

## Математическая модель оптимального размещения распределённой базы данных по узлам ЛВС на базе двухуровневой клиент-серверной архитектуры

*А.Н. Скоба, Е.В. Состина*

*Южно-Российский государственный политехнический университет (НПИ) им.*

*М. И. Платова, Новочеркасск*

**Аннотация:** в данной статье с использованием аппарата замкнутых экспоненциальных сетей массового обслуживания (СеМО) разработана математическая модель решения задачи об оптимальном размещении распределённой базы данных (РБД) по узлам локальной вычислительной сети (ЛВС) на базе двухуровневой клиент-серверной архитектуры по критерию минимума среднего времени реакции системы на запросы пользователей. Приведены результаты численных экспериментов.

**Ключевые слова:** распределённая база данных, селективная выборка информации, *SQL*-запрос, распределённая информационная система, транзакция, пространство состояний системы, стационарная вероятность, переходная вероятность, интенсивность обслуживания, матрица объёмов информации, время реакции системы.

Отличительная особенность разрабатываемой модели от модели, представленной в работе [1] состоит в том, что на сервере производится селективная выборка информации, и по каналу связи передаётся не полная база данных, а некоторое количество кортежей базы данных, удовлетворяющих условию поиска *SQL*-запроса [2-4]. Дополнительно к введённым при разработке модели распределённой информационной системы на базе файл-серверной архитектуры, исходным данным [1]:

$$\begin{aligned} U &= \{U_1, \dots, U_s, \dots, U_n\}; & A &= \{A_1, \dots, A_s, \dots, A_n\}; & R &= \{R_1, \dots, R_j, \dots, R_d\}; & \lambda &= \{\lambda_1, \dots, \lambda_s, \dots, \lambda_n\}; \\ Q &= \{Q_1, \dots, Q_b, \dots, Q_q\}; & V &= \{V_1, \dots, V_j, \dots, V_d\}; & VV &= \{VV_1, \dots, VV_s, \dots, VV_n\}; \\ VD &= \{VD_1, \dots, VD_s, \dots, VD_n\}; & PU &= \{PU_1, \dots, PU_s, \dots, PU_n\}; & \theta &= \theta_0; & \alpha &= \alpha_0; \\ F &= \left\| f_{A,Q_l} \right\|, (s = \overline{1, n}, l = \overline{1, q}); & B &= \left\| b_{Q,R_j} \right\|, (l = \overline{1, q}, j = \overline{1, d}); \end{aligned}$$

$\delta(B) = \|\delta_{ij}\|, (i = \overline{1, q}, j = \overline{1, d})$ , вводится матрица объёмов информации, получаемой после процессорной обработки:  $\bar{B} = \|\bar{b}_{Q,R}\|, (i = \overline{1, q}, j = \overline{1, d})$ .

При этом, дополнительно к концептуальной модели информационной системы на базе файл-серверной архитектуры [1], для сборки ответов на элементарные транзакции вводятся буферные памяти пользователей:  $BPA_1, \dots, BPA_s, \dots, BPA_n$  ёмкостью  $d$  и ключи:  $K_1, \dots, K_s, \dots, K_n$ . Каждый из ключей может находиться в двух состояниях: разомкнутом и замкнутом. В первом состоянии он находится до тех пор, пока в  $BPA_s, (s = \overline{1, n})$  не будет собран ответ на запрос пользователя  $A_s$ , т.е. будут получены ответы на все элементарные транзакции. После этого ключ  $K_s$  переходит в замкнутое состояние, и собранный ответ на весь запрос выводится на пользовательский терминал. Количество одновременно обслуживаемых в системе сообщений, с учётом “расщепления” запроса на элементарные транзакции, не может превосходить величины  $nd$ , а в период пиковых нагрузок точно совпадает с этой величиной. Концептуальная модель функционирования этой системы представлена на рис. 1.

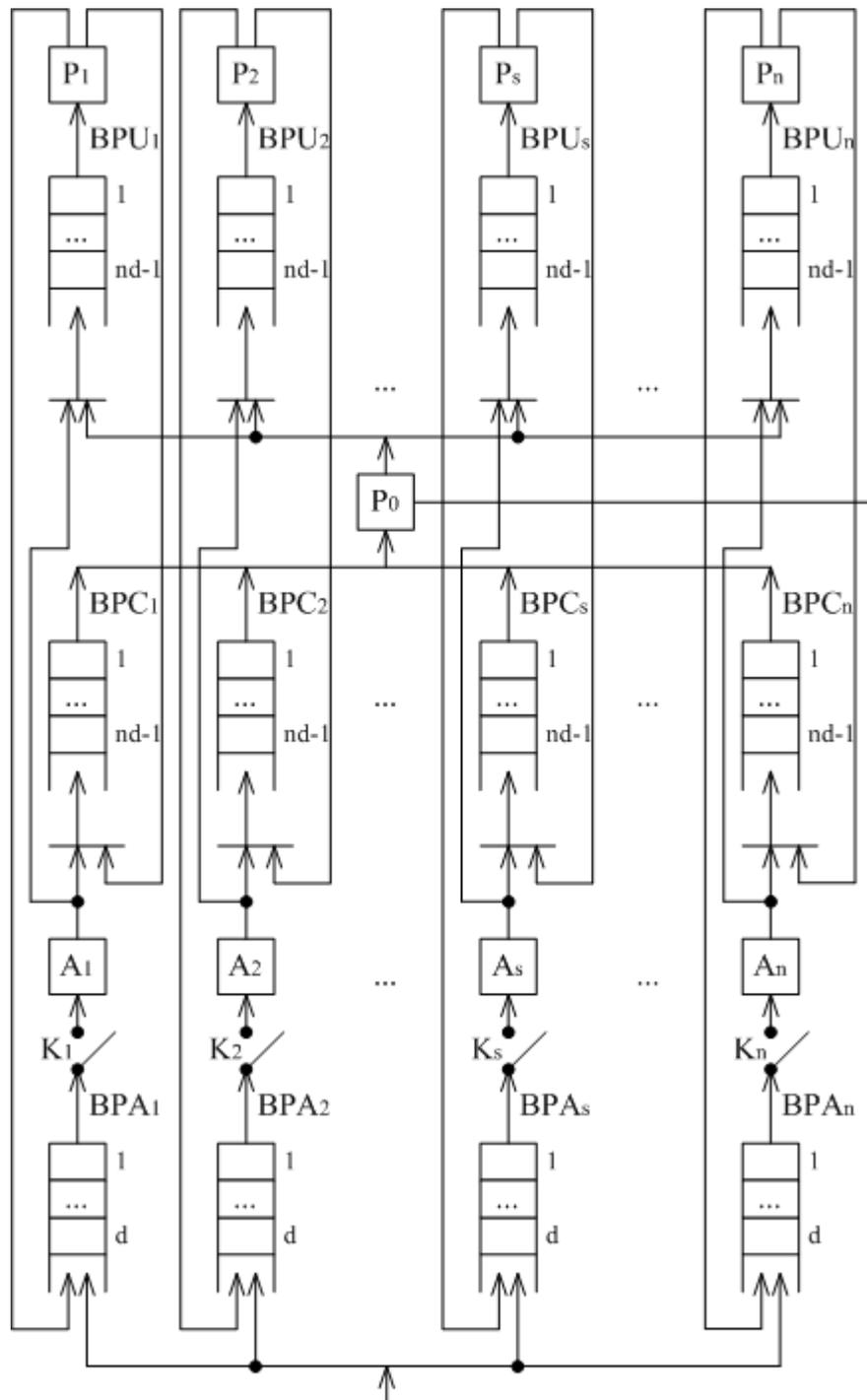


Рис. 1. Концептуальная модель системы на базе архитектуры «клиент-сервер»

По сравнению с моделью информационной системы на базе файл-серверной архитектуры [1], пространство состояний данной модели

модифицируется:

$E\{E_i(i_{01}, \dots, i_{0s}, \dots, i_{0n}; i_{11}, \dots, i_{1s}, \dots, i_{1n}; \dots; i_{2n+1,1}, \dots, i_{2n+1,s}, \dots, i_{2n+1,n}), i = \overline{1, \rho}\}$ , где  $\{i_{0s}, s = \overline{1, n}\}$  описывают состояние буфера  $BPA_s$ , при этом  $i_{0s}$  – количество обслуженных транзакций  $s$ -го пользователя;  $\{i_{1s}, s = \overline{1, n}\}$  – описывают состояние  $s$ -го пользователя, причём

$$i_{1s} = \begin{cases} d, \text{ если } s\text{-й пользователь находится в активном состоянии} \\ \text{(формирует запрос),} \\ 0 \leq i_{1s} \leq d, s\text{-й пользователь находится в пассивном состоянии} \\ \text{(ждёт ответа на запрос).} \end{cases}$$

Ограничения имеют следующий вид:

$$1) \sum_{k=0}^{2n+1} \sum_{r=1}^n i_{kr} = nd;$$

$$2) \sum_{k=0}^{2n+1} i_{kr} = d, r = \overline{1, n};$$

$$3) \forall s = \overline{1, n}, 0 \leq i_{0s} \leq d; 0 \leq i_{1s} \leq d; 0 \leq i_{0s} + i_{1s} \leq d; i_{ks} \geq 0, \forall k = \overline{0, 2n+1}.$$

Расчёт стационарных вероятностей состояний сети  $P(\vec{i})$  аналогичен расчёту данных величин для информационной системы на базе файл-серверной архитектуры [1] и производится на основе решения уравнения глобального баланса [5-9].

При расчёте величин  $\mu_{sr}, (s = \overline{1, 2n+1}, r = \overline{1, n})$  следует учитывать тот факт, что применение клиент-серверной архитектуры приводит к вариации содержания процедур обработки второй и третьей групп запросов, а обслуживание заявок первой группы не изменяется по отношению к модели информационной системы на базе файл-серверной архитектуры [1,10]. В частности, при выполнении заявок первой группы производится селективная выборка информации из баз данных, размещённых в  $U_s$ -м узле, а затем её процессорная обработка. Обслуживание запросов второй группы в  $U_s$ -м узле

состоит в выполнении операций процессорной обработки над кортежами, которые были считаны в других узлах и переданы в  $U_s$ -й узел по каналу связи. Реализация запросов третьей группы включает селективную выборку информации из баз данных, размещённых в  $U_s$ -м узле, и передачу кортежей отношений, удовлетворяющих условиям поиска, пользователям, прикрепленным к другим узлам. В соответствии с этим, интенсивности обслуживания запросов в узлах сети, по сравнению с [1] модифицируются следующим образом:

$$\mu_{sr} = \begin{cases} \left( \theta_0 + \frac{\sum_{l=1}^q f_{rl} \sum_{j=1}^d \bar{b}_{lj} x_{jz}}{\theta} \right)^{-1}, & r = \overline{1, n}, z = \overline{1, n}, s = 1 + z, z \neq r; \\ \left( \alpha_0 + \frac{\sum_{l=1}^q f_{rl} \sum_{j=1}^d \delta_{lj} v_j x_{jz}}{VV_z} + \frac{\sum_{l=1}^q f_{rl} \sum_{j=1}^d \delta_{lj} v_j x_{jz}}{PU_z} + \frac{\sum_{l=1}^q f_{rl} \sum_{\substack{k=1 \\ k \neq z}}^n \sum_{j=1}^d \bar{b}_{lj} x_{jk}}{VD_z} \right)^{-1}, & \begin{matrix} s = \overline{1, 2n+1}, \\ r = \overline{1, n} \end{matrix} \\ \left( \alpha_0 + \frac{\sum_{l=1}^q f_{rl} \sum_{j=1}^d \delta_{lj} v_j x_{jz}}{VV_z} + \frac{\sum_{l=1}^q f_{rl} \sum_{j=1}^d \delta_{lj} v_j x_{jz}}{PU_z} \right)^{-1}, & \begin{matrix} r = \overline{1, n}, z = \overline{1, n}, s = n + 1 + z, r = z; \\ r = \overline{1, n}, z = \overline{1, n}, \\ s = n + 1 + z, r \neq z. \end{matrix} \end{cases}$$

Элементы матрицы переходных вероятностей  $\|P_{ik}(s)\|$  для запросов  $s$ -го пользователя  $(i, k = \overline{1, 2n+1}, s = \overline{1, n})$ , по сравнению с моделью, представленной в [1], модифицируются следующим образом:

$$P_{ik}(s) = \left\{ \begin{array}{l} 1 - \sum_{j=1}^d \left\{ \frac{\sum_{l=1}^q f_{sl} \delta_{lj}}{\sum_{l=1}^q \sum_{j=1}^d f_{sl} \delta_{lj}} \right\} x_{js}, i = 1, k = s + 1; \\ \sum_{j=1}^d \left\{ \frac{\sum_{l=1}^q f_{sl} \delta_{lj}}{\sum_{l=1}^q \sum_{j=1}^d f_{sl} \delta_{lj}} \right\} x_{js}, i = 1, k = n + s + 1; \\ 1, \text{ если } \sum_{j=1}^d \left\{ \frac{\sum_{l=1}^q f_{sl} \delta_{lj}}{\sum_{l=1}^q \sum_{j=1}^d f_{sl} \delta_{lj}} \right\} x_{jz} \neq 0 \text{ и } z = \overline{1, n}, i = z + 1, i \neq s + 1, k = 1; \\ z = \overline{1, n}, i = n + 1 + z, z \neq s, k = z + 1 \text{ и } z = s, k = 1; \\ \frac{\sum_{j=1}^d \left\{ \frac{\sum_{l=1}^q f_{sl} \delta_{lj}}{\sum_{l=1}^q \sum_{j=1}^d f_{sl} \delta_{lj}} \right\} x_{jz}}{\sum_{j=1}^d \left\{ \frac{\sum_{l=1}^q f_{sl} \delta_{lj}}{\sum_{l=1}^q \sum_{j=1}^d f_{sl} \delta_{lj}} \right\} x_{js}} \neq 1, \text{ при } i = s + 1, z = \overline{1, n}, \\ 1 - \sum_{j=1}^d \left\{ \frac{\sum_{l=1}^q f_{sl} \delta_{lj}}{\sum_{l=1}^q \sum_{j=1}^d f_{sl} \delta_{lj}} \right\} x_{js} \\ k = n + 1 + z, k \neq n + s + 1; \\ 0, \text{ во всех остальных случаях.} \end{array} \right.$$

Расчёт среднего времени реакции системы производится по формуле

$$\bar{T} = \left( \frac{1}{\sum_{s=1}^n \lambda_s} \right) \sum_{s=1}^n \lambda_s \bar{T}_s,$$

где  $\lambda_s (s = \overline{1, n})$  – интенсивность формирования запросов  $s$ -м пользователем;  
 $\bar{T}_s (s = \overline{1, n})$  – среднее время реакции системы на запрос  $s$ -го пользователя.

Величину  $\bar{T}_s$  определим как  $\bar{T}_s = \frac{\bar{N}_s}{\bar{\lambda}_s}$ , где  $\bar{N}_s, (s = \overline{1, n})$  – среднее количество запросов  $s$ -го пользователя;  $\bar{\lambda}_s, (s = \overline{1, n})$  – средняя интенсивность формирования запросов  $s$ -м пользователем. Величины  $\bar{N}_s$  и  $\bar{\lambda}_s$  определим как:  $\bar{N}_s = 1 - P_s(d)$ ,  $\bar{\lambda}_s = \lambda_s P_s(d)$ , где  $P_s(d)$  – вероятность того, что  $s$ -й пользователь находится в активном состоянии (формирует запрос). Расчёт величины  $P_s(d)$  аналогичен расчёту величины  $P_s(l)$  для информационной системы на базе файл-серверной архитектуры с простым типом запросов [1]. Аналогично работе [1], расчёт величины  $\bar{T}$  также по существу сводится к расчёту нормализующей константы  $G(N_1, \dots, N_n)$ , для вычисления которой был использован рекуррентный метод Бузена [6, 11].

Для решения задачи оптимального размещения РБД по узлам ЛВС был использован эвристический алгоритм, приведённый в работе [1].

В таблице №1 приведены некоторые результаты машинных экспериментов.

Таблица №1

Результаты машинных экспериментов

| Размерность задачи $n \times d \times q$ | Начальное значение $\bar{T}$ | Число итераций МПП | Значение $\bar{T}^{(r)}$ | Время решения задачи МПП, с | Число итераций ЭА | Значение $\bar{T}^{(n)}$ | Время решения задачи ЭА, с | $\Delta, \%$ |
|--|------------------------------|--------------------|--------------------------|-----------------------------|-------------------|--------------------------|----------------------------|--------------|
| 3x4x5                                    | 7,1426                       | 81                 | 2,6875                   | 8,8                         | 2                 | 2,8603                   | 3,72                       | $\approx 6$  |
| 6x8x10                                   | 2,1347                       | $6^8$              | 0,9712                   | 223                         | 5                 | 1,1023                   | 34,11                      | $\approx 14$ |
| 8x13x15                                  | 3,1807                       | $8^{13}$           | -                        | -                           | 9                 | 2,1619                   | 192,2                      | -            |
| 10x15x20                                 | 2,1845                       | $10^{15}$          | -                        | -                           | 14                | 0,1826                   | 1332,5                     | -            |

Здесь МПП – метод полного перебора; ЭА – эвристический алгоритм;  $\bar{T}^{(r)}$  – среднее время реакции системы для оптимального размещения РБД,

полученного МПП;  $\bar{T}^{(n)}$  – среднее время реакции системы для оптимального размещения РБД, полученного ЭА;  $\Delta$  – относительная погрешность, выраженная в % между оптимальным размещением, полученным МПП и оптимальным размещением, полученным с помощью ЭА.

Разработанная модель оптимального размещения РБД по узлам ЛВС на базе двухуровневой клиент-серверной архитектуры также может быть использована при внедрении интегрированных информационно-справочных систем на промышленных предприятиях.

### Литература

1. Скоба А.Н., Состина Е.В. Математическая модель оптимального размещения распределённой базы данных по узлам ЛВС на базе файл-серверной архитектуры. // Инженерный вестник Дона. 2015. №2. URL:ivdon.ru/ru/magazine/archive/n2y2015/2881.
2. Богуславский Л.Б., Ляхов А.И. Оценка производительности распределённых информационно-вычислительных систем архитектуры «КЛИЕНТ-СЕРВЕР» // Автоматика и телемеханика.-1995.-С.160-175.
3. Воробьёв С.П., Горобец В.В. Исследование модели транзакционной системы с репликацией фрагментов базы данных, построенной по принципам облачной среды // Инженерный вестник Дона. 2012. №4. URL:ivdon.ru/ru/magazine/archive/n4y2012/1149.
4. Павлов С.В., Самойлов А.С. Проектирование структуры распределённой базы пространственных данных в сложно структурированных иерархических географических информационных системах // Инженерный вестник Дона. 2015. №1. URL:ivdon.ru/ru/magazine/archive/n1y2015/2755.
5. Вишневский В.М. Теоретические основы проектирования компьютерных сетей.- М.: Техносфера, 2003.- 512 с.

6. Жожикашвили В.А., Вишневецкий В.М. Сети массового обслуживания. Теория и применение к сетям ЭВМ. - М.: Радио и связь, 1988.-192с.
7. Герасимов А.И. Теория и практическое применение стохастических сетей. М.: Радио и связь., 1994. 175с.
8. Antunes C.H. et al. A Multiple Objective Routing Algorithm for Integrated Communication Network // Proc. ITC-16.-1999.V.3b. pp.1291-1300.
9. Chakka R., Harrison P.G. A Markov modulated multi-server queue with negative customers –The MM CPP/GE/c/LG-queue // Acta Informatika.2001. v.37. pp.785-799.
10. Черноморов Г.А. Теория принятия решений: Учебное пособие / Юж.-Рос.гос. техн.ун-т.-3-е изд.перераб. и доп.-Новочеркасск : Ред. журн. «Изв. Вузов. Электроомеханика», 2005. 448с.
11. Круглый З.Л. Алгоритмы расчёта моделей структур вычислительных систем с различными классами заданий // Управляющие системы и машины.- 1980. №4. С.73-79.

### References

1. Skoba A.N., Sostina E.V. Inzhenernyj vestnik Dona (Rus), 2015. №2.URL:ivdon.ru/ru/ magazine/archive/n2y2015/2881.
2. Boguslavsky L. B., Lyakhov A. I. The performance evaluation of distributed information systems of the "Client-server» architecture. Automation and telemechanics.1995. pp. 160-175.
3. Vorobyov S. P., Gorobets V. V. Inzhenernyj vestnik Dona (Rus), 2012. №. 4. URL: ivdon.ru/ru/magazine/archive/n4y2012/1149.
4. Pavlov S. V., Samoilov A. S. Inzhenernyj vestnik Dona (Rus), 2015. №1.URL: ivdon.ru/ru/magazine/archive/n1y2015/2755.

5. Vishnevskiy V.M. Teoreticheskie osnovy proektirovaniya komp'yuternykh setey [Theoretical foundations of computer network design].M: Tekhnosfera, 2003. 512p.
6. Zhozhikashvili V.A., Vishnevskiy V.M. Seti massovogo obsluzhivaniya. Teoriya i primeneniye k setyam EVM [Queueing networks. Theory and its network application].M: Radio i svyaz', 1988. 192p.
7. Gerasimov A.I. Teoriya i prakticheskoye primeneniye stokhasticheskikh setey [Theory and practical application of stochastic networks].M: Radio i svyaz', 1994. 175p.
8. Antunes C. H. et al. A Multiple Objective Routine Algorithm for Integrated Communication Network. Proc ITC-16.1999. V. 3b.pp. 1291-1300.
9. Chakka R., Harrison P. G. A Markov modulated multi-server queue with negative customers-The MM CPP/GE/c/L G-queue. Acta Informatica, 2001, V. 37.pp.785-799.
10. Chernomorov G.A. Teoriya prinyatiya resheniy[decision making theory]: Uchebnoye posobie. Yuzh.-Ros.gos. tekhn. un-t. 3-e izd. pererab. i dop. Novocherkassk: Red. zhurn."Izv.vuzov.Elektromekhanika", 2005.448p.
11. Kruglyy Z.L. Algoritmy rascheta modeley struktur vychislitel'nykh sistem s razlichnymi klassami zadaniy. Upravlyayushchie sistemy i mashiny.1980. №4.pp.73-79.