

УДК 519.9: 681.5

МОДЕЛЬНЫЙ АНАЛИЗ ФУНКЦИОНИРОВАНИЯ КОММУНИКАЦИОННОГО СЕРВЕРА В РАСПРЕДЕЛЕННОЙ СИСТЕМЕ НЕПРЕРЫВНОГО МОНИТОРИНГА СОСТОЯНИЯ СЛОЖНЫХ ТЕХНИЧЕСКИХ ОБЪЕКТОВ

Беневоленский С.Б., Карпов А.В., Чернова Т.А.

*ГОУ ВПО «МАТИ» «Российский государственный технологический университет
им. К.Э. Циолковского», Москва, e-mail: sbb13@mail.ru*

В работе рассмотрены результаты модельного анализа, и на этой основе предложен метод расчета максимальной нагрузки на коммуникационные сервера систем оперативного приема и обработки информации, позволяющий оценить реакцию системы на резко осциллирующие нагрузки.

Ключевые слова: распределенные информационно-вычислительные системы, сервера приложений, массовое обслуживание

Распределенные информационные системы, построенные на базе вычислительных и телекоммуникационных сетей, занимают в современном мире все более значимое место за счет расширения сферы применения, повышения качества обслуживания пользователей [1-3, 7]. Совершенствование сетевых технологий позволило создавать корпоративные информационные системы и вычислительные сети большой размерности, обслуживающие тысячи пользователей. Развитие этих систем, обслуживающих большое количество пользователей и реализующих большое количество разнообразных прикладных программ, привело к возникновению проблемы [4-6], связанной с прогнозированием уровня нагрузки на сервера приложений.

Коммуникационный сервер системы непрерывного мониторинга состояния сложных технических объектов является ключевым звеном при передаче информации от контролируемых объектов до рабочих мест операторов. Особенностью программной архитектуры коммуникационного сервера является то, что фактически ра-

бота модуля взаимодействия с регистраторами не зависит от работы модуля передачи данных. При этом для исследования реакции коммуникационного сервера на входящие потоки данных различной интенсивности достаточно ограничиться использованием схемы, изображенной на рис. 1. Выдача данных регистраторами производится при получении от коммуникационного сервера соответствующей команды в ответ на запрос авторизации. Этот запрос формируется при достижении наличия определенного количества записей во временном хранилище данных. Передача данных со стороны регистратора продолжается до тех пор, пока не будет полностью опустошено хранилище данных или когда соединение с коммуникационным сервером не будет разорвано по таймауту (т.е., в терминах теории массового обслуживания, не будет удалено из очереди входящих сообщений коммуникационного сервера за сроком давности). При переполнении очередей модуля авторизации и диспетчера данных обработка сообщений в корневой очереди приостанавливается.

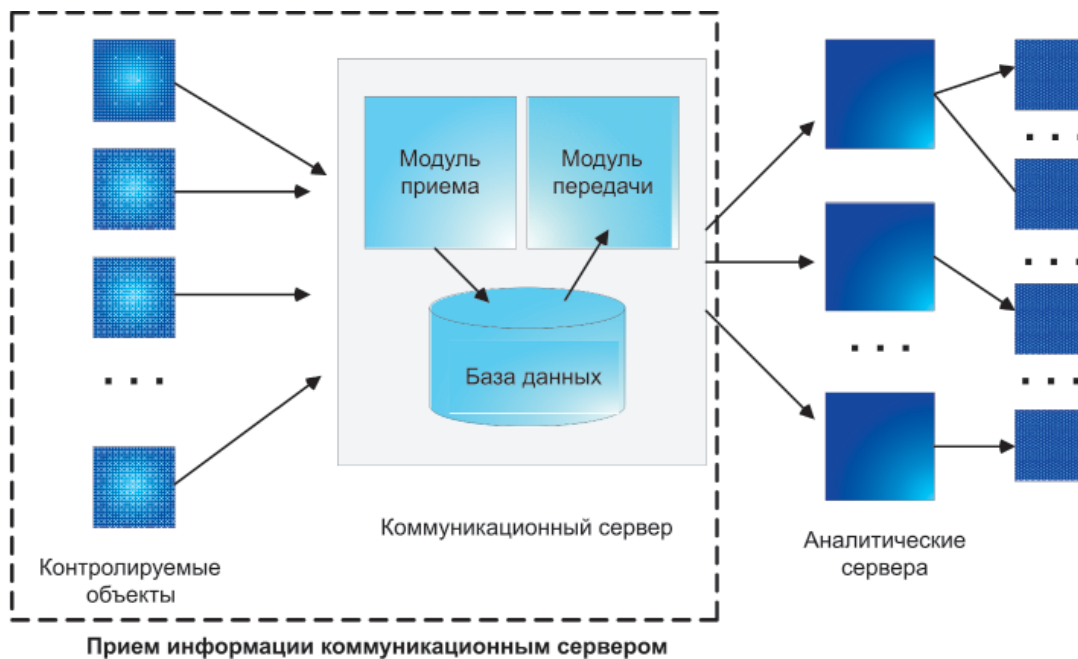


Рис. 1. Принципиальная схема передачи информации в системе непрерывного мониторинга состояния сложных технических объектов

Модельный анализ функционирования коммуникационного сервера, проводимый с учетом положений теории массового

обслуживания, позволяет представить сервер как систему, размеченный граф состояний которой изображен на рис. 2.

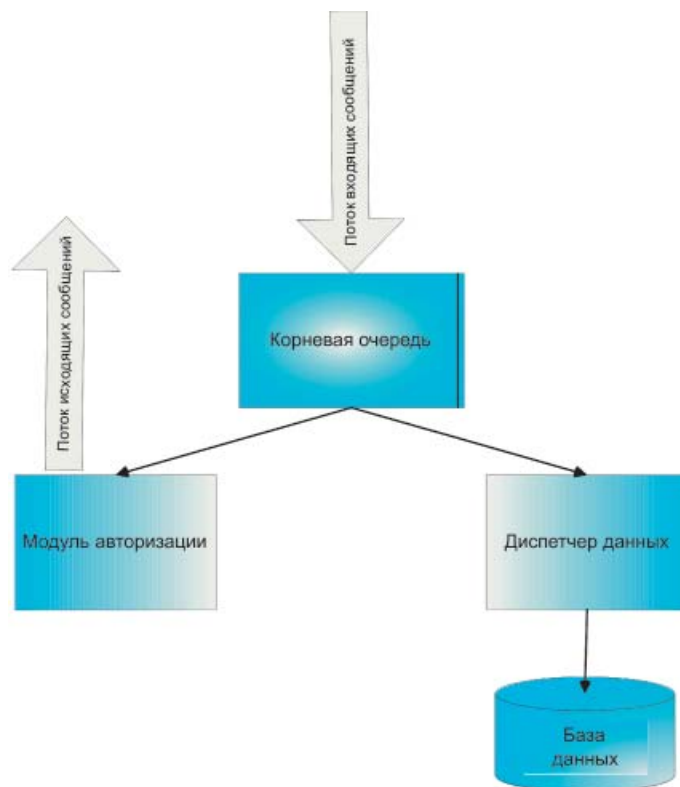


Рис. 2. Модель функционирования принимающего модуля Коммуникационного сервера

Интенсивности переходов системы из одного состояния в другое $\lambda_0, \lambda_1, \dots, \lambda_{g-1}$ можно интерпретировать как интенсивности возникновения заявок в системе. Аналогично величины $\mu_0, \mu_1, \dots, \mu_{g-1}$ – интенсивности переходов системы из состояния в состояние справа налево можно интерпретировать как интенсивности выполнения заявок в системе. Зная интенсивности переходов, можно найти величины $p_1(t), p_2(t), \dots, p_g(t)$ – вероятности нахождения системы S в состояниях S_1, S_2, \dots ,

S_g соответственно. При этом выполняется условие

$$\sum_{j=1}^n p_j(t) = 1. \quad (1)$$

Поскольку все состояния являются сообщающимися и существенными, существует [4] предельное распределение вероятностей состояний. Получим систему уравнений для такого случая распределения вероятностей состояний системы. В стационарных условиях для каждого состояния поток, втекающий в данное состояние, равен потоку, вытекающему из данного состояния. Таким образом, имеем:

для состояния S_0 :

$$p_0 \lambda_0 \Delta t = p_1 \mu_0 \Delta t; \quad \lambda_0 p_0 = \mu_0 p_1;$$

для состояния S_1 :

$$p_1 (\lambda_1 + \mu_0) \Delta t = p_0 \lambda_0 \Delta t + p_2 \mu_1 \Delta t; \quad (\lambda_1 + \mu_0) p_1 = \lambda_0 p_0 + \mu_1 p_2.$$

Последнее уравнение с учетом предыдущего можно привести к виду $\lambda_1 p_1 = \mu_1 p_2$. Аналогично можно получить уравнения

для остальных состояний системы. В результате состояния системы можно описать системой уравнений

$$\begin{cases} \lambda_0 p_0 = \mu_0 p_1, \\ \lambda_1 p_1 = \mu_1 p_2, \\ \dots \\ \lambda_k p_k = \mu_k p_{k+1}, \\ \dots \\ \lambda_{g-1} p_{g-1} = \mu_{g-1} p_g, \\ p_0 + p_1 + \dots + p_g = 1. \end{cases} \quad (2)$$

Последнее уравнение в (2) является очевидным условием (1).

Решение системы уравнений (2) имеет вид:

$$p_0 = \left(1 + \frac{\lambda_0}{\mu_0} + \frac{\lambda_0 \lambda_1}{\mu_0 \mu_1} + \frac{\lambda_0 \lambda_1 \lambda_2}{\mu_0 \mu_1 \mu_2} + \dots + \frac{\lambda_0 \lambda_1 \lambda_2 \dots \lambda_{g-1}}{\mu_0 \mu_1 \mu_2 \dots \mu_{g-1}} \right)^{-1}; \quad (3)$$

$$p_1 = \frac{\lambda_0}{\mu_0} p_0; \quad p_2 = \frac{\lambda_0 \lambda_1}{\mu_0 \mu_1} p_0; \quad p_3 = \frac{\lambda_0 \lambda_1 \lambda_2}{\mu_0 \mu_1 \mu_2} p_0; \quad \dots; \quad p_g = \frac{\lambda_0 \lambda_1 \lambda_2 \dots \lambda_{g-1}}{\mu_0 \mu_1 \mu_2 \dots \mu_{g-1}} p_0. \quad (4)$$

Рассмотрим модуль приема сообщений коммуникационного сервера как одноканальную систему массового обслуживания с ограниченной очередью, без ограничения нахождения заявки в очереди по времени. В системе массового обслуживания с ограниченной очередью число мест m в очереди ограничено. Следовательно, заявка, поступившая в момент времени, когда все места в очереди заняты, отклоняется и покидает систему массового обслуживания. Вероятность отказа в обслуживании заявки в такой системе равна:

$$p_{\text{отк}}(t) = \frac{\lambda(t)}{\mu} - \frac{1 - \frac{\lambda(t)}{\mu}}{1 - \left(\frac{\lambda(t)}{\mu}\right)^{m+2}}. \quad (5)$$

Анализируя динамику поступления и выполнения заявок в системе, необходимо отметить, что особенность системы мониторинга и оперативного анализа данных заключается в том, что необработанные сервером заявки накапливаются, увеличивая при этом поток входящих запросов. Обозначим количество поступающих заявок во входную очередь системы как $A(t)$, количество необработанных системой заявок – $B(t)$, а количество обработанных – $C(t)$. Новые заявки, формируемые регистраторами данных, обозначим $\varphi(t)$. На этой основе получим систему уравнений:

$$\begin{aligned} A(t) &= \varphi(t) + B(t); \\ B(t) &= A(t) \cdot p_{\text{отк}}(t) + B_0; \\ C(t) &= A(t) - B(t); \end{aligned} \quad (6)$$

$$\varphi(t) = \frac{N}{\tau} \cdot t,$$

где N – количество регистраторов данных; τ – временной интервал между формированием заявок одним регистратором; B_0 – количество заявок в системе в начальный момент времени.

Результаты численного эксперимента, представленные на рис. 3, позволяют

оценить уровень максимальной допустимой нагрузки на коммуникационный сервер системы при заданных технических характеристиках. Закономерности, полученные в результате проведенных натуральных экспериментов (рис. 4), позволяют оценить выход системы на ламинарный режим после резко осциллирующих нагрузок.

Сравнение результатов модельного анализа с показателями работы реально функционирующей системы показывает их достаточное согласование при исследовании загруженности коммуникационного сервера, функционирующего в системе мониторинга автопарков различных предприятий.

Таким образом, разработанный метод расчета максимальной нагрузки на коммуникационные сервера систем оперативного приема и обработки информации, а также оценки их реакции на резко осциллирующие нагрузки, основанный на предложенной математической модели коммуникационного сервера системы непрерывного мониторинга состояния сложных технических объектов, позволяет провести оптимизацию загруженности информационно-вычислительного комплекса автоматизированной системы управления многоагентной технологической средой.

Список литературы:

1. Афанасьев А.П., Батулин Ю.М., Еремченко Е.Н., Кириллов И.А., Клименко С.В. Информационно-аналитическая система для принятия решений на основе сети распределенных ситуационных центров // Информационные технологии и вычислительные системы. – 2010. – №2.
2. Афанасьев А.П., Волошинов В.В., Рогов С.В. и др. Проблемы вычислений в распределенной среде: Развитие концепции распределенных вычислительных сред // Труды института системного анализа РАН. – М.: УРСС, 2004.
3. Беневоленский С.Б., Жадан В.Г., Жадан И.В., Истомина Н.Л., Спыну М.В., Спыну С.К. Принципы создания программного обеспечения для систем распределенного вычисления. – Современные наукоемкие технологии – 2005. – №11. – С. 27–28.
4. Вишневецкий В.М. Теоретические основы проектирования компьютерных сетей. – М.: Техносфера, 2003.

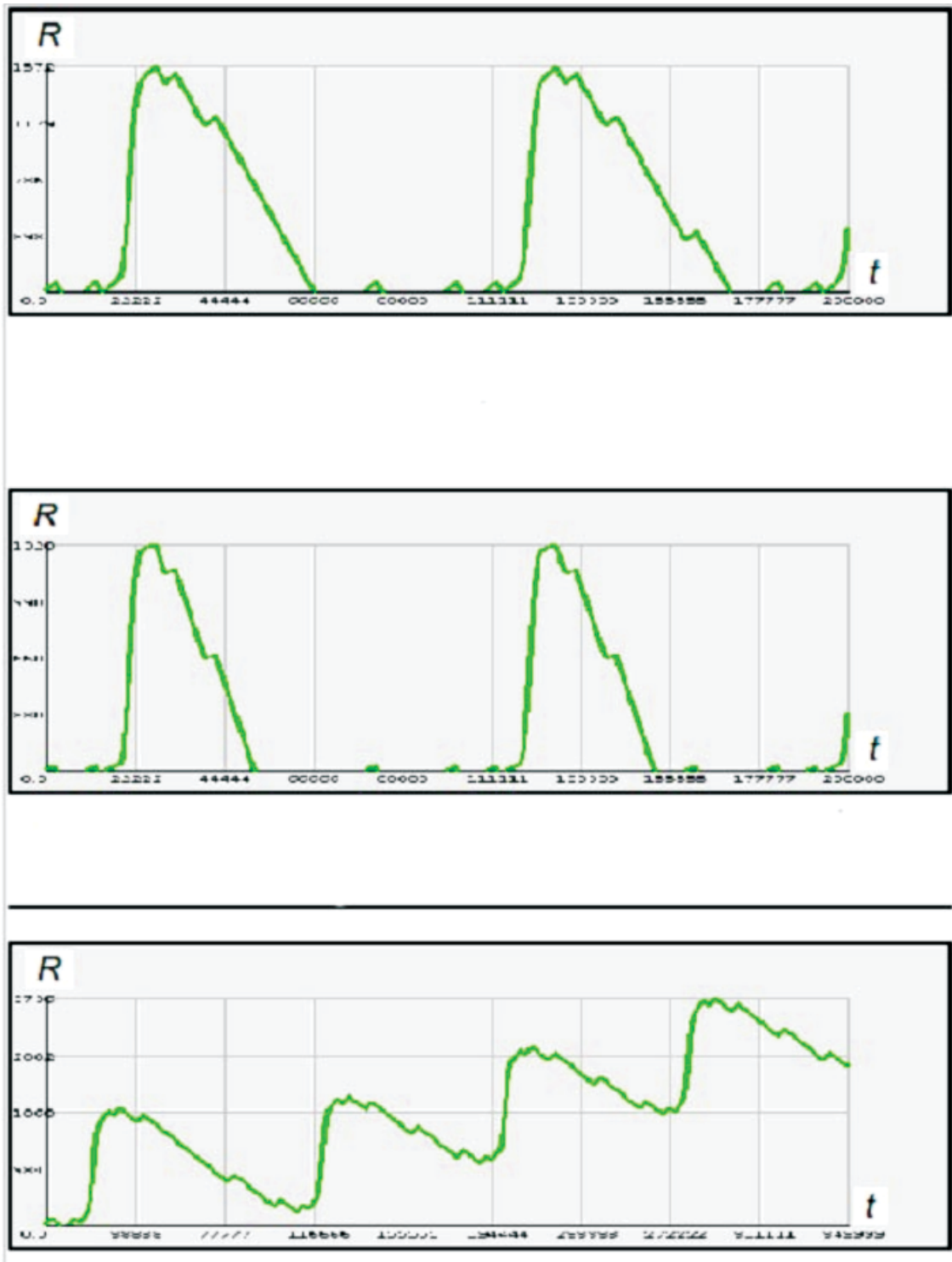


Рис. 3. Реакция серверов на резкое повышение нагрузок при разных пропускных способностях диспетчеров сообщений. Численный эксперимент

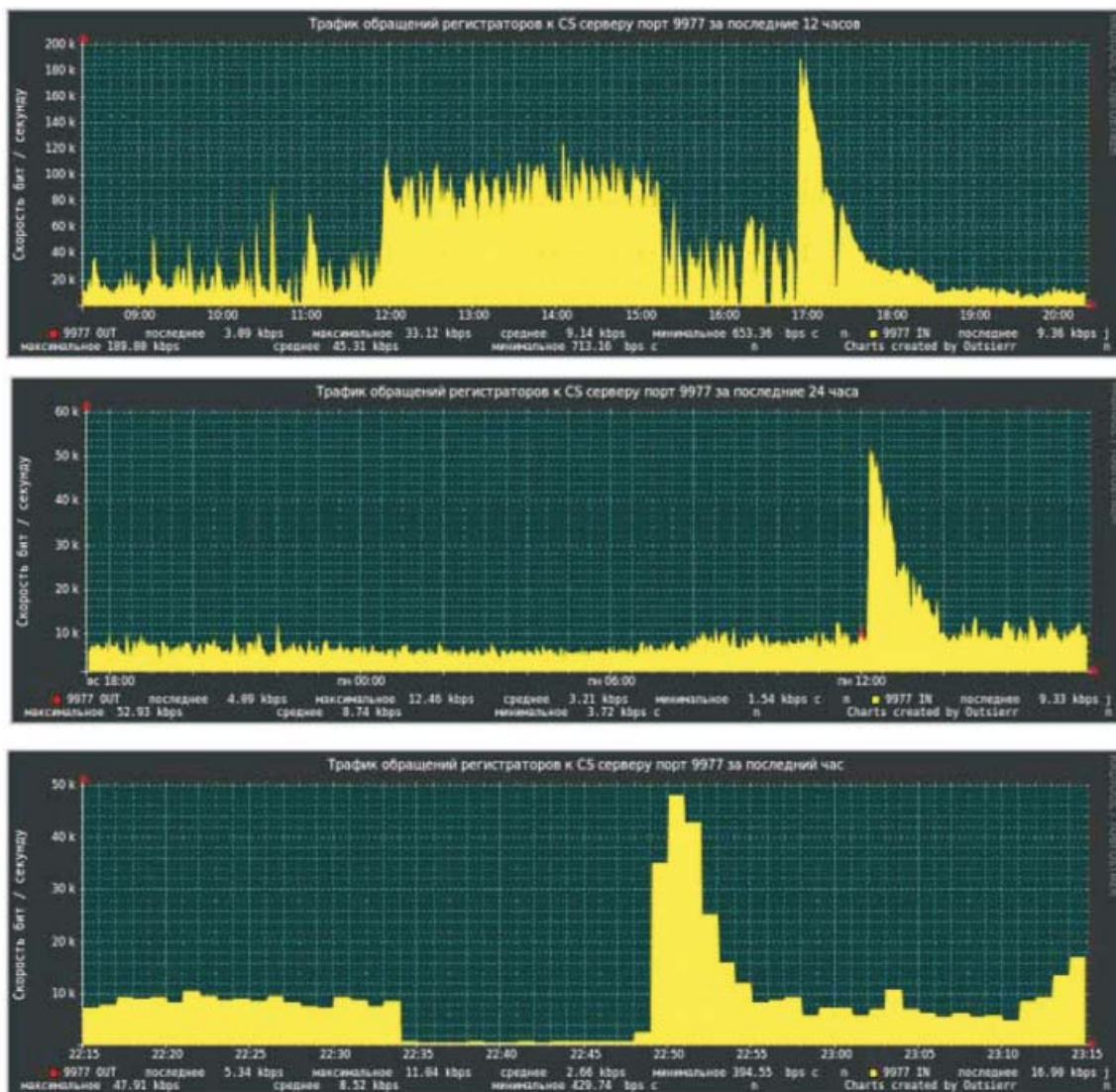


Рис. 4. Реакция серверов на резкое повышение нагрузок при разных пропускных способностях диспетчеров сообщений

5. Гнеденко Б.В., Коваленко И.Н. Введение в теорию массового обслуживания. – М.: Наука, 1966.

6. Григорьев О.Г. Модель работы сервера приложений в распределенной системе // Электросвязь. – 2004. – № 10.

7. Хританков А.С. Модели и алгоритмы распределения нагрузки. Алгоритм на основе сетей СМО // Информационные технологии и вычислительные системы. – 2009. – №3.

Рецензенты:

Марсов Вадим Израилевич, д.т.н., профессор Московского автомобильно-дорожно-

го государственного технического университета (МАДИ), кафедра «Автоматизация производственных процессов»;

Галушкин Александр Иванович, д.т.н., профессор, начальник лаборатории «Интеллектуальные информационные системы» Федерального государственного научного учреждения «Центр информационных технологий и систем органов исполнительной власти»;

Попов Федор Алексеевич, д.т.н., профессор, зам. директора по ИТ. Бийский технологический институт (филиал) ГОУ ВПО «Алтайский государственный технический университет им. И.И. Ползунова».

**MATHEMATICAL MODEL OF THE COMMUNICATION SERVER
IN THE DISTRIBUTED SYSTEM OF ON-LINE MONITORING
OF THE COMPLICATED TECHNICAL OBJECT'S CONDITIONS**

Benevolenskiy S.B., Chernova T.A., Karpov A.V.

Moscow state aviation technological university, Moscow, Russia,

e-mail: electron_inform@mail.ru

This work is focused on the results of model analysis of communication servers in operational systems for receiving and processing information. Based on that, methods of calculating the maximum load and response to rapidly oscillated load are proposed for these systems and their servers.

Keywords: distributed data-processing systems, application servers, queuing