылки произвольного объекта из R на объект из S как η_i .

Будем считать независимыми в совокупности события ссылки из каждого объекта из R на какой-либо объект из S .

$$\forall i = 1..|D_{REF}|, \exists j = 1..|D_{OID}|: a_{REF_i} = a_{OID_j} \quad (3)$$

здесь a_{REF} и a_{OID} – значения из домена ссылок и домена объектных идентификаторов соответственно.

Оценка времени выполнения соединения Forward Join

Для задания дискретных параметров модели – таких как число строк (объектов) в отношении (классе) S , число выбранных объектов и т. п., используются невырожденные распределения вероятностей. Поэтому в предлагаемом математическом методе применены производящие функции дискретных случайных величин (ПФСВ) [2, 6], определяемые следующей суммой:

$$G(z) = p_0 + p_1z + p_2z^2 + \dots = \sum_{i=0}^{\infty} p_i z^i. \quad (4)$$

Для оценки времени выполнения соединения, выполняемого по алгоритму Forward Join, необходимо определить число объектов, выбранных из отношения S . Лемма 1 представляет оценку числа объектов второго отношения, выбираемых при соединении пары R, S по алгоритму FJ.

Лемма 1. Пусть даны два отношения R и S , связанных между собой ассоциативной связью вида $R.REF \rightarrow S.OID$ (см. рис. 1), и на них выполняется запрос вида (1) с помощью алгоритма Forward Join. Пусть также выбор j -го объекта по условию F_S не зависит от выбора объекта по ссылке REF и значения ссылок (3) независимы в совокупности. Тогда:

1) ПФСВ числа отобранных объектов во втором отношении (S) определяется следующим выражением:

$$\Omega(z) = [z^0]V_S(z) + \sum_{k=1}^K [z^k]V_S(z) \binom{K}{k}^{-1} \sum_{(j_1, \dots, j_k) \in Q_k} \prod_{j=j_1, \dots, j_k} (1 - (1 - [z^0]V_R(1 - \eta_j(1 - z))) (1 - z)) \quad (5)$$

2) математическое ожидание числа отобранных объектов имеет следующий вид:

$$M(\omega) = \sum_{k=1}^K \left([z^k]V_S(z) \binom{K}{k}^{-1} \sum_{(j_1, \dots, j_k) \in Q_k} \sum_{j=j_1, \dots, j_k} (1 - [z^0]V_R(1 - \eta_j(1 - z))) \right), \quad (6)$$

где $V_R(z)$ - ПФСВ числа выбранных объектов в отношении R по условию F_R ;

$V_S(z)$ - ПФСВ числа выбранных объектов из S по условию ;

F_R, F_S - условия селекции, наложенные на отношения R и S (см. (1));

K - число объектов в отношении S ;

η_j - вероятность ссылки объекта из R на j -й объект отношения S ;

Q_k - множество комбинаций по " k " элементов всех идентификаторов (номеров) объектов из S .

Доказательство. Из свойства свертки ПФСВ следует, что ПФСВ числа попаданий при N независимых испытаниях по схеме Бернулли с параметром p [2, 6] определяется следующей ПФСВ:

$$G(z) = (1 - p(1 - z))^N. \quad (7)$$

Так как вероятность ссылки на j -й объект отношения (класса) S равна η_j , ПФСВ числа ссылок одного произвольного объекта класса R на j -й объект класса S будет равна:

$$(1 - \eta_j(1 - z)). \quad (8)$$

Так как число выбранных объектов в отношении R по условию F_R определяется ПФСВ V_R , то общее число ссылок на j -й объект может быть представлено как сумма независимых случайных величин с ПФСВ (8). Отсюда, учитывая свойство независимых слагаемых величин, ПФСВ числа ссылок на j -й объект отношения S будет иметь следующий вид:

$$G_j(z) = V_R(1 - \eta_j(1 - z)). \quad (9)$$

Вероятность того, что число ссылок χ_j на j -й объект будет равно нулю, может быть определена следующим образом:

$$P(\chi_j = 0) = [z^0]G_j(z).$$

Таким образом, ПФСВ выбора (0 - нет выбора, 1 - есть выбор) j -го объекта по ссылке может быть представлено следующим выражением:

$$(1 - P(\chi_j \neq 0)(1 - z)) = (1 - (1 - [z^0]G_j(z))(1 - z)). \quad (10)$$

Рассмотрим теперь условие селекции F_S , наложенное на отношение S (см. (1)). При условии, что предикату F_S удовлетворяет k объектов из S , для каждого набора объектов j_1, \dots, j_k из S ПФСВ числа отобранных объектов класса S (и по ссылке из R , и по предикату F_S) будет равна:

$$\prod_{j=j_1, \dots, j_k} \left(1 - (1 - [z^0]G_j(z))(1-z)\right). \quad (11)$$

Согласно условию леммы, множество всех комбинаций по k элементов всех идентификаторов (номеров) объектов из S обозначено как:

$$Q_k, \quad |Q_k| = \binom{K}{k} \quad (12)$$

Учитывая, что все комбинации по k элементов равновероятны с вероятностью $|Q_k|^{-1}$, получим ПФСВ числа отобранных объектов при условии, что предикату F_S удовлетворяет k объектов из S :

$$\binom{K}{k}^{-1} \sum_{(j_1, \dots, j_k) \in Q_k} \prod_{j=j_1, \dots, j_k} \left(1 - (1 - [z^0]G_j(z))(1-z)\right). \quad (13)$$

Отсюда по формуле полной вероятности для производящих функций получим:

$$\Omega(z) = [z^0]V_S(z) + \sum_{k=1}^K \left([z^k]V_S(z) \binom{K}{k}^{-1} \sum_{(j_1, \dots, j_k) \in Q_k} \prod_{j=j_1, \dots, j_k} \left(1 - (1 - [z^0]G_j(z))(1-z)\right) \right) \quad (14)$$

Подставляя (9) в (14), получим выражение (5) ПФСВ числа отобранных объектов отношения S .

Математическое ожидание (6) числа отобранных объектов из S можно найти из (5), используя технику дифференцирования произведения.

Доказательство леммы 1 завершено.

Из леммы 1 может быть получено, в частности, выражение для среднего числа объектов, отобранных из S при условии равных вероятностей ссылок η_j .

Следствие 1. Пусть вероятности ссылок η_j равны между собой и определяются следующим выражением:

$$\forall j \in 1 \dots K : \eta_j = \eta = \frac{1}{|D_{REF}|} = \frac{1}{|D_{OID}|} = \frac{1}{K},$$

а домены D_{REF} и D_{OID} совпадают, тогда:

1) ПФСВ числа объектов, отобранных из S , равна:

$$\Omega(z) = \left(1 - \left(1 - \left(1 - \frac{1}{K} \right)^{N_R} \right) (1-z) \right)^{K_S},$$

2) математическое ожидание числа объектов, отобранных из S , равно:

$$M(\omega) = K_S \left(1 - \left(1 - \frac{1}{K} \right)^{N_R} \right), \quad (15)$$

где N_R и K_S - число отобранных объектов из R и S по условиям F_R и F_S соответственно;

K - число объектов в отношении S .

Время выполнения запросов к ООСУБД определим с помощью функции распределения вероятностей $F(t)$ непрерывной случайной величины (н.с.в.) ξ , используя следующее преобразование Лапласа-Стилтьеса (ПЛС):

$$LST(s) = \int_0^{\infty} e^{-st} dF(t). \quad (16)$$

Алгоритм поиска объектов из отношения S по их объектным идентификаторам (OID) зависит от типа OID. В случае физических идентификаторов доступ осуществляется непосредственно к объектам отношения S . В случае же логических идентификаторов поиск сначала осуществляется во вспомогательной hash или b-tree структуре ($S_{OID} \rightarrow$ физический _адрес), после чего с помощью полученного физического указателя осуществляется доступ непосредственно к объекту.

Теорема 1 определяет формулы (17) и (18) для вычисления ПЛС времени выполнения соединения с учетом особенностей реализации идентификаторов OID.

Теорема 1. *Предположим, что выполнены предпосылки леммы 1 и соединение выполняется по физическим и логическим указателям. Пусть время доступа к объекту из отношения S по физическому указателю является н.с.в. θ с ПЛС $\Theta(s)$, а время поиска физического указателя во вспомогательной структуре является н.с.в. ω с ПЛС $\Omega(s)$, тогда:*

1) ПЛС времени выполнения соединения определяется следующим выражением:

$$\Xi(s) = \Psi(\Theta(s)\Omega(s)); \quad (17)$$

2) математическое ожидание времени выполнения соединения имеет следующий вид:

$$M = M(\xi)(M(\theta) + M(\omega)), \quad (18)$$

где: $\Psi(z) = \prod_{j=1}^{|D_{OID}|} V_R(1 - \eta_j(1 - z))$ - ПФСВ числа ссылок из R к S ;

$M(\xi)$ - математическое ожидание числа ссылок из R к S ;

$M(\theta)$ - математическое ожидание времени доступа к объекту из S по физическому указателю;

$M(\omega)$ - математическое ожидание времени поиска физического указателя во вспомогательной структуре.

Рассмотрим пример, иллюстрирующий полученные результаты.

Пример 1. Пусть таблица S содержит $K=100$ объектов, таблица R включает $N=1\dots 500$ объектов, а вероятности ссылок η_j равны между собой. Пусть также по условию F_R выбраны все строки, а по условию F_S - (а) 100%, (б) 75% и (в) 25% строк. Найти зависимость математического ожидания числа объектов таблицы S , которые удовлетворяют условию F_S и на которые ссылается хотя бы один объект из R , от N .

Решение. Условия, перечисленные в примере 1, соответствуют предпосылкам следствия 1. Поэтому для нахождения математического ожидания числа объектов отношения S используем выражение вида (15).

Зависимости математического ожидания числа выбранных объектов от количества объектов в таблице R для вариантов (а, б, в) представлены на рис. 2.

Для нахождения асимптотической прямой в окрестности нуля заменим дискретную переменную N_R на непрерывную переменную x и вычислим следующий предел:

$$\begin{aligned} \lim_{x \rightarrow 0} \frac{K_S \left(1 - \left(1 - \frac{1}{K} \right)^x \right)}{ax} &= \lim_{x \rightarrow 0} \frac{\frac{\partial}{\partial x} K_S \left(1 - \left(1 - \frac{1}{K} \right)^x \right)}{\frac{\partial}{\partial x} ax} = \\ \frac{K_S}{a} \left[- \lim_{x \rightarrow 0} \frac{\partial}{\partial x} \left(1 - \frac{1}{K} \right)^x \right] &= \frac{K_S}{a} \left[- \lim_{x \rightarrow 0} \left(1 - \frac{1}{K} \right)^x \ln \left(1 - \frac{1}{K} \right) \right] = \\ \frac{K_S}{a} [\ln(K) - \ln(K-1)] & \end{aligned} \quad (19)$$

Приравняв предел (19) к единице, найдем уравнение асимптотической прямой:

$$f(N_R) = K_S [\ln(K) - \ln(K-1)] N_R.$$

Отклонение кривых вниз от асимптоты показывает, что выражение (15) учитывает тот факт, что с увеличением числа объектов, отобранных из R , увеличивается повторная выборка некоторых объектов из отношения S .

Уравнение горизонтальной асимптоты имеет следующий вид:

$$g(N_R) = \lim_{N_R \rightarrow \infty} K_S \left(1 - \left(1 - \frac{1}{K} \right)^{N_R} \right) = K_S.$$

Горизонтальная асимптота показывает, что с увеличением выборки из R число выбранных объектов из S стремится к максимуму, равному K_S .

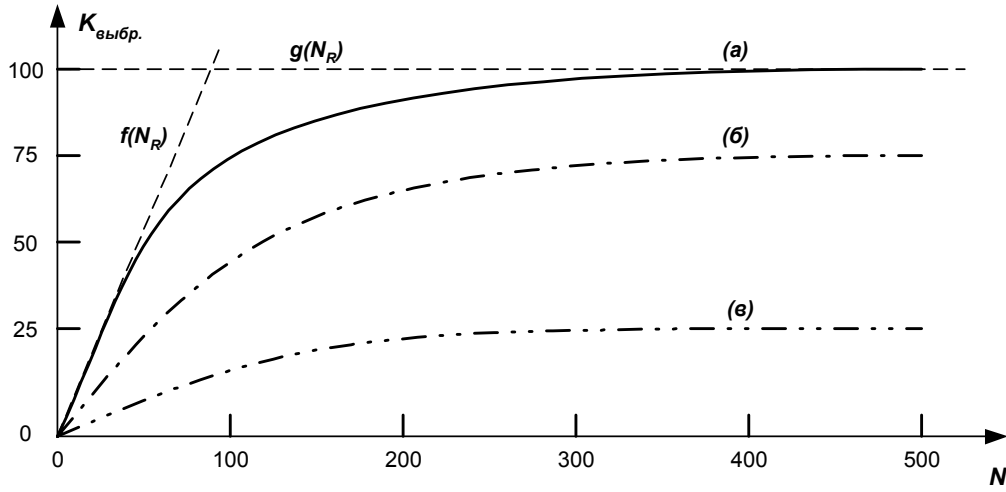


Рис. 2. Зависимость математического ожидания числа выбранных объектов

Оценка времени выполнения соединения Reverse Join

Для оценки времени выполнения запросов, выполняемых по алгоритму Reverse Join, можно воспользоваться леммой 2 и теоремой 2, которые сформулированы ниже. Они доказываются по аналогии с леммой 1 и теоремой 1.

Лемма 2. Пусть выполняются те же предпосылки, что и в случае леммы 1, а соединение выполняется по алгоритму Reverse Join, тогда:

1) ПФСВ числа выбранных объектов из отношения R определяется следующим выражением:

$$\Psi(z) = V_R \left(1 - \left(1 - \left[\sum_{k=1}^K \left[z^k \right] V_S(z) \binom{K}{k}^{-1} \sum_{(j_1, \dots, j_k) \in Q_k} \prod_{j=j_1, \dots, j_k} (1 - \eta_j) \right] \right) \right) (1 - z); \quad (20)$$

2) математическое ожидание числа выбранных объектов равно:

$$M(\Psi) = M(\chi_R) \left(1 - \left[\sum_{k=1}^K \left[z^k \right] V_S(z) \binom{K}{k}^{-1} \sum_{(j_1, \dots, j_k) \in Q_k} \prod_{j=j_1, \dots, j_k} (1 - \eta_j) \right] \right), \quad (21)$$

где $M(\chi_R)$ - математическое ожидание числа объектов в R .

Следствие 2. Пусть выполняются предпосылки леммы 2, вероятности ссылок η_j равны между собой, а домены D_{REF} и D_{OID} совпадают, тогда:

1) ПФСВ числа выбранных объектов из R имеет следующий вид:

$$\Psi(z) = \left(1 - \left(1 - \left(1 - \frac{1}{K} \right)^{K_S} \right) \right) (1 - z)^{N_R};$$

2) математическое ожидание числа объектов равно:

$$M(\psi) = N_R \left(1 - \left(1 - \frac{1}{K} \right)^{K_s} \right).$$

Теорема 2. Пусть выполнены предпосылки леммы 2, соединение выполняется по индексу и физическим указателям. Пусть также время доступа к объекту из отношения R по физическому указателю ($R_{OID} \rightarrow$ физический_адрес) является н.с.в. θ_{RJ} с ПЛС $\Theta_{RJ}(s)$, а время поиска указателя во вспомогательной структуре (индексе $R_{REF} \rightarrow R_{OID}$), является н.с.в. ω с ПЛС $\Omega(s)$, тогда:

1) ПЛС времени выполнения соединения Reverse Join определяется выражением

$$\Xi(s) = V_S(\Theta_{RJ}(s)\Omega(s));$$

2) математическое ожидание времени выполнения соединения:

$$M = M(\xi)(M(\theta) + M(\omega)),$$

где $M(\xi)$ - математическое ожидание числа объектов, отобранных из S по условию F_S ;

$M(\theta)$ - математическое ожидание времени доступа по указателю;

$M(\omega)$ - математическое ожидание времени поиска;

V_S - ПФСВ числа объектов, отобранных из S по условию F_S .

Заключение

В работе предложен математический метод оценки объемно-временных характеристик выполнения запросов к объектно-ориентированным базам данных. Предложенный метод учитывает специфику выполнения запросов в ООСУБД, а также параметры схемы БД и запросов.

Показано, что основной особенностью выполнения запросов к ООСУБД является наличие явно задаваемых ассоциативных связей между объектами классов и специальных алгоритмов, используемых для выполнения OQL-запросов. Выделены основные алгоритмы выполнения запросов: прямое и обратное соединение (Forward Join и Reverse Join).

Получено выражение для ПФСВ числа объектов, выбираемых из парного экстенда при выполнении запроса. Производящая функция получена для произвольного распределения числа объектов экстендов и объектных ссылок. Доказано следствие для частного случая вырожденного распределения числа объектов и ссылок.

Выражение для ПФСВ числа отобранных объектов было использовано для доказательства формулы ПЛС времени выполнения запросов по алгоритмам FJ и RJ. Также получена оценка математического ожидания времени выполнения запросов для каждого из этих алгоритмов.

Предложенный метод расчёта характеристик времени выполнения запросов к ООСУБД реализован в экспертной системе КСАМ (комплекс

инструментальных средств анализа моделей), позволяющей выполнять оценку распределенных систем обработки данных с узлами обработки на основе как реляционных, так и объектно-ориентированных СУБД.

ЛИТЕРАТУРА

1. *Андреев А., Березкин Д., Самарев Р.* Внутренний мир объектно-ориентированных СУБД//Открытые системы. 2001. №3(59). С. 47-57.
2. *Грэхем Р., Кнут Д., Паташник О.* Конкретная математика. Основания информатики: Пер. с англ. М.: Мир, 1998.
3. *Braumandl R., Claussen J., Kemper A.* Evaluating functional joins along nested reference sets in objectrelational and object-oriented databases. In Proc. of the Conf. on Very Large Data Bases (VLDB). New York. USA. August 1998. P. 110-121.
4. *Carey M.J., et.al.* O-O, What Have They Done to DB2, VLDB. 1999. P.542-553.
5. *Delis A., Roussopoulos N.* Performance and scalability of client-server database architectures. In Proceedings of the Eighteenth International Conference on Very Large Date Bases. Vancouver, British Columbia, August 1992.
6. *Flajolet P., Sedgewick R.* The Average Case Analysis of Algorithms: Counting and Generation Functions, 1993.
7. *Gardarin G., Sha F., Tang Z.H.* Calibrating the query optimizer cost model of IRO-DB, an objectoriented federated database system. In Proc. of the Conf. on Very Large Data Bases (VLDB), 1996. P.378-389
8. *Graefe G.* Query evaluation techniques for large databases. ACM Computing Surveys. Vol. 25. No. 2. 1993. P.73-170.
9. *Harris E.P., Ramamohanarao K.* Join algorithm costs revisited. VLDB Journal. Vol. 5(1). 1996. P.64-84.
10. *Ozsu M.T., Straube D.D., Peters R.* Query processing issues in object-oriented knowledge base systems. In F.E. Petry and L.M. Delcambre, editors, Volume 1 - Intelligent Database Technology: Approaches and Applications, Advances in Databases and Artificial Intelligence. JAI Press. 1994. P.79-144.
11. *Teeuw W. B., Rich C., Scholl M. H., Blanken H.M.* An Evaluation of Physical Disk I/Os for Complex Object Processing. Proceedings of the Ninth International Conference on Data Engineering, April 19-23, 1993, Vienna, Austria. IEEE Computer Society, 1993. P.363-371.
12. *Григорьев Ю.А., Плутенко А.Д.* Оценка времени выполнения запросов к реляционной системе управления базами данных. М.: Изд-во МГТУ им. Н.Э. Баумана, 2000.