РАЗМЕЩЕНИЕ РЕПЛИК ВЗАИМОСВЯЗАННЫХ МАССИВОВ ДАННЫХ В НЕНАДЕЖНЫХ РАСПРЕДЕЛЕННЫХ СИСТЕМАХ

[1]Микрин Е.А., [2]Сомов С.К.

[1] ПАО «Ракетно-космическая корпорация «Энергия» им. С.П. Королева»

Eugeny.Mikrin@rsce.ru

[2]Институт проблем управления им. В.А. Трапезникова РАН

ssomov2016@ipu.ru

Аннотация. Рассматривается задача оптимального размещения реплик взаимосвязанных массивов данных в распределенной системе обработки данных с ненадежными каналами связи. Критерий оптимизации задачи - минимум среднего времени ответа системы на запросы к данным. Предложен эвристический алгоритм решения сформулированной задачи.

Ключевые слова: распределенная система обработки данных, ненадежные каналы связи, оптимальное распределение реплик массивов данных.

Введение

К характеристикам распределенных систем обработки данных (РСОД) предъявляется много различных требований. Наиболее важными из них являются следующие: 1) высокая надежность работы системы, 2) приемлемое время обработки запросов к данным системы, 3) высокий уровень сохранности информации.

Применение в РСОД механизма репликации массивов данных, используемых в системе, позволяет выполнить данные требования. Метод репликации заключается в оптимальном размещении по некоторым узлам системы реплик массивов данных (идентичных копий массивов данных). Идентичность данных в репликах поддерживается механизмом репликации изменений в данных. Оптимальность размещения реплик может оцениваться по разным критериям оптимальности. Например, таким, как: минимальное время реакции системы на запрос пользователя; минимум стоимости функционирования системы, максимум вероятности получения ответа на запросы пользователей.

Задачи поиска оптимального размещения реплик по узлам распределенных систем обладают большой вычислительной сложностью (NP). В силу этой причины для решения таких задач применяются различные методы, которые уменьшают их вычислительную сложность [1-5].

В данной работе рассматривается задача поиска оптимального распределения реплик нескольких взаимосвязанных массивов данных в узлах РСОД с ненадежными каналами связи. Критерием оптимизации для данной задачи является минимум среднего времени реакции системы на запросы к массивам данных. При этом используется ограничение на максимальное количество реплик, размещенных в одном узле сети, и ограничение на максимальную стоимость функционирования распределенной системы.

1 Формальная модель распределенной системы

В различных узлах РСОД (пользовательские сервера) выполняется несколько прикладных процессов (задач) разных типов. Для каждого узла системы задана частота выполнения/решения в нем задач разного типа. В каждом узле системы может одновременно решаться несколько задач. Каждая задача генерирует определенный поток информационных запросов и/или запросов на модификацию массивов данных. Предполагаем, что во время решения задачи запросы генерируются и обрабатываются РСОД последовательно. Каждый информационный запрос обращается за данными реплики одного массива данных. Запрос на модификацию данных массива пересылается во все узлы с репликами этого массива. В случае, если решаемой задаче необходимы данные из нескольких массивов данных, то задача генерирует последовательно соответствующее количество запросов.

РСОД работает на базе компьютерной сети, которая состоит из N узлов. Топология сети описывается взвешенным графом $G = (X, \Gamma)$. Для дуг графа определены их длины. Для своей работы РСОД использует M $(m = \overline{1, M})$ массивов данных. Массивы разного типа, отличаются друг от друга набором атрибутов записей. Сеть передачи данных построена на базе Φ $(\varphi = \overline{1, \Phi})$ каналов связи. Для каждого φ-го канала связи задана пропускная способность C_{φ} единиц данных в единицу времени. Запросы и ответы на них пересылаются по каналам связи с помощью метода коммутации сообщений. В каналах сети передачи данных возможно появление ошибок. При передаче единицы данных по каналу φ связи единичной длины ошибка возникает с вероятностью $q_{\varphi}^{*} = 1 - \rho_{\varphi}^{*}$, $(\varphi = \overline{1, \Phi})$. Успешная доставка сообщения узлу-адресату подтверждается квитанцией АСК (*ACKnowledgement*). Обозначим через t_{ACK} - время ожидания узлом-отправителем квитанции АСК от узла-адресата. В случае, если за время t_{ACK} узел-отправитель не получил квитанцию, то производится повторная передача сообщения. Сделаем следующие предположения: размер квитанции АСК незначителен, она передается за пренебрежимо малое время, вероятность возникновения ошибки при передаче квитанции равна нулю.

В РСОД решается несколько пользовательских задач, каждая из которых принадлежит одному из J типов. Матрица $F^{*} = \|f_{nj}^{*}\|$ $(n = \overline{1, N};\ j = \overline{1, J})$ определяет частоту решения задач. Здесь f_{nj}^{*} это частота решения задачи j-го типа в n-м узле сети. Матрица $E^{*} = \|e^{*}{}_{jm}\|$, задает значения $e^{*}{}_{jm}$ – частот генерации задачей j-го типа информационных запросов к массиву m-го типа. Матрица $U^{*} = \|u^{*}{}_{jm}\|$ определяет частоты генерации запросов на модификацию данных. Здесь $u^{*}{}_{jm}$ - частота генерации задачей j-го типа запросов на модификацию массива данных m-го типа. Полагаем, что в узлах сети бесконечные очереди на обслуживание запросов, система работает в установившемся режиме, и все запросы обрабатываются успешно. Будем считать, что время T_{pr} обработки любого запроса в любом узле системы одинаково.

Возникший в некотором узле системы запрос на модификацию массива данных одновременно пересылается во все узлы, имеющие реплики данного массива. Информационный запрос из узла, в котором он сгенерирован, пересылается по кратчайшему пути в ближайший узел с требуемой репликой. Между всеми парами вершин взвешенного графа G определены кратчайшие пути с помощью одного из известных алгоритмов поиска кратчайших путей (например, алгоритмы Флойда — Уоршелла, Беллмана – Форда, [6]). Для подсчета длин путей в графе используется метрика - условная «длина» l_{φ} канала связи (дуги графа), значение которой рассчитывается по формуле: $l_{\varphi} = l_{\varphi}^{*} / (C_{\varphi} * q_{\varphi}^{*})$. Здесь l_{φ}^{*} - физическая длина канала связи, q_{φ}^{*} - вероятность возникновения ошибки в канале, C_{φ}- пропускная способность канала. В результате для графа G определяется матрица $SP = (sp_{nk})_{N \times N}$ кратчайших путей, где элемент sp_{nk} — это длина кратчайшего пути/маршрута между узлами n и k.

Элементы матрицы $A = \|a_{nm}\|, (a_{nm} \in \{0,1\})$ задают распределение реплик M массивов данных по узлам системы. При этом элемент a_{nm} равен 1, если в n-м узле размещена реплика m-го массива. Каждый массив данных может иметь несколько реплик, размещенных в разных узлах системы. Количество p_m реплик массива m равно: $p_m = \sum_{n=1}^{N} a_{nm}$.

Матрица $B = \|b_{nkm}\|$, имеющая размерность $N \times N \times M$, определяет «близость» узлов сети. Т. е. элемент матрицы $b_{nkm} \in \{0,1\}$, и $b_{nkm} = 1$, если в узле k размещена реплика массива данных m, и узел k является ближайшим узлом для узла n. Это означает, что инициированные в узле n информационные запросы к массиву m, адресуются для обработки в узел k.

Заданы значения: $\bar{l}_m^q$ - средней длины информационного запроса к массиву m-го типа, и $\bar{l}_m^r$ - средняя длина ответа на данный запрос. Аналогично заданы значения: средней длины $\bar{\bar{l}}_m^q$ запроса на модификацию m-го массива, и $\bar{\bar{l}}_m^r$ - средней длины ответа на данный запрос.

2 Формулировка задачи размещения реплик в узлах распределенной системы.

Необходимо сформулировать задачу размещения в узлах распределенной системы реплики M массивов данных таким образом, чтобы обеспечивалось минимальное значение функционала задачи. В качестве функционала $F(A)$ задачи будем использовать среднее время обработки запросов, которые генерируются задачами пользователей в единицу времени. Значение функционала *F(A)* рассчитывается по формуле:

$$(1) \quad F(A) = \tilde{T} + T_{pr} + \tilde{T}$$

здесь:

$\tilde{T}$ - Среднее время передачи запроса пользователя в узел с репликой требуемого массива данных.

T_{pr} – Среднее время обработки запроса сервером реплик.

$\tilde{T}$ – Среднее время передачи ответа на запрос до пользовательского сервера.

Следовательно, в задаче необходимо найти минимум функционала *F(A)*:

$$(2) \quad min\, F(A) = min\,(\tilde{T} + T_{pr} + \tilde{T})$$

В данной задаче используются следующие ограничения:

- ограничение на стоимость OP_{cost} функционирования системы:

$$(3) \quad OP_{cost} \le COT_{MAX}$$

- ограничение RN_{MAX_n} на максимальное количество реплик, размещенных в узлах системы.

$$(4) \quad \sum_{m=1}^{M} a_{nm} \le RN_{MAX_n}, \qquad n = \overline{1,N}$$

3 Величина средней задержки сообщений в системе

Определим значение средней задержки сообщений в рассматриваемой распределенной системе с учетом ненадежности каналов связи. Под сообщениями будем понимать запросы пользователей и ответы на эти запросы. Обмен квитанциями АСК не рассматриваем.

Рассмотрим два произвольных узла сети *i* и *k*. В каждом из этих узлов могут решаться задачи пользователей, которые генерируют запросы к репликам, и могут храниться реплики массивов данных. Информационные запросы адресуются из узла *i* в узел *k*, если узел *k* это ближайший узел для узла *i* и узел *k* имеет реплику нужного массива. В этом случае количество y_{ik}^q генерируемых в узле *i* запросов (информационных и запросов на модификацию данных) равно:

$$y_{ik}^q = \sum_{j=1}^{J} f_{ij}^* * \sum_{m=1}^{M} \left(e_{jm}^* * b_{ikm} + a_{km} * u_{jm}^*\right)$$

В тоже время узел *i* может хранить реплики массивов данных и получать запросы из узла *k*. В результате обработки полученных запросов узел *i* отправляет в узел *k* ответы с частотой, равной:

$$y_{ik}^r = \sum_{j=1}^{J} f_{kj}^* * \sum_{m=1}^{M} \left(e_{jm}^* * b_{kim} + a_{im} * u_{jm}^*\right)$$

За единицу времени из узла *i* в узел *k* будет передано сообщений:

$$y_{ik} = y_{ik}^q + y_{ik}^r$$

Тогда во всей сети в единицу времени генерируется трафик *Y* сообщений, величина которого равна:

$$Y = \sum_{i=1}^{N} \sum_{k=1}^{N} \left(y_{ik}^{q} + y_{ik}^{r}\right)$$

Средняя длина $1/\mu$ сообщений, передаваемых по каналам сети, равна:

$$(5) \quad 1/\mu = Y^{-1} \left\{ \sum_{i=1}^{N} \sum_{k=1}^{N} \sum_{j=1}^{N} f_{ij}^{*} \left[\sum_{m=1}^{M} a_{km} * d^{1} + \sum_{m=1}^{M} a_{im} * d^{2} \right] \right\},$$

$$d^{1} = \left(\bar{l}_{m}^{q} * e^{*}{}_{jm} * sp_{ik} + \bar{\bar{l}}_{m}^{q} * u^{*}{}_{jm} \right), d^{2} = \bar{\bar{l}}_{m}^{q} * e^{*}{}_{jm} * sp_{ki} + \bar{\bar{l}}_{m}^{r} * u^{*}{}_{jm}$$

Определим величину T^{*}средней задержки сообщений в системе, для чего воспользуемся результатами работ [7,8]. В данных работах показано, что в сетях со средней связностью любое сообщение, исходящее из внутреннего узла сети, имеет случайную длину с плотностью распределения:

$$p(b) = \mu e^{-\mu b} \; (b \geq 0), \text{где } \mu^{-1} \text{ это средняя длина сообщения.}$$

Это означает, что при передаче очередного сообщения по каналу связи, это сообщение имеет длину, величина которой распределена по показательному закону распределения. Время, затрачиваемое на передачу сообщения по каналам связи, также является независимой случайной величиной. Допустим, что все сообщения обрабатываются в соответствии с дисциплиной *FCFS* [7,9]. В этом случае каждый φ -й канал связи может быть рассмотрен как система массового обслуживания (СМО) вида $M|M|1|\infty$. Данная СМО имеет пуассоновский входной поток со средним значением γ_{ik} сообщений в секунду, а время обслуживания сообщения распределено по показательному закону со средним значением μ^{-1}/C_{φ} (бит в секунду). В соответствии со сделанными предположениями средняя задержка сообщения будет равна [7]:

$$(6) \qquad \sum_{i=1}^{N} \sum_{k=1}^{N} \frac{\gamma_{ik}}{\gamma} Z_{ik}$$

Здесь Z_{ik} – средняя задержка сообщения, которое передается по кратчайшему пути sp_{ik} из узла i в узел k.

Считаем, что если сообщения, передаваемые по пути sp_{ik}, используют при этом канал связи φ, то этот канал связи включен в данный путь. В этом случае считаем, что нам задано множество матриц $X^{\varphi} = \left\| x_{ik}^{\varphi} \right\|, (\varphi = \overline{1, \Phi}, \;\; i, k = \overline{1, N})$. Если в данной матрице элемент $x_{ik}^{\varphi} = 1$, то канал связи φ входит в состав пути sp_{ik}. Поток сообщений, который проходит через канал φ, имеет интенсивность λ_{φ}^{*}:

$$(7) \qquad \lambda_{\varphi}^{*} = \sum_{i=1}^{N} \sum_{k=1}^{N} \gamma_{ik} \, x_{ik}^{\varphi}$$

Обозначим как T_{φ}^{*} среднее время ожидания сообщения в очереди на обслуживание и на передачу его по φ-му каналу связи. Тогда величина Z_{ik} средней задержки сообщения, переданного из узла i в узел k по маршруту sp_{ik}, будет равна сумме времен T_{φ}^{*} его обслуживания во всех каналах связи, входящих в этот маршрут, т.е.:

$$Z_{ik} = \sum_{\varphi=1}^{\Phi} x_{ik}^{\varphi} * T_{\varphi}^{*}$$

Используя формулу (6) получим, что значение T^{*}средней задержки сообщения в сети будет равна:

$$T^{*} = \sum_{i=1}^{N} \sum_{k=1}^{N} \frac{\gamma_{ik}}{Y} \sum_{\varphi=1}^{\Phi} x_{ik}^{\varphi} * T_{\varphi}^{*}$$

Сменив порядок суммирования, получим:

$$T^{*} = \sum_{\varphi=1}^{\Phi} \frac{T_{\varphi}^{*}}{Y} \sum_{i=1}^{N} \sum_{k=1}^{N} \gamma_{ik} \, x_{ik}^{\varphi}$$

И с учетом формулы (7), будем иметь:

$$(8)\quad T^* = \sum_{\varphi=1}^{\Phi} \frac{\lambda_\varphi^*}{Y} T_\varphi^*$$

Определим величину T_φ^*. Каждый канал связи φ можно трактовать как СМО вида $M|M|1|\infty$. Такая СМО имеет пуассоновский входной поток с интенсивностью λ_φ^* и показательным законом распределения времени обслуживания со средним значением μ^{-1}/C_φ. Показано [7], что в такой СМО время пребывания заявки в системе (в φ -м канале) равно:

$$(9)\quad T_\varphi^* = 1\ /\ \left(\mu C_\varphi - \lambda_\varphi^*\right)$$

Формула (9) справедлива для сетей, имеющих надежные каналы связи. Тогда в сетях с надежными каналами связи средняя задержка T^* сообщения в сети будет равна:

$$T^* = \sum_{\varphi=1}^{\Phi} \frac{\lambda_\varphi^*}{\gamma} T_\varphi^* = \frac{1}{\gamma} \sum_{\varphi=1}^{\Phi} \frac{\lambda_\varphi^*}{\mu C_\varphi - \lambda_\varphi^*}$$

В распределенных системах с ненадежными каналами связи с вероятностью $q_\varphi^* = 1 - \rho_\varphi^*$ при передаче единицы данных в φ-м канале может возникнуть ошибка. В этом случае производится повторная передача сообщения. Очевидно, что при этом увеличивается общее время обслуживания сообщения в канале связи, и, следовательно, снижается его фактическая пропускная способность. Вероятность $q_\varphi = 1 - \rho_\varphi$ возникновения ошибки в φ-м канале при передаче сообщения со средней длиной $l = \mu^{-1}$ равна:

$$q_\varphi = 1 - \rho_\varphi = 1 - \left(1 - q_\varphi^*\right)^l = 1 - \left(\rho_\varphi^*\right)^l$$

Следовательно, среднее время t'_φ передачи сообщения длиной $l = \mu^{-1}$ по ненадежному φ-му каналу связи равно:

$$(10)\quad t'_\varphi = \rho_\varphi T_\varphi^* + \left(t_{ACK} + \rho_\varphi T_\varphi^*\right) q_\varphi \left(1 - q_\varphi\right)^{-1},\ \text{где}\ q_\varphi = 1 - \left(\rho_\varphi^*\right)^l$$

Обозначим фактическую пропускную способность ненадежного φ-го канала связи через C'_φ. Эта пропускная способность зависит от надежности канала, величины времени t_{ACK} и физической пропускной способности C_φ канала.

Значение среднего времени t'_φ передачи сообщения по ненадежному каналу определяется формулой (10). С другой стороны, это время равно отношению длины сообщения к фактической пропускной способности канала, т. е.:

$$t'_\varphi = \mu^{-1} / C'_\varphi$$

Тогда фактическая пропускная способность C'_φ ненадежного канала связи равна:

$$C'_\varphi = \mu^{-1}/t'_\varphi,$$

где t'_φ определяется по формуле (10).

Среднее время T'_φ задержки сообщения в φ-м ненадежном канале связи будет равно:

$$T'_\varphi = 1 / \left(\mu C'_\varphi - \lambda_\varphi^*\right)$$

Тогда, в системе с ненадежными каналами величина $\tilde{T}$среднего времени задержки сообщения при передаче по каналам сети равна:

$$(11)\quad \tilde{T} = \sum_{i=1}^{N} \sum_{k=1}^{N} \frac{\gamma_{ik}}{\gamma} \tilde{Z}_{ik}, \qquad \text{где} \quad \tilde{Z}_{ik} = \sum_{\varphi=1}^{\Phi} x_{ik}^{\varphi} * T'_\varphi$$

Здесь $\tilde{Z}_{ik}$ – значение средней задержки сообщения, передаваемого по ненадежным каналам связи сети из узла i в узел k.

4 Стоимость функционирования распределенной системы с ненадежными каналами.

Стоимость функционирования распределенной системы зависит от нескольких ее характеристик: количество реплик и их распределение по узлам системы, стоимость использования каналов связи для передачи данных, стоимость хранения реплик в узлах системы, интенсивность запросов к системе и стоимость их обработки.

Введем следующие обозначения:

$dst(x_i, x_k)$ – длина кратчайшего пути между узлами x_i и x_k, равная значению sp_{ik};

dtc –стоимость передачи данных по пути единичной длины;

v^e, v^u– средний объем данных, передаваемых по каналам связи при обработке информационного запроса и запроса на модификацию данных.

$S\_cst_m(x_k)$ – стоимость хранения в узле x_k одной реплики m-го массива данных;

$E\_cst_m(x_k)$, $U\_cst_m(x_k)$ – затраты на обработку информационного запроса и запроса на модификацию реплики m-го массива, размещенной в узле x_k;

$W = \{w_1, \dots, w_k, \dots, w_N\}$ – вектор «весов» узлов графа. Вес w_k узла k это средний объем данных, которыми узел k обменивается с другими узлами системы при обработке запросов, которые сгенерированы в этом узле за единицу времени. Вес w_k равен:

$$w_k = \sum_{j=1}^{J} f_{kj}^* \left(\sum_{m=1}^{M} e^*{}_{jm}\, v^e + \sum_{m=1}^{M} u^*{}_{jm}\, v^u \right)$$

Используя множество X всех узлов сети и матрицу A, можно получить M множеств X_m^r $(m = \overline{1, M})$, которые содержат номера узлов множества X, в которых размещены реплики M массивов данных. Одно множество X_m^r содержит p_m номеров узлов из множества X, в которых размещены p_m реплик массива m. В этих узлах производится обработка запросов к репликам.

Пусть $d^m(x_n, X_m^r)$ это минимальное расстояние от узла x_n множества X до одного из узлов множества X_m^r:

$$d^m(x_n, X_m^r) = \min_{x_j \in X_m^r} dst(x_n, x_j) \ , \quad m = \overline{1, M}$$

Для множеств X_m^r определим σ - передаточное число, равное стоимости обмена данными между узлами множества X_m^r и узлами множества X в процессе обработки запросов в узлах из X_m^r:

$$\sigma(X_m^r) = \sum_{n=1}^{I} d^m(x_n, X_m^r) * dtc * w_n$$

Затраты системы $Cost_{dtr}$ на обмен данными по каналам связи сети при обработке запросов к репликам всех M массивов данных в единицу времени равны:

$$Cost_{dtr} = \sum_{m=1}^{M} \sigma(X_m^r)$$

Величина затрат на хранение всех реплик, размещенных в узлах системы равны:

$$Cost_{St} = \sum_{n=1}^{N} \sum_{m=1}^{M} S_{cst_m}(x_k) * a_{nm}$$

Стоимость обработки всех запросов (информационных и на модификацию данных), которые генерируются в системе, равна:

$$Cost_{EU} = \sum_{n=1}^{N} \sum_{j=1}^{J} f_{nj}^* * \sum_{m=1}^{M} \left(e_{jm}^* * \sum_{k=1}^{N} b_{nkm} * E_{cst_m}(x_k) + u_{jm}^* \sum_{k=1}^{N} a_{km} * U\_cst(x_k) \right)$$

Суммарная величина *OP_cost* затрат на функционирование распределенной системы равна:

$$\text{OP\_cost} = Cost_{dtr} + Cost_{St} + Cost_{EU}$$

5 Алгоритм решения задачи.

Для решения сформулированной задачи (2)-(4), имеющей большую вычислительную сложность, предлагается использовать эвристический алгоритм. Блок схема алгоритма представлена на рисунке 1, а детальное описание алгоритма дано ниже рисунка 1.

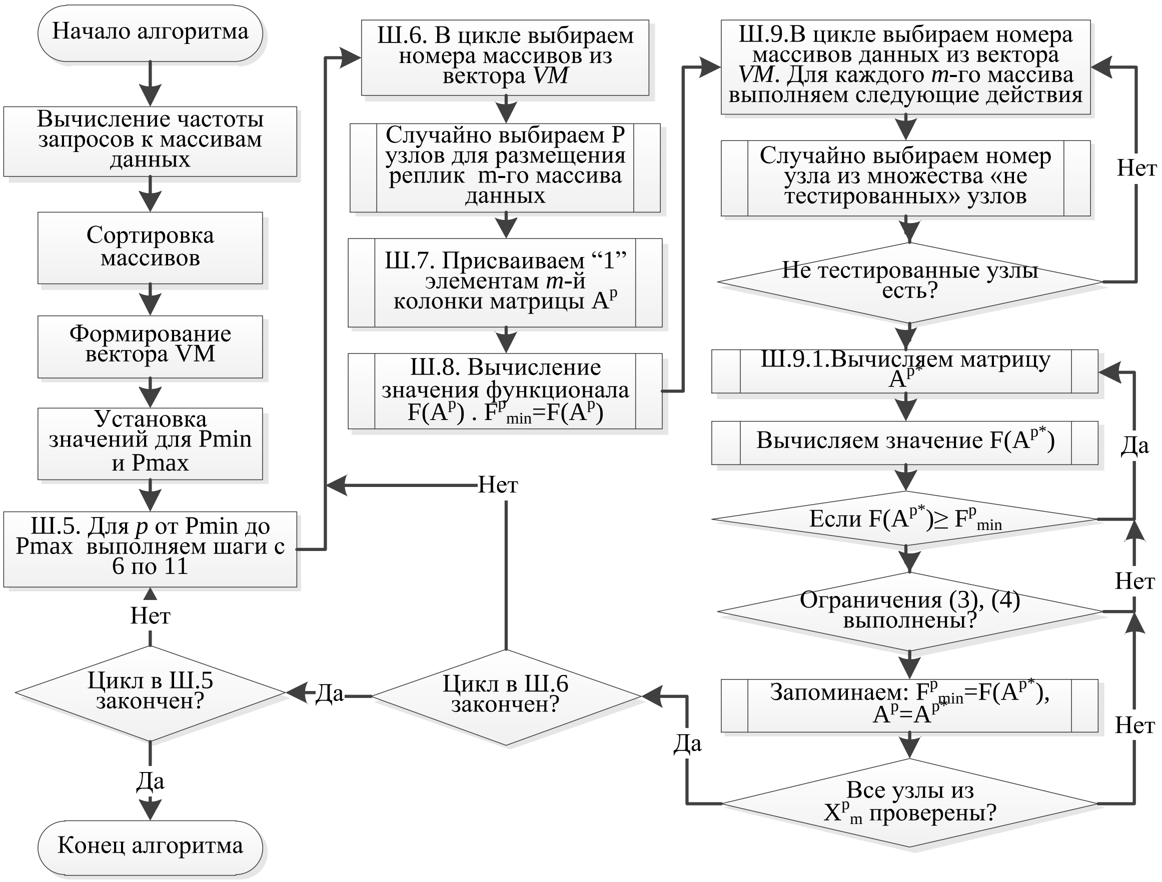

Рис. 1. Блок-схема алгоритма решения задачи

Описание алгоритма решения.

Шаг 1. Вычисляем значение fr_m - частота запросов к m-му массиву данных, генерируемых при решении задач в узлах системы:

$$fr_m = \sum_{n=1}^{N} f^*_{nj} \left(v^*{}_{jm} + u^*{}_{jm}\right); \;\; m = \overline{1, M}$$

Шаг 2. Сортируем номера массивов данных в порядке убывания соответствующих значений частот fr_m $(\mathrm{m} = \overline{1, \mathrm{M}})$.

Шаг 3. Заполняем вектор $VM = \langle vm_k \rangle, k = \overline{1, \mathrm{M}}$. При этом первый элемент вектора vm_1 принимает значение номера того массива, к которому генерируется наибольшее количество запросов. Последний элемент vm_M вектора имеет значение, равное номеру массива данных, к которому создается наименьшее количество запросов.

Шаг 4. Определяем значения ограничений на минимальное P_{min} и максимальное P_{max} количество реплик каждого из M массивов данных, которые могут быть размещены в узлах сети.

Шаг 5. В цикле по количеству p узлов с репликами, принимающем значения в интервале от P_{min} до P_{max}, выполняем шаги с Шаг 6 по Шаг 11:

Шаг 6. В цикле шаг за шагом выбираем номера vm_k массивов из вектора VM.

Для каждого очередного m-го массива выполняются следующие действия:

Случайным образом из множества X выбираются p узлов для распределения в них реплик m-го массива.

Номера этих узлов формируют множество $X_p^m = \{x_{pi}^m \,|i = \overline{1, p}\}$. Каждый элемент x_{pi}^m содержит номер узла сети, в котором размещается реплика m-го массива. Номера этих узлов запоминаются в множестве X_m «m-протестированных» узлов. Остальные номера узлов множества X, которые не вошли в множество X_p^m, запоминаются в $\bar{X}_m$ - множестве номеров «не протестированных» узлов. Т.е.

$$\bar{X}_m = \{X \backslash X_p^m\}.$$

Шаг 7. Заполняем значением "*1*" те элементы a_{nm}^{p} в m-ом столбце матрицы A^{p}, у которых номер n строки матрицы совпадает с номером x_{pi}^{m} узла сети с репликой. Данная матрица $A^{p} = \|a_{nm}^{p}\|$ описывает распределение p реплик массивов данных по узлам сети. При этом элемент $a_{nm}^{p} \in \{0,1\}$ равен *"1"*, если реплика m-го массива размещена в узле n.

Шаг 8. Рассчитываем значение функционала $F(A^{p})$. Для расчета используем формулу (1).

Полученное значение запоминаем: $F_{min}^{p} = F(A^{p})$

Шаг 9. Последовательно выбираем номера vm_k $(k = \overline{1, \mathrm{M}})$массивов данных из вектора VM.

Для каждого очередного выбранного m-го массива выполняются следующее:

Из множества $\bar{X}_m$ номеров случайным образом выбираем номер узла сети. Запоминаем номер в переменной x_i^{m*}.

Если все узлы из $\bar{X}_m$ выбраны, то возвращаемся к Шаг 9.

Шаг 9.1. В цикле каждую вершину x_{pi}^{m} из $X_p^m = \{x_{pi}^{m} \mid i = \overline{1,p}\}$ заменяем на вершину x_i^{m*}.

- В результате получаем множество X_p^{m*} в котором одна из вершин заменена на x_i^{m*}.
- Копируем матрицу A^{p} в матрицу A^{p*}.
- По аналогии с пунктом Шаг 7 присваиваем значение "*1*" элементам a_{nm}^{p*} m-го столбца матрицы A^{p*} в соответствии со значениями элементов множества X_p^{m*}.
- Вычисляем значение функционала $F(A^{p*})$.
- Если $F(A^{p*}) \geq F_{min}^{p}$ то возврат к Шагу 9.1.
- Проверяем ограничения (3) и (4) для распределения реплик A^{p*}.
- Если одно или оба ограничения не выполнены, то возврат к Шаг 9.1.
- Запоминаем новое полученное значение функционала и новое распределение реплик:

- $F_{min}^{p} = F(A^{p*}), A^{p} = A^{p*}$

- Если цикл по множеству X_p^m не завершен, то возврат к Шаг 9.1.

Иначе возврат к Шаг 9.

Шаг 10. Если не все элементы в VM просмотрены, то возврат к Шаг 6.

Шаг 11. Если цикл по количеству p реплик не завершен $(p < P_{max})$, то возврат к Шаг 5.

Шаг 12. Поиск закончен. Найдено близкое к оптимальному распределение A^{p} реплик по узлам распределенной системы. Этому распределению соответствует значение среднего времени ответа системы на запрос, равное F_{min}^{p}.

Конец работы.

Представленный алгоритм реализован на языке программирования C++ (MS Visual Studio). Алгоритм используется в автоматизированном комплексе поддержки принятия решений по обеспечению сохранности данных в распределенных системах.

Заключение

В работе рассмотрена задача оптимального размещения реплик нескольких массивов данных по узлам распределенной системы с ненадежными каналами связи. Формулировка задачи обладает несколькими особенностями. В частности, при возникновении сбоя в ненадежном канале связи выполняется повторная отправка сообщения в узел – адресат. Повторная отправка сообщений снижает производительность всей системы. В узлах распределенной системы выполняется несколько прикладных процессов (задач) разных типов по обработке запросов пользователей. В ходе решения каждой задачи генерируется некоторое множество информационных запросов и запросов на модификацию данных в репликах различных массивов данных. Поэтому в формулировке задачи поиска оптимального распределения реплик учитывается необходимость одновременного доступа отдельных прикладных процессов к некоторому подмножеству реплик. В качестве критерия оптимизации в задаче применяется величина среднего времени задержки сообщения (запросы к репликам и ответы на эти запросы) при передаче его по ненадежным каналам сети. Для решения сформулированной оптимизационной задачи, которая имеет большую вычислительную сложность, предложен оригинальный эвристический алгоритм, реализованный на языке C++.

Использование данного алгоритма позволяет улучшать время реакции системы на запросы на 5-8% и сокращать издержки на эксплуатацию системы на 8-10%.

Литература

1. *Чернышев Г. А.* Обзор подходов к организации физического уровня в СУБД // Труды СПИИРАН. 2013. - Санкт-Петербург, 2013. Вып. 1(24). – С. 222 – 275.
2. *Azzam Sleit and oth.* A Dynamic Object Fragmentation and Replication Algorithm In Distributed Database Systems // American Journal of Applied Sciences. 2007.Vol. 4 (8). P.613-618.
3. *Loukopoulos T., Ahmad I., Papadias D.* An Overview of Data Replication on the Internet // Proceedings of the International Symposium on Parallel Architectures, Algorithms and Networks (ISPAN.02). 2002. 6 p.
4. *Sahoo J., Salahuddin M.A., Glitho R.* A Survey on Replica Server Placement Algorithms for Content Delivery Networks // IEEE Communications Surveys & Tutorials. 2016. – P. 30. URL: https://arxiv.org/ftp/arxiv/papers/1611/1611.01729.pdf (дата обращения 2018-07-06).
5. *Mansouri N.* Adaptive data replication strategy in cloud computing for performance improvement // Frontiers of Computer Science. 2016. — Vol. 10, N 5. — P. 925-935.
6. *Cormen T.H., et al*. Introduction to Algorithms, Third Edition // The MIT Press, 2009. — 1313 p.
7. *Kleinrock L.* Queueing systems. Volume II: Computer applications // New York: Wiley, 1976. – 569p
8. *Kleinrock L.* Communication Nets: Stochastic Message Flow and Design // McGraw-Hill, 1964. — 220 p.
9. *Hamdy A. Taha.* Operations Research: An introduction. 8-th Edition // Pearson Prentice. Hall., 2007. — 838 p.