

Моделирование распределенных неоднородных СУБД при решении разнородных высокодинамичных задач

Н.С. Дудаков, К.В. Макаров, С.А. Пулято

ОАО «РТИ»

При проектировании современных информационных систем одним из наиболее важных моментов является управление доступом, хранением и использованием информации. С увеличением объемов информации, сложности вычислений при анализе и обработке данных возрастают требования к СУБД. В связи с высокой стоимостью оборудования и использования корпоративных СУБД, при интенсивной нагрузке на информационную систему актуальной является задача разработки или подбора СУБД с учетом характеристик обрабатываемых данных.

Одним из вариантов решения задачи хранения и обработки данных при высокой нагрузке может быть оптимизация структуры СУБД при неизменных программных и аппаратных средствах. Далее, при возможности кластеризации обрабатываемых данных по темпам обновления и доступа на малосвязанные области, актуальным является использование распределенных неоднородных СУБД (РН СУБД). Компоненты РН СУБД, подобранные с учетом объема данных, интенсивности доступа к информации и различного времени жизни объектов обеспечивают лучшую производительность, чем универсальные решения [1],[2].

Применение РН СУБД при обработке данных с широким спектром характеристик обеспечивает возможность оптимизации структуры СУБД. Ключевым является инструментарий количественной оценки производительности СУБД, для чего предлагается использовать аппарат теории массового обслуживания. Так, с высокой степенью точности СУБД может быть представлена, как система массового обслуживания с пуассоновским входным потоком и рекуррентным временем обслуживания [3]. Время обслуживания предлагается считать дискретной случайной величиной, распределенной аналогично типу запроса (класса): вероятность поступления запроса к конкретному классу определяется следующим образом:

$$P(n_j) = \frac{\lambda_j}{\sum_{i=1}^N \lambda_i}, \quad (1)$$

где λ_j – интенсивность потока запросов объектов класса n_j . Далее, при распределении $n_j \in N, j = \overline{1, N}$ классов данных по $m_i \in M, i = \overline{1, M}$ компонентам РН СУБД для оценки эффективности работы компонента m_i с группой классов $n_j \in N_k \subseteq N$ предлагается использовать время ожидания выполнения запроса:

$$w_i = \frac{\sum_j b_{ij}^2 \lambda_j x_{ij}}{2(1 - \sum_j b_{ij} \lambda_j x_{ij})}, \quad (2)$$

где b_{ij} – время выполнения запроса класса n_j в компоненте РН СУБД m_i ;
 x_{ij} – переменная, определяющая факт обработки класса n_j в компоненте РН СУБД m_i . Для системы из M компонент в качестве критерия предлагается использовать суммарное реальное время пребывания запросов в системе:

$$J(X) = \sum_j \tilde{b}_{ij} = \sum_i \sum_j (b_{ij} + w_i) x_{ij}. \quad (3)$$

Соответственно, под оптимальной структурой РН СУБД понимается распределение классов по компонентам СУБД $X = \{x_{ij}\}_{i=\overline{1, M}, j=\overline{1, N}}$, обеспечивающее минимум критерия (3) на множестве всевозможных распределений X при ограничениях:

$$\left\{ \begin{array}{l} \forall i = \overline{1, M}: \sum_j (\lambda_j b_{ij} x_{ij}) < 1 \\ \forall j = \overline{1, N}: \sum_{i=1}^M x_{ij} = 1 \end{array} \right. , \quad (4)$$

а именно – обработке каждого класса в единственном компоненте и отсутствии накопления очереди запросов.

Задача оптимизации (3),(4) является задачей псевдоболева программирования. В качестве метода решения задачи предлагается сведение исходной нелинейной задачи к дробно-линейной путём замены каждого попарного произведения переменных $x_{i_1 j_2} * x_{i_3 j_4}$ на новую переменную $u_{i_1 j_2 i_3 j_4} \in \{0,1\}$ с добавлением дополнительных ограничений [4]. Далее, частично линейаризованная задача сводится к широко известной проблеме выполнимости булевых формул (SAT-задача). Суть преобразований при решении задачи оптимизации сводится к получению решения, удовлетворяющего ограничениям (построение приемлемого решения) и итерационному добавлению ограничения для отсекаемого найденного решения. Также, вариантом решения задачи оптимизации может быть поиск локального оптимального значения критерия (3) при начальном приближении, соответствующем минимуму загрузки компонент СУБД:

$$P(X) = \sum_i^M \sum_j^N b_{ij} \lambda_j x_{ij} \rightarrow \min_X. \quad (5)$$

Линейная зависимость выражения (5) от переменных x_{ij} обеспечивает быстрый поиск начального приближения, корреляция загрузки (5) и критерия (3) – быструю сходимость к минимуму критерия.

Практическая реализация распределения данных с использованием современных реляционных СУБД *PostgreSQL* и БД *SQLite* показывает значительный прирост производительности: по сравнению с использованием только клиент-серверной СУБД *PostgreSQL*, оптимальное распределение классов данных уменьшает среднее время выполнения запросов вдвое.

Таким образом, представленная модель позволяет количественно оценить производительность СУБД, используя математический аппарат теории массового обслуживания, а разработанная методика – построить оптимальное распределение данных по РН СУБД, значительно улучшая возможности СУБД по обработке данных за счет наилучшего подбора компонент СУБД в соответствии с характеристиками обрабатываемой информации.

Литература

1. Кузнецов С.Д. Основы современных баз данных. - Интернет-университет информационных технологий, 2007. 488 с.
2. Дейт, К.Дж. Введение в системы баз данных. - М.: Вильямс, 2001. 1328 с.
3. Клейнрок, Л. Теория массового обслуживания - М.: Машиностроение, 1979. 432 с.
4. Boros, E., Peter L. Hammer. Pseudo-Boolean Optimization // Discrete Applied Mathematics, 2002. V. 123. P. 155-225.