



Задача определения оптимальной реактивности распределенных систем обработки информации при фиксированной стоимости их простоя

А. Н. Скоба, Айеш Ахмед Нафеа Айеш, В. К. Михайлов

Южно-Российский государственный политехнический университет (НПИ) им. М. И. Платова, Новочеркасск

Аннотация: в данной статье, в терминах языка сетей массового обслуживания, сформулирована задача нахождения оптимального значения среднего времени реакции системы на запросы всех пользователей для распределенных систем обработки информации, реализованных на базе архитектуры «файл-сервер», двухуровневой и трехуровневой архитектуры «клиент-сервер» при фиксированной стоимости простоя, включающей стоимость простоя пользователей, стоимость простоя канала передачи данных и стоимость простоя обслуживающих узлов. Представлен алгоритм решения данной оптимизационной задачи.

Ключевые слова: распределенная система обработки информации, архитектура «файл-сервер», двухуровневая архитектура «клиент-сервер», трехуровневая архитектура «клиент-сервер» пространство состояний, среднее время реакции системы, стоимость потерь, стационарные вероятности.

При проектировании распределенных систем обработки информации (СОИ) на базе различных архитектур их реализации возникает необходимость производить их сравнительный анализ с точки зрения надежности, технико-эксплуатационных и технико-экономических характеристик [1,2]. Так, например, из-за ненадежных компонентов распределенных СОИ происходит ухудшение их технико-эксплуатационных характеристик: пропускная способность каналов передачи данных уменьшается, а время реакции системы на запросы пользователей увеличивается. Кроме того, технико-эксплуатационные характеристики функционирования распределенных СОИ также существенным образом влияют и на технико-экономические показатели их работы, такие как потери из-за неработоспособности отдельных узлов распределенных СОИ, затраты на их вынужденный простой, затраты на их ремонт (восстановление) и т.д. [3,4].

В данной статье представлена математическая модель решения оптимизационной задачи выбора оптимального среднего времени реакции системы на запросы пользователей для различных вариантов архитектурного построения распределенных СОИ при фиксированной стоимости простоя всей системы с использованием аппарата экспоненциальных сетей массового обслуживания (СеМО) [5-7].

Математическая модель решения задачи

Дана распределённая СОИ, реализованная на базе архитектуры «файл-сервер», исходные данные для моделирования которой подробно описаны в работе [8]. В качестве математической модели функционирования данной распределённой СОИ была выбрана экспоненциальная СеМО. Для идентификации состояний случайного процесса описывающего переходы заявок в СеМО было введено векторное пространство состояний:

$$E = \{E_i(i_{11}, \dots, i_{1s}, \dots, i_{1n}; i_{21}, \dots, i_{2s}, \dots, i_{2n}; \dots; i_{s+1,1}, \dots, i_{s+1,n}, \dots, i_{n+1,1}, \dots, i_{n+1,s}, \dots, i_{n+1,n}, \dots, i_{n+2,1}, \dots, i_{n+2,s}, \dots, i_{n+2,n}; \dots; i_{n+s+1,1}, \dots, i_{n+s+1,s}, \dots, i_{n+s+1,n}; \dots; i_{2n+1,1}, \dots, i_{2n+1,s}, \dots, i_{2n+1,n}; \dots, i - \overline{1, p}\}, \text{ где}$$

$$\bar{i}_{1r}^{(1)} = \begin{cases} 1, \text{ если } r\text{-й пользователь } (r = \overline{1, n}) \text{ находится в активном состоянии} \\ \text{(формирует сообщение);} \\ 0, \text{ в противном случае,} \end{cases}$$

$\{\bar{i}_{sr}^{(2)}, s = \overline{2, n+1}, r = \overline{1, n}\}$ – описывает очереди к каналу и состояние канала,

где \bar{i}_{sr} – количество запросов r -го пользователя (тонкого клиента) в s -ой буферной памяти канала и на обслуживании в канале;

$\{i_{sr}^{(3)}, s = \overline{n+2, 2n+1}, r = \overline{1, n}\}$ – описывает очереди к узлам и состояния узлов

(ПЭВМ), где i_{sr} – количество сообщений r -го пользователя в буферной

памяти узла и на обслуживании в узле. При этом имеют место следующие ограничения:

- 1) $\sum_{k=1}^{2n+1} \sum_{r=1}^n i_{kr} = n;$
- 2) $\sum_{k=1}^{2n+1} i_{kr} = 1, r = \overline{1, n};$
- 3) $i_{kr} \geq 0, \forall k = \overline{1, 2n+1}, r = \overline{1, n}.$

В работе [8] было показано, что расчет величины \bar{T} – среднего времени реакции системы на запросы всех пользователей распределенной СОИ может быть сведен к вычислению стационарных вероятностей СеМО $P\{\bar{i}_{1r}^{(1)}, \bar{i}_{sr}^{(2)}, \bar{i}_{sr}^{(3)}\}$, которые имеют мультипликативный вид и являются решением уравнения глобального баланса.

Пусть $\bar{a}_r (r = \overline{1, n})$ – средняя стоимость ожидания ответа на запрос r -го пользователя распределенной СОИ. Тогда среднюю стоимость ожидания ответа всех пользователей распределенной СОИ за время \bar{T} можно определить как

$$\sum_{s=2}^{2n+1} \sum_{r=1}^n P\{\bar{i}_{1r}^{(1)} = \bar{0}, \bar{i}_{sr}^{(2)}, \bar{i}_{sr}^{(3)}\} \cdot \bar{a}_r \bar{T};$$

$\bar{b}_r (r = \overline{1, n})$ – средняя стоимость простоя канала передачи данных распределенной СОИ из-за отсутствия в нем сообщений r -го пользователя. Тогда среднюю стоимость простоя канала передачи данных за время \bar{T} определим как

$$\sum_{s=2}^{2n+1} \sum_{r=1}^n P\{\bar{i}_{1r}^{(1)}, \bar{i}_{sr}^{(2)} = \bar{0}, \bar{i}_{sr}^{(3)}\} \cdot \bar{b}_r \bar{T};$$

$\bar{c}_r (r = \overline{1, n})$ – средняя стоимость простоя r -ой ПЭВМ распределенной СОИ за время \bar{T} , тогда среднюю стоимость простоя всех ПЭВМ за время \bar{T} определим как

$$\sum_{s=2}^{2n+1} \sum_{r=1}^n P\{\bar{i}_{1r}^{(1)}, \bar{i}_{sr}^{(2)}, \bar{i}_{sr}^{(3)} = \bar{0}\} \cdot \bar{c}_r \bar{T}.$$

Тогда суммарные затраты от простоя всех составляющих компонентов распределенной СОИ, реализованной на базе архитектуры «файл-сервер» за время \bar{T} равны

$$C = \sum_{s=2}^{2n+1} \sum_{r=1}^n P\{\bar{i}_{1r}^{(1)} = \bar{0}, \bar{i}_{sr}^{(2)}, \bar{i}_{sr}^{(3)}\} \cdot \bar{a}_r \bar{T} + \\ + \sum_{s=2}^{2n+1} \sum_{r=1}^n P\{\bar{i}_{1r}^{(1)}, \bar{i}_{sr}^{(2)} = \bar{0}, \bar{i}_{sr}^{(3)}\} \cdot \bar{b}_r \bar{T} + \\ + \sum_{s=2}^{2n+1} \sum_{r=1}^n P\{\bar{i}_{1r}^{(1)}, \bar{i}_{sr}^{(2)}, \bar{i}_{sr}^{(3)} = \bar{0}\} \cdot \bar{c}_r \bar{T}.$$

И тогда задача определения оптимального среднего времени реакции системы на все запросы пользователей распределенной СОИ при фиксированной стоимости потерь от простоя ее компонентов будет иметь вид:

$$\left. \begin{aligned} \bar{T} &\rightarrow opt, \\ C &\leq C^*. \end{aligned} \right\} \quad (1)$$

где C^* – величина фиксированной стоимости потерь от простоя компонентов распределенной СОИ.

Исходные данные для моделирования работы распределённой СОИ, реализованной на базе двухуровневой архитектуры «клиент-сервер» также приведены в работе [8]. Для идентификации состояний случайного процесса, происходящего в СеМО было введено векторное пространство состояний :

$$E = \{E_i(i_{01}, \dots, i_{0s}, \dots, i_{0n}; i_{11}, \dots, i_{1s}, \dots, i_{1n}; \dots; i_{2n+1,1}, \dots, i_{2n+1,s}, \dots, i_{2n+1,n}), \mathbf{i} = \overline{1, p}\}, \text{ где}$$

$\{i_{0s}, s = \overline{1, n}\}$ описывают состояние буфера s -го пользователя, при этом i_{0s} –

количество обслуженных транзакций s -го пользователя; $\{i_{1s}^{(1)}, s = \overline{1, n}\}$

описывают состояние s -го пользователя, причем

$$\bar{i}_{1s}^{(1)} = \begin{cases} d, \text{ если } s\text{-й пользователь находится в активном состоянии} \\ \text{(формирует сообщение);} \\ 0 \leq i_{1s} \leq d, s\text{-й пользователь находится в пассивном состоянии} \\ \text{(ждет ответа на запрос),} \end{cases}$$

$\{\bar{i}_{sr}^{(2)}, s = \overline{2, n+1}, r = \overline{1, n}\}$ – описывает очереди к каналу и состояние канала,

где \bar{i}_{sr} – количество запросов r -го пользователя в s -ой буферной памяти канала и на обслуживании в канале;

$\{\bar{i}_{sr}^{(3)}, s = \overline{n+2, 2n+1}, r = \overline{1, n}\}$ – описывает очереди к узлам и состояния узлов

(ПЭВМ), где i_{sr} – количество сообщений r -го пользователя в буферной памяти узла и на обслуживании в узле. При этом имеют место следующие ограничения:

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

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

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

И тогда суммарные затраты от простоя узлов распределенной СОИ, реализованной на базе двухуровневой архитектуры «клиент-сервер» можно представить как

$$C = \sum_{s=2}^{2n+1} \sum_{r=1}^n P\{\bar{i}_{1r}^{(1)} < d, \bar{i}_{sr}^{(2)}, \bar{i}_{sr}^{(3)}\} \cdot \bar{a}_r \bar{T} + \\ + \sum_{s=2}^{2n+1} \sum_{r=1}^n P\{\bar{i}_{1r}^{(1)}, \bar{i}_{sr}^{(2)} < d, \bar{i}_{sr}^{(3)}\} \cdot \bar{b}_r \bar{T} + \\ + \sum_{s=2}^{2n+1} \sum_{r=1}^n P\{\bar{i}_{1r}^{(1)}, \bar{i}_{sr}^{(2)}, \bar{i}_{sr}^{(3)} < d\} \cdot \bar{c}_r \bar{T}.$$

Для модели функционирования распределённой СОИ на базе трёхуровневой архитектуры «клиент-сервер» исходные данные приведены в

работе [9]. Для идентификации состояний случайного процесса, описываемого экспоненциальной СеМО, было введено векторное пространство состояний \bar{E} , которое согласно [9] имеет следующий вид:

$$\bar{E} = \{E_i(i_{11}, \dots, i_{1s}, \dots, i_{1n}; i_{21}, \dots, i_{2s}, \dots, i_{2n}; \dots; i_{s+1,1}, \dots, i_{s+1,n}, \dots, i_{n+1,1}, \dots, i_{n+1,s}, \dots, i_{n+1,n}, \dots, i_{n+2,1}, \dots, i_{n+2,s}, \dots, i_{n+2,n}; \dots; i_{n+s+1,1}, \dots, i_{n+s+1,s}, \dots, i_{n+s+1,n}; \dots; i_{2n+1,1}, \dots, i_{2n+1,s}, \dots, i_{2n+1,n}; i_{2n+2,1}, \dots, i_{2n+2,s}, \dots, i_{2n+2,n}; \dots; i_{2n+s+1,1}, \dots, i_{2n+s+1,s}, \dots, i_{2n+s+1,n}; \dots; i_{3n+1,1}, \dots, i_{3n+1,s}, \dots, i_{3n+1,n})\}, \bar{i} = \overline{1, s}, \text{ где}$$

$$\bar{i}_{1s}^{(1)} = \begin{cases} d, \text{ если } s\text{-й пользователь (тонкий клиент, } s = \overline{1, n}) \text{ находится в активном} \\ \text{состоянии (формирует запрос на инициализацию бизнес-приложения);} \\ 0, \text{ в противном случае,} \end{cases}$$

$\{\bar{i}_{sr}^{(2)}, s = \overline{2, n+1}, r = \overline{1, n}\}$ – описывает очереди к каналу и состояние канала,

где \bar{i}_{sr} – количество запросов r -го пользователя (тонкого клиента) в s -ой буферной памяти канала и на обслуживании в канале;

$\{i_{sr}^{(3)}, s = \overline{n+2, 2n+1}, r = \overline{1, n}\}$ – описывает очереди запросов к серверам приложений и состояния серверов приложений, где i_{sr} – количество

запросов r -го пользователя в буферной памяти s -го сервера приложений и на обслуживании в s -ом сервере приложений;

$\{i_{sr}^{(4)}, s = \overline{2n+2, 3n+1}, r = \overline{1, n}\}$ – описывает очереди к серверам баз данных и

состояние серверов баз данных, где i_{sr} – количество сообщений r -го пользователя в буферной памяти s -го сервера баз данных и на обслуживании в s -ом сервере баз данных.

При этом имеют место следующие ограничения:

$$1) \sum_{k=1}^{3n+1} \sum_{r=1}^n i_{kr} = n \cdot r \cdot d;$$

$$2) \sum_{k=1}^{3n+1} i_{kr} = r \cdot d.$$

Введем в рассмотрение величину $\bar{d}_r (r = \overline{1, n})$ – среднюю стоимость простоя r -го сервера-приложений распределенной СОИ, тогда средняя стоимость простоя всех серверов-приложений распределенной СОИ за время

$$\bar{T} \text{ равна } \sum_{s=2}^{3n+1} \sum_{r=1}^n \{\bar{i}_{1r}^{(1)}, \bar{i}_{sr}^{(2)}, \bar{i}_{sr}^{(3)} < d, \bar{i}_{sr}^{(4)}\} \cdot \bar{d}_r \bar{T},$$

а суммарные потери за время \bar{T} будут равны

$$\begin{aligned} C = & \sum_{s=2}^{3n+1} \sum_{r=1}^n \{\bar{i}_{1r}^{(1)} < d, \bar{i}_{sr}^{(2)}, \bar{i}_{sr}^{(3)}, \bar{i}_{sr}^{(4)}\} \cdot \bar{a}_r \bar{T} + \\ & + \sum_{s=2}^{3n+1} \sum_{r=1}^n \{\bar{i}_{1r}^{(1)}, \bar{i}_{sr}^{(2)} < d, \bar{i}_{sr}^{(3)}, \bar{i}_{sr}^{(4)}\} \cdot \bar{b}_r \bar{T} + \\ & + \sum_{s=2}^{3n+1} \sum_{r=1}^n \{\bar{i}_{1r}^{(1)}, \bar{i}_{sr}^{(2)}, \bar{i}_{sr}^{(3)} < d, \bar{i}_{sr}^{(4)}\} \cdot \bar{d}_r \bar{T} \\ & + \sum_{s=2}^{3n+1} \sum_{r=1}^n \{\bar{i}_{1r}^{(1)}, \bar{i}_{sr}^{(2)}, \bar{i}_{sr}^{(3)}, \bar{i}_{sr}^{(4)} < d\} \cdot \bar{c}_r \bar{T}. \end{aligned}$$

Так как $\bar{T} = \bar{T}(X)$, где X – матрица, задающая размещение отношений $R = \{R_1, \dots, R_j, \dots, R_d\}$ по узлам $U = \{U_1, \dots, U_s, \dots, U_n\}$ распределенной СОИ, то задача (1) является задачей комбинаторной оптимизации и на концептуальном уровне сводится к задаче оптимального размещения информационных ресурсов коллективного пользования по её узлам [8,9].

Алгоритм решения задачи

1. Положить $k = 0$, задать начальное распределение информационных ресурсов (отношений R) по узлам U распределенной СОИ выбранной архитектуры – матрицу $X^{(k)}$.

2. С использованием разработанного пакета [10] найти распределение стационарных вероятностей $P\{\bar{i}_{sr}\}$ и среднее время реакции системы на запросы пользователей – $\bar{T}(X^{(k)})$.

3. Вычислить величину C – суммарные затраты от простоя узлов распределенной СОИ при заданном значении $\bar{T}(X^{(k)})$.

4. Если условие $C \leq C^*$ (C^* – заданная величина) выполняется, то полученное значение $\bar{T}^{(k)}$ – искомое решение задачи вида (1). В противном случае, полагаем $k = k + 1$ и с использованием встроенного алгоритма оптимизации в программный комплекс [10] получаем новое значение $\bar{T}^{(k)}$ и, соответственно, новые значения $P\{\bar{i}_{sr}\}$. Осуществляем переход на п. 3.

5. Если за приемлемое число шагов k , решение задачи (1) найдено не будет, то необходимо провести корректировку значения C^* .

Вывод. Разработанные модели решения задачи определения оптимальной реактивности распределённых СОИ при фиксированной стоимости простоя их компонент, реализованные на базе различных архитектур их построения (файл-серверной, двухуровневой клиент-серверной, трёхуровневой клиент-серверной), могут быть использованы для получения наилучших технико-эксплуатационных и технико-экономических характеристик при внедрении интегрированных информационно-справочных систем на предприятиях для планировании более рациональной организации вычислительного процесса.

Литература

1. Клепов А.В., Гузенко В.Л., Миронов Е.А. Постановка задачи обеспечения технико-экономической эффективности распределенной информационной системы // Современные проблемы науки и образования. – 2014. – №2. URL: science-education.ru/ru/article/view?id=12361

2. Zibin Zhend, M.R.Lyu, WS-DREAM: A distributed reliability assessment mechanism for web services, in: IEEE International Conference on Dependable

Systems and Networks with FTCS and SCC, DSN 2008. 24-27 June 2008, pp.392-397.

3. Авен О.И., Гурин Н.И., Котан Я.А. Оценка качества и оптимизации вычислительных систем. – М.: Наука, 1982. – 464 с.
4. Кузнецов Н.А., Кульба В.В., Косяченко С.А. Методы анализа и синтеза модульных информационно-управляющих систем. -М.: ФИЗМАТЛИТ, 2002. -880с.
5. Вишневский В.М. Теоретические основы проектирования компьютерных сетей.- М.: Техносфера, 2003.- 512 с.
6. Жожикашвили В.А., Вишневский 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. Скоба А.Н., Состина Е.В. Математическая модель оптимального размещения распределенной базы данных по узлам ЛВС на базе двухуровневой клиент-серверной архитектуры. // Инженерный вестник Дона. 2015. № 2. URL: ivdon.ru/ru/magazine/archive/n2y2015/2882.
9. Скоба А. Н., Айеш Ахмед Нафеа Айеш. Математическая модель функционирования распределённой информационной системы на базе трёхуровневой клиент-серверной архитектуры // Инженерный вестник Дона.2017. №4. URL:ivdon.ru/ru/magazine/archive/n4y2017/4482.
10. Михайлов В.К., Скоба А.Н., Айеш Ахмед Нафеа Айеш. Программный комплекс для решения задач оптимального размещения информационных ресурсов и моделирования влияния основных интегральных показателей на реактивность распределенных систем обработки информации. // Изв. вузов. Электромеханика. 2018. Т. 61. №6. С.102-109.

References

1. Klepov A.V., Guzenko V.L., Mironov E.A. Sovremennye problemy nauki i obrazovaniya, 2014. №2. URL: science-education.ru/ru/article/view?id=12361
 2. Zibin Zhend, M.R.Lyu, IEEE International Conference on Dependable Systems and Networks with FTCS and SCC, DSN 2008. 24-27 June 2008, pp.392-397.
 3. Aven O.I., Gurin N.I., Kotan Ya.A. Ocenka kachestva i optimizacii vy`chislitel`ny`x sistem [Evaluation of the quality and optimization of computing systems]. M.: Nauka, 1982. 464p.
 4. Kuznecov N.A., Kul`ba V.V., Kosyachenko S.A. Metody analiza i sinteza modul'nyh informacionno-upravlyayushchih sistem [Methods of analysis and synthesis of modular information and control systems]. M.: FIZMATLIT, 2002. 880 p.
 5. Vishnevskij V.M. Teoreticheskie osnovy` proektirovaniya komp`yuterny`x setej [Theoretical foundations of computer network design]. M.: Texnosfera, 2003. 512p.
 6. Zhozhikashvili V.A., Vishnevskij V.M. Seti massovogo obsluzhivaniya. Teoriya i primeneniye k setyam E`VM [Queueing networks. Theory and its network application]. M.: Radio i svyaz', 1988. 192p.
 7. Gerasimov A.I. Teoriya i prakticheskoye primeneniye stoxasticheskix setej.- M.: Radio i svyaz`. [Theory and practical application of stochastic networks]. 1994. 175p.
 8. Skoba A.N., Sostina E.V. Inzenernyj vestnik Dona, 2015. №2. URL:ivdon.ru/ru/magazine/archive/n2y2015/2882.
 9. Skoba A.N., Ayesh Achmed Nafea Ayesh. Inzenernyj vestnik Dona, 2017. №4. URL:ivdon.ru/ru/magazine/archive/n4y2017/4482.
 10. Mikhaylov V.K., Skoba A.N., Ayesh Achmed Nafea Ayesh. Izv. vuzov. E`lektromexanika, 2018. T. 61. №6. pp. 102-109.
-

