Инженерный вестник Дона, №4 (2016)
ivdon.ru/ru/magazine/archive/n4y2016/3784

Построение аналитико-численной модели распределенных информационных систем с высоким уровнем сетевого трафика

В.Н.Ковалевский

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

Аннотация: в данной статье на основе теории систем массового обслуживания (СМО) разработана аналитико-численная модель для распределенной информационной системы (ИС)в условиях высокого уровня сетевого трафика. Приведены выражения целевой функции для задачи многопараметрической оптимизации передаваемых пакетов, формулы расчета характеристик СМО на основе вектора стационарного распределения вероятностей (СРВ).
Ключевые слова: распределённая информационная система, сетевой трафик, адаптационный протокол, система массового обслуживания, марковский процесс, целевая функция, многопараметрическая оптимизация,пространство состояний, инфинитезимальная матрица, вектор стационарного распределения вероятностей.

Одним из распространенных способов анализа информационных процессов в вычислительных сетях (ВС) распределенных информационных систем [1-3] является построение вероятностных моделей в виде *систем и сетей массового обслуживания*. При разработке таких моделей применяется понятие марковского процесса и такие методы как вложенных цепей Маркова, экспоненциальных сетей [4-6] и т.д.

Рассматриваемая аналитико-численная модель описывает функционирование вычислительной сети распределенных ИС с высоким, изменяющимся сетевым трафиком в линиях связи. Выбор оптимальных длин пакетов (адаптация) для всех линий связи (ЛС) осуществляется с учетом диапазона варьирования этого трафика. Данная модель поддерживает схему раздельных буферных накопителей в центральной ЭВМ (сервере БД) и дисциплину обслуживания в порядке очереди. Протокол включает на транспортном уровне обмена адаптацию длины прямых сообщений. В линиях связи реализуется стартстопный тип передачи и используется адаптивно-настраиваемая длина прямых сообщений и фиксированная длина ответных сообщений. Предлагаемая модель имеет следующие особенности.

Во-первых, процесс передачи прямых сообщений при адаптационном протоколе обмена сопровождается их разбиением на группы пакетов и самостоятельной транспортировкой каждого из них по линии связи. Эти пакеты анализируются на достоверность в центральном узле так же, как и целое прямое сообщение.

Во-вторых, процессы разбиения приводят к резкому увеличению количества блоков сообщений в сети. Эта особенность обуславливает возникновение перед обслуживающим устройством очереди ожидающих проверки пакетов от различных узлов сети.

В-третьих, наличие этой очереди усиливает влияние независимо функционирующих линий связи друг на друга. Это влияние проявляется на одном из уровней работы устройства, связанном с анализом достоверности пакетов. Данный уровень имеет абсолютный приоритет по отношению к другому уровню, предоставляющему обслуживание прямого сообщения.

В предлагаемой модели анализ пакетов в устройстве описывается самостоятельным процессом, т.к. в противном случае это приведет к большим погрешностям результатов моделирования. В качестве исходных данных для построения модели использованы следующие параметры и законы распределения[7,8]:$\boldsymbol{m}$,$\boldsymbol{L_l^{ps}}$,$\boldsymbol{L_l^{os}}$, $\boldsymbol{C_l}$, $\overline{\boldsymbol{t_{po_l}^{ps}}}$,$\overline{\boldsymbol{t_{pk_l}^{ps}}}$, $\overline{\boldsymbol{t_{po_l}^{os}}}$, $\boldsymbol{W_{wneh_l}}$, $\boldsymbol{W_{wnt_l}}$, $\boldsymbol{F_l^{ps}(t)}$, $\boldsymbol{P_l^b}$,$\boldsymbol{F_l^{os}(t)}$,$\boldsymbol{B_l\ (t)}$,$\boldsymbol{F_l^n(t)}$, а также интервал среднего количества битовых ошибок передачи для различных уровней сетевого трафика$\left[\boldsymbol{\alpha_l^{min}}, \boldsymbol{\alpha_l^{max}}\right]$. (2.1)

Дополнительно введены следующие исходные данные: длина квитанции $\boldsymbol{L_l^{kv}}$ и пакета $\boldsymbol{L_l^{pak}}$, используемые в l -м канале. Параметры $\boldsymbol{P_l^b}$, $\boldsymbol{L_l^{pak}}$могут принимать различные текущие значения, т.е. являются непостоянными. Параметр $\boldsymbol{P_l^b}$ определяется, как$\boldsymbol{P_l^b = 1 - e^{-\alpha_l/c_l}}$. Также в модели используется: закон распределения времени передачи пакета в l -м канале $\boldsymbol{F_l^{pak}(t) = 1 - e^{-\eta_l^{pak} t}}$, где$\boldsymbol{\eta_l^{pak}} = \boldsymbol{C_l/L_l^{pak}}$ –интенсивность передачи пакета в l -м канале; закон распределения времени анализа достоверности пакетаl -го канала $\boldsymbol{B_l^{pak}(t) = 1 - e^{-\mu_l^{pak} t}}$,где $\boldsymbol{\mu_l^{pak}} = \boldsymbol{k_l \cdot L_l^{pak}}$ – интенсивность процесса анализа пакета l -го канала, $\boldsymbol{k_l}$ – некоторый линейный коэффициент зависимости параметров $\boldsymbol{\mu_l^{pak}}$ и $\boldsymbol{L_l^{pak}}$.

Среднее время занятости l-го канала при передаче прямого сообщения $\boldsymbol{\overline{t_{zan_l}^{ps}}}$, будет определяться следующим образом:

$$\overline{t_{zan_l}^{ps}} = \overline{t_{po_l}^{ps}} + \overline{T_{otk_l}^*}, \qquad (2.2)$$

где$\boldsymbol{\overline{T_{otk_l}^*}}$ - минимальное среднее время отклика $\boldsymbol{\overline{T_{otk_l}}}$при передаче прямого сообщения пакетом, длина $\boldsymbol{L_l^{pak}}$которого является оптимальной для текущего значения битовой ошибки $\boldsymbol{P_l^b}$ в l -м канале с учетом значений битовых ошибок и используемых длин пакетов в других каналах. Перед определением $\boldsymbol{\overline{T_{otk_l}^*}}$ необходимо: построить целевую функцию для решения задачи многопараметрической оптимизации для пакетов в линиях связи $\boldsymbol{L_l^{pak}(l = \overline{1,m})}$, определить место поиска экстремума этой функции и провести отыскание этого экстремума с целью получения вектора оптимальных длин пакетов $\boldsymbol{L_{opt_l}^{pak}}$. Данная оптимизационная задача относится к многокритериальным сложным задачам. Поэтому при ее решении необходимо воспользоваться методом свертки, позволяющим упростить процесс поиска экстремума за счет рассмотрения целевой функции в виде суммы критериев $\overline{\boldsymbol{T}_{\boldsymbol{otk_l}}}, (\boldsymbol{l} = \overline{\boldsymbol{1, m}})$. Выбор метода свертки обусловлен вышеуказанными особенностями организации процесса передачи сообщений при использовании адаптации в моделируемой ВС и, в частности, влиянием линий связи друг на друга. Поэтому характеристика $\overline{\boldsymbol{T}_{\boldsymbol{otk_l}}}$ определяется, как

$$\overline{T_{otk_l}} = \left[\left(\frac{L_l^{pak}}{C_l} + \overline{t_{og}^{pak}} + \frac{1}{\mu_l^{pak}} + \frac{L_l^{kv}}{C_l}\right) \cdot \left(M_l^{pak}(\varphi) + 1\right)\right] \cdot \theta_l. \quad (2.3)$$

Тогда целевая функция выглядит следующим образом:

$$\sum_{l=1}^{m} \overline{T_{otk_l}}(L_i^{pak}) = \sum_{l=1}^{m} \left[\left(\frac{L_l^{pak}}{C_l} + \overline{t_{og}^{pak}} + \frac{1}{\mu_l^{pak}} + \frac{L_l^{kv}}{C_l}\right) \cdot \left(M_l^{pak}(\varphi) + 1\right)\right] \cdot \theta_l, \quad (2.4)$$

где $\overline{\boldsymbol{t}_{\boldsymbol{og}}^{\boldsymbol{pak}}}$ — среднее время ожидания пакетов всех каналов в очереди при анализе в устройстве; $\boldsymbol{M}_{\boldsymbol{l}}^{\boldsymbol{pak}}(\boldsymbol{\varphi})$ - математическое ожидание количества неуспешных передач $\boldsymbol{\varphi}$ одного пакета длиной $\boldsymbol{L}_{\boldsymbol{l}}^{\boldsymbol{pak}}$ в l -м канале; $\boldsymbol{\theta}_{\boldsymbol{l}}$- количество пакетов длиной $\boldsymbol{L}_{\boldsymbol{l}}^{\boldsymbol{pak}}$, необходимое для передачи прямого сообщения длиной $\boldsymbol{L}_{\boldsymbol{l}}^{\boldsymbol{ps}}$.Математическое ожидание количества неуспешных передач $\boldsymbol{\varphi}$ одного пакета в l -м канале определяется выражением

$$M_l^{pak}(\varphi) = \frac{P_{og_l}^{''\prime}}{1 - P_{og_l}^{''\prime}},$$

где $P_{og_l}^{''\prime} = 1 - \left(1 - P_l^b\right)^{L_l}$ - вероятность искажения одного пакета в l -м канале; $\boldsymbol{P}_{\boldsymbol{l}}^{\boldsymbol{b}}$ - вероятность битовой ошибки в l -м канале. Количество пакетов, используемых для передачи одного прямого сообщения в l -м канале, вычисляется, как $\theta_l = L_l^{ps}/L_l^{pak}$. Для получения характеристики $\boldsymbol{t_{og}^{pak}}$ необходимо рассматривать процесс анализа пакетов на достоверность в устройстве независимо от процесса обслуживания. Правомерность этого допущения обуславливается абсолютным приоритетом, который имеет данный процесс. Учитывая вышесказанное, а также то, что в каждом канале выполняется стартстопная передача пакета, т.е. пакет выдается в канал по получению квитанций на ранее посланный пакет, этот процесс в изучаемой вычислительной сети можно концептуально представить в виде замкнутой системы массового обслуживания. СМО включает $\boldsymbol{m}$ источников без накопителей и один - обслуживающий прибор. Источниками являются каналы, интенсивность $\boldsymbol{\eta_l^{pak}}$ которых обратно пропорциональна длине пакета $\boldsymbol{L_l^{pak}}$. Обслуживающий прибор работает с интенсивностью $\boldsymbol{\mu_l^{pak}}$, где$(\boldsymbol{\mu_l^{pak}})^{-1}$- среднее время анализа пакета l -го канала. Ввиду заданных экспоненциальных законов распределения времени передачи пакета $\boldsymbol{F_l^{pak}(t)}$и времени его анализа в устройстве $\boldsymbol{B_l^{pak}(t)}$, рассматриваемая СМО относится к системе с разнородными источниками. Для нахождения ее характеристик, и в частности среднего времени ожидания пакетов в очереди на обслуживании $\boldsymbol{t_{og}^{pak}}$, необходимо разработать аналитико-численную модель на базе метода экспоненциальных сетей. Конструирование этой модели включает формирование пространства состояний СМО $E = \{E_j, i = \overline{1, \rho_o}\}$, вычисление элементов инфинитезимальной матрицы $\boldsymbol{Q_{E_iE_j}}$.Пространство$\boldsymbol{E}$ рассматриваемой системы массового обслуживания имеет вид $E = \{E_i(i_{1,} \ldots\ldots, i_{l,} \ldots\ldots, i_m)$, $i = \overline{1, \rho_o},\ \rho_o = 2^m\}$,где $\boldsymbol{i_{l,}}$- состояниеl -го источника (0- пассивное, 1 -

активное). Элементы матрицы $\boldsymbol{Q}_{E_iE_j}$, формальное описание условий и интенсивностей перехода системы из состояния $\boldsymbol{E_i}$, соответствующее моменту времени $\boldsymbol{t}$, в состояние $\boldsymbol{E_j}$ для момента времени $\boldsymbol{t}+\Delta\boldsymbol{t}$ определяются следующим образом:

$$q_{E_iE_j} = \begin{cases} \eta_f^{pak}, & \exists \mathrm{f}: i_f = 1\,, j_f = 0;\ \ \forall_{s\neq f}^{s}: i_s = j_s\,, s = \overline{1, m-1}\,; \\ \mu_f^{pak} \cdot \dfrac{\eta_f^{pak}}{\sum_{i=1}^{m} \eta_f^{pak}}, & \exists f: i_f = 0\,, j_f = 1; \forall_{s\neq f}^{s}: i_s = j_s\,, s = \overline{1, m-1}\,; \\ -\sum\limits_{\substack{r=1 \\ r\neq i}}^{p} q_{ir}\,, & \forall s: i_f = j_s,\quad s = \overline{1, m}\,; \\ 0, & \text{в остальных случаях.} \end{cases}$$

Далее необходимо последовательно рассчитать[8]: вектор стационарного распределения вероятностей $\pi = \left(\pi_{1,}\,\ldots\ldots, \pi_{i,}\,\ldots\ldots, \pi_p\right)$;закон распределения состояний *l*-го источника $P_{l,z}(z = \overline{0{,}1})$;среднее количество пакетов в системе от *l*-го источника $\overline{\omega_l}$;среднюю интенсивность *l*-го источника $\eta_{sr\ \ l}^{pak} = \eta_l^{pak} \cdot \left(1 - P_{l,o}\right)$;среднее время нахождения пакета *l*-го источника в системе $\overline{t_{sis\ \ l}^{pak}} = \overline{\omega_l}/\eta_{sr\ \ l}^{pak}$. Среднее время ожидания пакета *l*-го источника в очереди определяется, как $\overline{t_{og_i}^{pak}} = \overline{t_{sis_l}^{pak}} - 1/\mu_l^{pak}$.

Однако приведенная модель не позволяет получить окончательное аналитическое выражение для вычисления характеристики $\overline{\boldsymbol{t_{og_i}^{pak}}}$ и, следовательно, выразить ее через вектор длин пакетов L_1^{pak} ,L_2^{pak} ,...,

L_m^{pak}.Это условие необходимо для построения целевой функции (2.4). Данное аналитическое выражение можно найти на базе аналитической модели СМО с однородными источниками. Но для ее использования необходимо ввести следующие допущения:

$$\eta_{sr}^{pak} = \sum_{l=1}^{m} \eta_l^{pak} / m \, , \mu_{sr}^{pak} = \sum_{l=1}^{m} \mu_l^{pak} / m \, .$$

Тогда среднее количество пакетов в системе на обслуживаниеопределяется, как$\overline{\omega}$ = m – (1 - P_o)/ρ_{pak} ,

где $\rho_{pak} = \eta_{sr}^{pak} / \mu_{sr}^{pak}$ -коэффициент загрузкисистемы; $P_o = \left(1 + m\rho_{pak} + m(m-1)\rho_{pak}^2 + \ \dots + m!\,\rho_{pak}^m\right)^{-1}$ – вероятность того, что устройство свободно. Среднее время нахождения пакетов системе вычисляется следующей образом:$\overline{t_{sis}^{pak}} = \overline{\omega} / \eta_{sr}^{pak} \ (m - \overline{\omega})$.Среднее время ожидания в очереди находится как

$\overline{t_{og}^{pak}} = \overline{\omega} / \eta_{sr}^{pak} \ (m - \overline{\omega})$ – 1/μ_{sr}^{pak}.

Для нахождения экстремума целевой функции применение прямого метода его отыскания через систему уравнений$\partial\left(\sum_{l=1}^{m} \overline{T_{otk_l}} \left(L_l^{pak}\right) / \partial L_l^{pak}\right)$ = 0,(l = $\overline{1,m}$) затруднено ввиду сложности функции (2.4). Поэтому для решения данной задачи необходимо использовать один из численных итерационных методов поиска безусловного экстремума - метод сопряженных направлений Дэвидона–Флетчера–Пауэлла [9]. Этот метод скомпонован программно в виде библиотечного модуля, который был успешно также включен в программную реализацию модели, представленной в работе [10]. Таким образом, при заданных значениях вероятностей битовых ошибок $\boldsymbol{P_l^b}$ (l = $\overline{1,m}$)можно получить вектор оптимальных длин пакетов$L_{opt_1}^{pak}, L_{opt_2}^{pak}, \ldots, L_{opt_m}^{pak}$, на основе которого по выражениям (2.2), (2.3) легко рассчитать соответственно характеристики $\overline{t_{zan_l}^{ps}}$ и $\overline{T_{otk_l}^{*}}$.

Литература

1.Таненбаум Э. Распределенные системы. Принципы и парадигмы / Э. Таненбаум, М. ван Стеен. – СПб.: Питер, 2003. – 877 с.

2. Wiesmann M., Pedone F., Schiper A., Kemme B., Alonso G. Database Replication Techniques: a Three Parameter Classification // Proc. 19-th {IEEE} Symp. on Reliable Distributed Systems. 2000. pp. 206–218.

3. Holliday J., Steinke R.., Agrawal D., Amr E. A. Epidemic Algorithms for Replicated Databases // IEEE Transactions on Knowledge and Data Engineering. 2003. Vol. 15, N. 3. pp. 1218–1238.

4. 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.

5. Матвеев В.Ф., Ушаков В.Г. Системы массового обслуживания. – М.: Изд-во МГУ, 1984. – 240 с.

6. Скоба А.Н., Состина Е.В. Математическая модель оптимального размещения распределённой базы данных по узлам ЛВС на базе двухуровневой клиент-серверной архитектуры // Инженерный вестник Дона, 2015, №2 URL:ivdon.ru/ru/magazine/archive/n2y2015/2882.

7.Ковалевский В.Н. Аналитико-численное моделирование распределенных информационных систем с низким уровнем сетевого трафика // Инженерный вестник Дона, 2015, № 3URL:ivdon.ru/ru/magazine/archive/n3y2015/3174.

8. Ковалевский В.Н., Воробьёв С.П. Построение аналитико-численных моделей распределенных информационных систем с невысоким уровнем

сетевого трафика // Изв. вузов. Сев.- Кавк. регион. Техн. науки. 2015. № 2. С. 23-29.

9. Базара М., Шетти К. Нелинейное программирование. Теория иалгоритмы. –М: Мир, 1982.– 583 с.

10.Зуев В.А., Ковалевский В.Н., Черноморов Г.А. Программное моделирование систем: учеб. пособие/ Новочерк. политехн. ин-т. - Новочеркасск, 1992. – 109 с.

References

1. Tanenbaum Je. Raspredelennye sistemy. Principy i paradigmy [The distributed systems. Principles and paradigms]. SPb.:Piter, 2003. 877 p.

2. Wiesmann M., Pedone F., Schiper A., Kemme B., Alonso G. Database Replication Techniques: a Three Parameter Classification. Proc. 19-th {IEEE} Symp. on Reliable Distributed Systems. 2000. pp. 206–218.

3. Holliday J., Steinke R.., Agrawal D., Amr E. A. Epidemic Algorithms for Replicated Databases. IEEE Transactions on Knowledge and Data Engineering. 2003. Vol. 15, N. 3. pp. 1218–1238.

4. Chernomorov G.A. Teorija prinjatija reshenij: Uchebnoe posobie [Theory of decision-making: Manual]. Juzh. -Ros.gos. tehn.un-t.-3-e izd. pererab. i dop. Novocherkassk: Red. zhurn. «Izv. Vuzov. Jelektroomehanika», 2005. 448p.

5.Matveev V.F., Ushakov V.G. Sistemy massovogo obsluzhivanija. [Systems of mass service]. M.: Izd-vo MGU, 1984. 240 p.

6. Ckoba A.N., Sostina E.V. Inženernyj vestnik Dona (Rus), 2015, №2 URL:ivdon.ru/ru/magazine/archive/n2y2015/2881.

7. Kovalevskij V.N. Inženernyj vestnik Dona (Rus), 2015, №3 URL:ivdon.ru/ru/magazine/archive/ n3y2015/3174.

8. Kovalevskij V.N., Vorob'jov S.P. Izv. vuzov. Sev.-Kavk. region. Tehn. nauki. 2015. № 2. pp. 23–29.

9. Bazara M., SHetti K. Nelinejnoe programmirovanie. Teoriya i algoritmy. [Nonlinear Programming. Theory and Algorithms].M: Mir, 1982. 583p.

10. Zuev V.A., Kovalevskij V.N., Chernomorov G.A. Programmnoe modelirovanie sistem: ucheb. posobie [Program modeling of systems: ucheb.posoby]. Novocherk. politehn. in-t. Novocherkassk, 1992. 109 p.