УДК 004.7

МАТЕМАТИЧЕСКОЕ ОБЕСПЕЧЕНИЕ В КОМПЬЮТЕРНЫХ СЕТЯХ И СИСТЕМАХ: МОДЕЛИ ОБРАБОТКИ ИНФОРМАЦИОННЫХ ПРОЦЕССОВ

Исаева Гульнара Бостановна
кандидат педагогических наук
Каптагай Гулбану Алибековна
доктор философских наук
Смагулова Лаура Амангельдиевна
кандидат педагогических наук
Катаев Назбек Сейсенбайевич
кандидат педагогических наук
Ермаков Анатолий Семенович
кандидат технических наук
Мишина Айгерим Ержановна
магистр
Байдильдинов Дамир Талгатович
магистр
Каспийский общественный университет
Казахский национальный исследовательский технический университет
им. К.И. Сатпаева, Алматы (Казахстан)

Аннотация. Математические методы теории массового обслуживания позволяют рассчитать показатели производительности различных ресурсов вычислительных систем и сетей. В основном это задачи оценки вероятностно-временных характеристик функционирования узлов. Традиционный подход к решению задач анализа производительности вычислительной сети заключается в решении уравнений баланса системы: равновесия потоков относительно их интенсивностей в модели для ее декомпозиции на отдельные узлы и в нахождении характеристик узлов и всей сети в целом. В статье приведен сравнительный анализ нескольких моделей на примере трехфазных систем массового обслуживания, а также значения пропускной способности для каждой из описанных моделей.

Ключевые слова: модели обработки процессов, производительность, теория массового обслуживания, цепи Маркова, пропускная способность системы, коэффициент загрузки.

Для анализа производительности вычислительных систем и сетей может быть проведено аналитическоевероятностноемоделирование-наосноветеориимассовогообслуживания. Расчет характеристик систем и сетей при динамическом распределении нагрузки в системах массового обслуживания с большим числом узлов представляет собой сложную математическую задачу. Аналитическое вероятностное моделирование позволяет получить расчетную схему для определенного числа узлов; при возрастании количества узлов используются приближенные методы, например, итерационный метод [1].

С целью проведения оценки систем обработки и сетей передачи информации взяты трехфазные системы массового обслуживания, отражающие взаимодействие процессов ввода, обработки и вывода информационных потоков. Рассмотрим несколько вариантов.

Трехфазная модель без буфера. В такой модели предполагается, что только перед первой фазой располагается бесконечный буфер. Данная модель рассмотрена в [2] для частного случая, когда интенсивность обслуживания в каждой фазе одинакова $\mu_1 = \mu_2 = \mu_3 = 1$. В такой модели стационарное состояние достигается в случае, когда интенсивность поступления запросов на фазу меньше интенсивности обслуживания в фазе.

Для анализа модели на основе предположений о законах распределения случайных величин определены вероятности состояний P(i,j,h), где i, j и h – состояния первой, второй и третьей фаз соответственно. Первая фаза может находиться в состояниях i = {1,β}. Состояние i = 1 это состояние обслуживания запроса. Состояние i = β это состояние блокировки первой фазы по занятости второй фазы обслуживанием запроса. Вторая фаза может находиться в состояниях j = {1,0,β}. Состояние j = 1 это состояние обслуживания запроса. Состояние j = 0 это состояние ожидания поступления запроса из первой фазы. Состояние j = β это состояние блокировки второй фазы по занятости третьей фазы обслуживанием запроса. Третья фаза может находиться в состояниях h = {1,0}. Состояние h = 1 это состояние обслуживания запроса. Состояние h = 0 соответствует состоянию ожидания поступления запроса из второй фазы.

Процесс обслуживания запросов может быть представлен цепью Маркова, состоящей из 8 состояний. В стационарном режиме обслуживания функционирование модели описывается системой разностных уравнений баланса. Решение системы было получено для частного случая, когда интенсивности обслуживания запросов в каждой фазе одинаковы и равны единице, т.е.$\mu_1 = \mu_2 = \mu_3 = 1$, и определены характеристики системы: коэффициенты загрузки первой фазы (η_1), второй фазы (η_2) и третьей фаз (η_3), коэффициент загрузки системы (η), пропускная способность системы (*W*).

Коэффициенты загрузки фаз определяются суммарными вероятностями их обслуживания. Приведенные формулы используются для расчета коэффициентов и для последующих моделей. Коэффициент загрузки первой фазы определяется отсутствием состояний блокировки $\eta_1 = 1 - \Sigma P_{\beta\text{xx}} = 1 - 0{,}4359 = 0{,}5641$. Коэффициент загрузки второй фазы определяется отсутствием состояний её блокировки и простоя $\eta_2 = 1 - (\Sigma P_{\text{x}\beta\text{x}} + \Sigma P_{\text{x0x}}) = 1 - (0{,}2308 + 0{,}2051) = 0{,}5641$. Коэффициент загрузки третьей фазы определяется отсутствием состояний её простоя из-за ожидания $\eta_3 = 1 - \Sigma P_{\text{xx0}} = 1 - 0{,}4359 = 0{,}5641$.

Коэффициент загрузки системы η равен $\eta = \eta_1 = \eta_2 = \eta_3 = 0{,}5641$.

Пропускная способность системы *W* вычисляется через интенсивности обслуживания фаз μ_i и коэффициенты загрузки фазη_i:

$$W = \mu_i\eta_i = \mu_1\left(1 - \Sigma P_{\beta\text{xx}}\right) = \mu_2 1 - \left(\Sigma P_{\text{x}\beta\text{x}} + \Sigma P_{\text{x0x}}\right) = \mu_3(1 - \Sigma P_{\text{xx0}}) = 0{,}5641 \quad (1)$$

Трехфазная модель с одним общим буфером. Предполагается, что перед первой фазой располагается бесконечный буфер. Общий буфер используется второй фазой для считывания (*r*) запроса, помещенного туда первой фазой, или записи (ω) обработанного запроса для передачи в третью фазу.

Для анализа модели определены вероятности состояний $P_{i,j,h}(r/\omega)$, где i, j и h – состояния первой, второй и третьей фаз соответственно, $r=\overline{0,1}$ число запросов в буфере, r – буферизация по вводу, ω – буферизация по выводу.Первая фаза может находиться в состояниях $i=\{1,\beta\}$. Состояние $i=1$ – обслуживание. Состояние $i=\beta$ – блокировка фазы по занятости буфера или по занятости второй фазы обслуживанием. Вторая фаза может находиться в состояниях j= $\{1,0,\beta\}$. Состояние $j=1$ – обслуживание. Состояние $j=0$ – ожидание поступления запроса из первой фазы или из-за отсутствия запроса в буфере. Состояние $j=\beta$ – блокировка фазы по занятости буфера или по занятости третьей фазы обслуживанием. Третья фаза может находиться в состоянияx$h=\{1,0\}$. Состояние $h=1$ – обслуживание. Состояние $h=0$ соответствует ожиданию поступления запроса из второй фазы или из-за отсутствия запроса в буфере.

Процесс обслуживания запросов представлен цепью Маркова, состоящей из 15 состояний. Решение системы разностных уравнений баланса было получено для частного случая, когда интенсивности обслуживания запросов в каждой фазе одинаковы и равны единице ($\mu_1=\mu_2=\mu_3=1$) и емкость буфера $r=1$, а также определены характеристики системы: коэффициенты загрузки первой фазы (η_1), второй фазы (η_2) и третьей фаз (η_3), коэффициент загрузки системы (η), пропускная способность системы (*W*).

Коэффициент загрузки первой фазы равен $\eta_1=1-\Sigma P_{\beta\text{xx}}=1-0{,}3785=0{,}6215$. Коэффициент загрузки второй фазы равен$\eta_2=1-(\Sigma P_{\text{x}\beta\text{x}}+\Sigma P_{\text{x0x}})=1-(0{,}2203+0{,}1582)=0{,}6215$. Коэффициент загрузки третьей фазы равен$\eta_3 = 1 - \Sigma P_{\text{xx0}} = 1 - 0{,}3785 = 0{,}6215$. Коэффициент загрузки системы η равен $\eta = \eta_1 = \eta_2 = \eta_3 = 0{,}6215$.

Пропускная способность системы W вычисляется через интенсивности обслуживания фаз μ_i и коэффициенты загрузки фаз η_i:

$$W = \mu_i \eta_i = \mu_1\left(1 - \Sigma P_{\beta\text{xx}}\right) = \mu_2 1 - \left(\Sigma P_{\text{x}\beta\text{x}} + \Sigma P_{\text{x0x}}\right)$$
$$= \mu_3(1 - \Sigma P_{\text{xx0}}) = 0{,}6215 \qquad (2)$$

Для поиска альтернативных вариантов организации ввода/вывода и обработки рассмотрены две модификации приведенной выше модели.

Первая модификация. Оптимизируем модель, удалив из нее одно состояние, $P_{1\beta 1}(r)$, для исключения состояния блокировки. Процесс обработки запросов представляем цепью Маркова, состоящей уже из 14 состояний. Решение системы линейных алгебраических уравнений баланса было получено для частного случая, когда интенсивности обслуживания запросов в каждой фазе одинаковы и равны единице ($\mu_1 = \mu_2 = \mu_3 = 1$) и емкость буфера $r = 1$.

Получены следующие характеристики системы. Коэффициент загрузки первой фазы равен $\eta_1 = 1 - \Sigma P_{\beta\text{xx}} = 1 - 0{,}3579 = 0{,}6421$. Коэффициент загрузки второй фазы равен $\eta_2 = 1 - (\Sigma P_{\text{x}\beta\text{x}} + \Sigma P_{\text{x0x}}) = 1 - (0{,}1705 + 0{,}1874) = 0{,}6421$. Коэффициент загрузки третьей фазы равен $\eta_3 = 1 - \Sigma P_{\text{xx0}} = 1 - 0{,}3579 = 0{,}6421$. Суммарная загрузка системы равна η=$0{,}6421$.Пропускная способность системы W вычисляется через интенсивности обслуживания фаз μ_i и коэффициенты загрузки фаз:

$$W = \mu_i \eta_i = \mu_1\left(1 - \Sigma P_{\beta\text{xx}}\right) = \mu_2 1 - \left(\Sigma P_{\text{x}\beta\text{x}} + \Sigma P_{\text{x0x}}\right)$$
$$= \mu_3(1 - \Sigma P_{\text{xx0}}) = 0{,}6421 \qquad (3)$$

Вторая модификация. Оптимизируем модель, удалив из нее два состояния, $P_{1\beta 1}(r)$ и $P_{\beta\boldsymbol{\beta 1}}(r)$, для исключения состояния блокировки. Процесс обслуживания запросов представляем цепью Маркова, состоящей из 13 состояний. Решение системы линейных алгебраических уравнений баланса было получено для частного случая, когда интенсивности обслуживания запросов в каждой фазе одинаковы и равны единице ($\mu_1 = \mu_2 = \mu_3 = 1$) и емкость буфера $r = 1$.

Получены следующие характеристики системы. Коэффициент загрузки первой фазы равен $\eta_1 = 1 - \Sigma P_{\beta \text{xx}} = 1 - 0{,}3492 = 0{,}6508$. Коэффициент загрузки второй фазы равен $\eta_2 = 1 - (\Sigma P_{\text{x}\beta\text{x}} + \Sigma P_{\text{x0x}}) = 1 - (0{,}1538 + 0{,}1954) = 0{,}6508$. Коэффициент загрузки третьей фазы равен $\eta_3 = 1 - \Sigma P_{\text{xx0}} = 1 - 0{,}3492 = 0{,}6508$. Суммарная загрузка системы равна η = $0{,}6508$. Пропускная способность системы W вычисляется через интенсивности обслуживания фаз μ_i и коэффициенты загрузки фаз:

$$W = \mu_i \eta_i \ = \mu_1\left(1 - \Sigma P_{\beta\text{xx}}\right) = \mu_2 1 - \left(\Sigma P_{\text{x}\beta\text{x}} + \Sigma P_{\text{x0x}}\right)$$
$$= \mu_3(1 - \Sigma P_{\text{xx0}}) = 0{,}6508 \qquad (4)$$

Рассмотрим трехфазную модель с буферами ввода/вывода перед второй и третьей фазой. Предполагается, что перед первой фазой располагается бесконечный буфер. Данная модель приведена в [1] для частного случая $\mu_1 = \mu_2 = \mu_3 = 1$ и емкости буфера $r = 1$.

Для анализа модели определены вероятности состояний $P_{i,j,h}(r, \omega)$, где i, j и h – состояния первой, второй и третьей фаз соответственно, $r = \overline{0{,}1}$и $\omega = \overline{0{,}1}$число запросов в буфере, r – буферизация по вводу, ω – буферизация по выводу. Первая фаза может находиться в состояниях $i = \{1, \beta\}$. Состояние $i = 1$ – обслуживание. Состояние $i = \beta$ – блокировка фазы по занятости буфера или по занятости второй фазы обслуживанием. Вторая фаза может находиться в состояниях j = $\{1{,}0, \beta\}$. Состояние $j = 1$ – обслуживание. Состояние $j = 0$ – ожидание поступления запроса из первой фазы или из-за отсутствия запроса в буфере. Состояние $j = \beta$ – блокировка фазы по занятости буфера или по занятости третьей фазы обслуживанием. Третья фаза может находиться в состояниях $h = \{1,0\}$. Состояние $h = 1$ – обслуживание. Состояние $h = 0$ соответствует ожиданию поступления запроса из второй фазы или из-за отсутствия запроса в буфере.

Процесс обработки запросов представляем цепью Маркова, состоящей из 15 состояний. Решение системы линейных алгебраических уравнений баланса для стационарного режима обслуживания было получено для частного случая, когда интенсивности обслуживания запросов в каждой фазе одинаковы и равны единице ($\mu_1 = \mu_2 = \mu_3 = 1$) и емкость буфера $r = 1$.

Получены следующие характеристики системы. Коэффициент загрузки первой фазы равен $\eta_1 = 1 - \Sigma P_{\beta\text{xx}} = 1 - 0{,}3295 = 0{,}6705$. Коэффициент загрузки второй фазы равен $\eta_2 = 1 - (\Sigma P_{\text{x}\beta\text{x}} + \Sigma P_{\text{x0x}}) = 1 - (0{,}1732 + 0{,}1563) = 0{,}6705$. Коэффициент загрузки третьей фазы равен $\eta_3 = 1 - \Sigma P_{\text{xx0}} = 1 - 0{,}3295 = 0{,}6705$. Коэффициент загрузки системы η равен $\eta = \eta_1 = \eta_2 = \eta_3 = 0{,}6705$. Пропускная способность системы W вычисляется через интенсивности фаз μ_i и коэффициенты загрузки фазη_i:

$$\begin{aligned} W = \mu_i\eta_i &= \mu_1\left(1 - \Sigma P_{\beta\text{xx}}\right) = \mu_2 1 - \left(\Sigma P_{\text{x}\beta\text{x}} + \Sigma P_{\text{x0x}}\right) \\ &= \mu_3(1 - \Sigma P_{\text{xx0}}) = 0{,}6705 \end{aligned} \qquad (5)$$

Для упрощения анализа трехфазные схемы сводим к двухфазным схемам [3]: модель с асинхронным вводом и синхронным выводом (AISO), объединяя вторую фазу с третьей, и модель синхронного ввода и асинхронного вывода (SIAO), объединяя первую фазу со второй.

В модели AISO интенсивность обслуживания объединенной фазы равна $\mu_2^* = 1/\nu_2^*$, где ν_2^*это среднее время обслуживания во второй и третьей фазах.Для анализа модели были определены вероятности состояний $P_{i,j}(n)$, где iиj – состояния первой и второй объединенной фаз соответственно, n= $\overline{0,K}$ число запросов в буфере. Первая фаза может находиться в состояниях $i = \{1,\beta\}$. Состояние $i = 1$ – обслуживание. Состояние $i = \beta$ – блокировка обработки в первой фазе по занятости буфера или по занятости второй фазы обслуживанием. Вторая фаза может находиться в состояниях j= $\{1,0\}$. Состояние $j = 1$ – обслуживание. Состояние $j = 0$ – ожидание поступления запроса из первой фазы или из-за отсутствия запроса в буфере.

Процесс обработки запросов представляем цепью Маркова, состоящей из 15 состояний, и составляем систему линейных алгебраических уравнений баланса для стационарного режима обслуживания. Для частного случая $\mu_1 = \mu_2 = \mu_3 = 1$ и $n = 1$коэффициент загрузки первой фазы равен $\eta_1 = 0{,}6667$, коэффициент загрузки второй фазы равен $\eta_2^* = 0{,}3333$. Пропускная способность системы W вычисляется через интенсивность обработки и коэффициент загрузки первой фазы:

$$W = \mu_1 \eta_1 = 0{,}6667 \quad (6)$$

В модели SIAO [3] интенсивность обслуживания объединенной фазы равна $\mu_1^* = 1/\nu_1^*$, где ν_1^* это среднее время обслуживания в первой и во второй фазах. Анализ модели аналогичен анализу модели AISO. Только пропускная способность системы определяется через интенсивность обработки и коэффициент загрузки третьей фазы.

Для частного случая $\mu_1 = \mu_2 = \mu_3 = 1$ и $n = 1$ коэффициент загрузки первой фазы равен $\eta_1^* = 0{,}6667$, коэффициент загрузки второй фазы равен $\eta_3 = 0{,}3333$. Пропускная способность системы W вычисляется через интенсивность обработки и коэффициент загрузки третьей фазы:

$$W = \mu_3 \eta_3 = 0{,}3333 \quad (6)$$

В таблице 1 сведены значения различных характеристик и коэффициентов рассмотренных моделей.

Таблица 1

Значения пропускной способности для различных моделей

Модель	$\Sigma P_{\beta xx}$	$\Sigma P_{x\beta x}$	ΣP_{x0x}	ΣP_{xx0}	$\Sigma P_{\beta x}$	ΣP_{x0}	η_1	η_2	η_3	W
Трехфазная модель без буфера	0,4359	0,2308	0,2051	0,4359	-	-	0,5641	0,5641	0,5641	0,5641
Трехфазная модель с одним общим буфером	0,3785	0,2203	0,1582	0,3785	-	-	0,6215	0,6215	0,6215	0,6215
Трехфазная модель с одним общим буфером. Модификация 1	0,3579	0,1705	0,1874	0,3579	-	-	0,6421	0,6421	0,6421	0,6421
Трехфазная модель с одним общим буфером. Модификация 2	0,3492	0,1538	0,1954	0,3492	-	-	0,6508	0,6508	0,6508	0,6508
Трехфазная модель с двумя буферами	0,3295	0,1732	0,1563	0,3295	-	-	0,6705	0,6705	0,6705	0,6705
Модель AISO	-	-	-	-	0,6667	0,3333*	0,6667	0,3333*		0,6667
Модель SIAO	-	-	-	-	0,3333**	0,6667	0,6667**		0,3333	0,3333

Примечание: * – значение для объединенной второй и третьей фазы; ** – значение для объединенной первой и второй фазы.

В настоящее время развитие технологий позволило появиться большому количеству мобильных устройств, которые так же могут объединяться в сети. Появляются возможности построения новых сервисов и услуг на базе таких гибридных сетей, решать задачи отказоустойчивости и резервирования маршрутов и каналов передачи данных, распределённых вычислений и т.д. Важную роль в построении алгоритмов маршрутизации займёт итерационный метод, поскольку позволяет прогнозировать количество итераций, а значит и время для оценки при построении альтернативных маршрутов, поскольку высока вероятность того, что практическая пропускная способность системы в целом будет ниже ожидаемой при использовании других методов оценки.

Список использованных источников

1. Kan W., Ning Z. Analysis of dual tandem queues with a finite buffer capacity and non-overlapping service times and subject to breakdowns // IIE Transactions. 2015. V. 47. Is. 12.
2. Hunt G.C. Sequential arrays of waiting lines // Operations research. 1956. V. 4. № 6. P. 674-683.
3. Ермаков А.С. Анализ способов синхронно-асинхронного взаимодействия процессов ввода/вывода и обработки информации // Информационно-инновационные технологии: Интеграция науки, образования и бизнеса. Тр. междунар. науч.-практ. конф., посв. 75-л. КазНТУим. К.И. Сатпаева (27-28 ноября 2008). Алматы, 2008. С. 343-349.