

Метод сокращения трафика в каналах передачи данных и увеличения загрузки узлов распределенной системы

Прихожий А.А., Труханович Т.Л.

Белорусский национальный технический университет

Сокращение трафика в каналах передачи данных и сокращение загрузки узлов достигается оптимизацией назначения объектов и их репликаций на узлы по критерию минимума среднего времени обработки одного запроса к объекту.

Пусть $N = \{1, \dots, n\}$ – множество узлов, $M = \{1, \dots, m\}$ – множество объектов. Время обработки запроса к объекту j , поступившего с узла i и обрабатываемого узлом k , можно оценить с достаточной степенью точности выражением $f_{ij} * (r_{kj} + c_{ik} * send(l_j)) * x_{kj}$, где f_{ij} – число запросов к объекту j с узла i в течение времени наблюдения $Time$, r_{kj} – время обработки на узле k запроса к объекту j , c_{ik} – временная задержка передачи единицы данных в канале (i, k) , l_j – длина объекта j , $send(l_j)$ – часть длины объекта, пересылаемая в ответ на запрос, x_{kj} – переменная, принимающая значение 1, если объект j размещается на узле k , и принимающая значение 0 в противном случае. Среднее время обработки запроса по всем узлам, всем объектам и всем запросам зависит в первую очередь от матрицы назначения X и может быть оценено выражением

$$T_{request}(D) = \frac{1}{\Sigma(F)} \sum_{i=1}^n \sum_{j=1}^m f_{ij} * (r_{d(i,j),j} + c_{i,d(i,j)} * send(l_j)),$$

где $d(i,j) \in N$ – узел, на котором обрабатывается запрос к объекту j , поступивший с узла i .

В задаче оптимизации целевая функция $\min_{D \in \Lambda} \{T_{request}(D)\}$ минимизирует время запроса по всем возможным размещениям объектов на узлах, где Λ – множество всех возможных корректных матриц D . Ограничениями являются ограничение на объем памяти, занимаемой объектами на одном узле, а также условие размещения каждого объекта хотя бы на одном узле.

Для решения задачи предложен алгоритм оптимизации, базирующийся на методе ветвей и границ. Разработана программа, минимизирующая среднее время запроса при варьировании всех основных входных параметров задачи оптимизации. Проведенные исследования показывают, что алгоритм сокращает трафик в каналах передачи данных и увеличивает загрузку узлов, обрабатывающих запросы к объектам. Он может быть использован при оптимальном проектировании распределенной системы.