

УДК 681.518.5
DOI

ОБ ОДНОМ ПОДХОДЕ К МОДЕЛИРОВАНИЮ СИСТЕМЫ, ОБЕСПЕЧИВАЮЩЕЙ СВОЕВРЕМЕННЫЙ ВЫХОД ИНФОРМАЦИИ В ЭФИР

Малахова Н.С.

ФГКВОВ ВО «Ярославское высшее военное училище противовоздушной обороны имени Маршала Советского Союза Л.А. Говорова» Министерства обороны Российской Федерации, Ярославль;
ФГАОУ ВО «Санкт-Петербургский государственный электротехнический университет „ЛЭТИ“ им. В.И. Ульянова (Ленина)», Санкт-Петербург, e-mail: natashamalakhova95@mail.ru

В статье описывается подход к моделированию системы, обеспечивающей своевременный выход информации в эфир. Представленное исследование имеет цель подтвердить гипотезу о том, что решение задачи структурно-параметрического синтеза с оптимизацией функционала стоимости при выполнении временных ограничений даст возможность проектирования системы передачи мультимедийных сообщений в реальном времени с использованием Интернета вещей. В результате синтеза вычислительного комплекса, помимо его комплектации, определяется уровень мультипрограммирования, размеры зон оперативной памяти в области прикладных программ и размещение страниц в иерархической внешней памяти, причем размеры зон и объемы блоков памяти кратны размеру страницы и измеряются числом страниц. Математическая модель представляет собой задачу оптимизации структуры многоуровневого графа системы по критерию стоимости устройств при ограничении на производительность каждого узла. Алгоритм расчета производительности каждого узла каждого уровня учитывает ограничения задержек, которые технически могут реализоваться в настройках соответствующего протокола. Оптимизация многоуровневого графа проводится методом ветвей и границ. Модельные расчеты включают анализ интенсивностей информационных потоков и оценку средних времен проведения транзакций, что подтверждает эффективность предложенной модели. Описанный подход может быть применен для проектирования и дальнейшей оптимизации сложных информационно-телекоммуникационных систем, реализованных с применением аппаратных устройств, технически ограниченных по уровню оперативной и постоянной памяти, а также систем хранящих информацию в распределенных базах данных.

Ключевые слова: Интернет вещей, распределенные вычислительные сети, методы обработки телеинформации, медиаданные, метод «ветвей и границ»

ABOUT ONE APPROACH TO MODELING A SYSTEM THAT ENSURES TIMELY INFORMATION OUTPUT ON THE AIR

Malakhova N.S.

Yaroslavl Higher Military School of Air Defense named after Marshal of the Soviet Union
L.A. Govorov of the Ministry of Defense of the Russian Federation, Yaroslavl;
Saint Petersburg State Electrotechnical University "LETI" named after V.I. Ulyanov (Lenin),
Saint Petersburg, e-mail: natashamalakhova95@mail.ru

The article describes an approach to modeling a system that ensures timely broadcasting of information. The presented study is aimed at confirming the hypothesis that solving the problem of structural-parametric synthesis with optimization of the cost functional while meeting time constraints will make it possible to design a system for transmitting multimedia messages in real time using the "Internet of Things". As a result of the synthesis of the computing complex, in addition to its configuration, the level of multiprogramming, the sizes of RAM zones in the area of application programs and the placement of pages in the hierarchical external memory are determined, with the sizes of zones and volumes of memory blocks being multiples of the page size and measured by the number of pages. The mathematical model is the problem of optimizing the structure of a multi-level graph of the system according to the criterion of device cost with a limitation on the performance of each node. The algorithm for calculating the performance of each node at each level takes into account the delay limitations that can technically be implemented in the settings of the corresponding protocol. Optimization of the multi-level graph is carried out using the branch and bound method. Model calculations include an analysis of the intensities of information flows and an estimate of the average transaction times, which confirms the effectiveness of the proposed model. The described approach can be applied to the design and further optimization of complex information and telecommunication systems implemented using hardware devices that are technically limited in terms of operational and permanent memory, as well as systems storing information in distributed databases.

Keywords: Internet of things, distributed computing networks, methods of processing teleinformation, media data, branch and bound method

Введение

Телекоммуникационные сети для передачи сообщений в эфир работают в системе реального времени, в которой приоритетные временные характеристики процесса.

Очевидно, что каждая информационная единица, передаваемая в подобной сети, имеет ограничение срока доставки, после которого информация становится бесполезной. Система, в которой происходит об-

мен данными в реальном времени, состоит из передающих устройств, называемых узлами (нодами), соединенными между собой для передачи данных, как оптоволоконными, так и радиоканалами. Организация взаимодействия описывается специальными протоколами. Архитектура такого взаимодействия требует тщательной настройки, что не всегда удобно в режиме реального времени, да еще при наличии жестких временных ограничений, поэтому задача определения наиболее подходящей структуры сети и определение значений параметров сети остается актуальным направлением.

Требования, на которые неоднократно ссылаются авторы, например [1], в том или ином виде формируют ограничения в задаче структурного и параметрического синтеза системы сводятся к понижению вероятностей искажений и потерь пакетов сообщений, повышению общей производительности системы при одновременном большом количестве соединений и, конечно, снижении накладных расходов как на обслуживание, так и на штатную работу системы. Для системы реального времени критическое значение имеет также существенное ограничение на величину задержки сообщения.

Необходимо указать, что протоколы TCP и UDP не соответствуют этим требованиям в полной мере. В настоящее время протоколом передачи медианых, в большей или меньшей степени отвечающим перечисленным выше требованиям, является протокол RTP (Real-Time Transport Protocol).

Однако и применение протокола RTP в сложных мультимедийных системах связано с необходимостью постоянной настройки и расстановки временных меток пакетов, так чтобы минимизировать общую задержку при передаче данных. А это в свою очередь требует наличия дополнительных узлов сбора и обработки информации, что усложняет структуру сети и увеличивает сложность проектирования подобной системы.

Таким образом, решение задачи оптимального параметрического и структурного синтеза системы передачи мультимедиа данных в реальном времени при условии минимизации задержек передачи остается актуальным в контексте разработки методики своевременного выхода сообщения в эфир [2].

Исходя из вышесказанного, **целью данного исследования** можно считать подтверждение гипотезы о том, что решение задачи структурно-параметрического синтеза с оптимизацией функционала стоимости при выполнении временных ограничений. Подтверждение данной гипотезы позволит на практике реализовать возможность про-

ектирования системы передачи мультимедийных сообщений в реальном времени с использованием новейших разработок из области IoT (Internet of things) – Интернета вещей.

Материалы и методы исследования

В типичном процессе сетевого приложения реального времени на основе Интернета вещей аудиоданные генерируются конечным устройством, подключенным к сети по IP-протоколу, с минимальной скоростью. Эти пакеты согласно протоколу UDP, указанному выше, доставляются на сервера, составляющие магистральную сеть. Здесь имеет место флуктуация времени передачи, то есть пакеты не доставляются в фиксированные моменты времени. Эта неравномерность нивелируется буферизацией, реализованной в протоколе RTP, играющей роль очереди, что позволяет с постоянной скоростью передавать данные устройству воспроизведения мультимедиа контента [3, 4].

Рассматриваемый мультимедийный вычислительный комплекс (далее – МВК) должен состоять из одного либо нескольких однотипных узлов, это могут быть мультимедийные сервера относительно небольшой производительности или же мобильные рабочие станции (поддерживающие обработку и сжатие данных формата HEVC и др.). Общий поток запросов на прием сообщений разделяется, используя механизм оркестрирования, подробно описанный в [5]. В основе модели лежит балансировка нагрузки с использованием подходов теории очередей.

В качестве исходных данных задаются технические мультимедийные характеристики серверов, пропускные способности каналов, средние частоты поступления сообщений, количества узлов [6, 7].

По аналогии с системами контейнеризации в данной задаче необходимо реализовать программно-аппаратный комплекс, осуществляющий диспетчирование потока запросов для передачи сообщений. Задержки, связанные с обслуживанием узла, учитываются при расчете быстродействия как известной системы «с охлаждением» [8, 9].

Комплектация МВК модулями подходит, исходя из требований к производительности системы при минимальной общей стоимости комплекса [10]. В результате синтеза МВК помимо его комплектации определяется максимально возможное число запросов, обрабатываемых одновременно на каждом узле, размеры зон оперативной памяти в области прикладных программ для каждого узла, размещение запросов в иерархической внешней памяти.

Результаты исследования и их обсуждение

Предполагается, что длительности обслуживания запросов в узлах подчиняются экспоненциальному закону [11].

Пусть r_k – число запросов в очереди, r_{n+1} – число запросов на обслуживание в узле. Число активных запросов вычисляется как

$$r = r_{n+1} + \sum_{k=1}^{k=n} r_k,$$

где $r = 0, \dots, R$, где R – принятое максимально возможное число запросов, обрабатываемых одновременно на каждом узле.

Величина $r_0 = R - r$ определяет число запросов, которые еще не зафиксированы системой передачи в определенный момент t .

Интенсивность появления сообщений в источнике обозначим λ . Интенсивность обслуживания σ_{n+1} – это величина, обратная среднему, взятому по сумме временных промежутков между моментами начала и окончания обработки запросов, за вычетом времени обработки случайных прерываний.

После обработки на сервере запрос поступает в очередь на дальнейшее обслуживание к одному из устройств внешней памяти. Интенсивность обслуживания λ_k в устройстве внешней памяти определяется как величина, обратная значению общего времени, необходимого для страничного обмена между оперативной памятью и блоком памяти, подключенным к k -му каналу без учета ожидания в очередях на обмен. По окончании обслуживания во внешнем устройстве запрос может вновь попасть на сервер для дополнительной обработки. Такое может произойти, например, в случае несовпадения контрольных меток в сообщениях большого объема. Если этого не произошло, то с вероятностью λ_0 завершает свое обслуживание в модуле мультимедийного вычислительного комплекса. Алгоритм расчета оценочных значений пропускной способности мультимедийного вычислительного комплекса описан в [12].

Итак, известны размер страницы оперативной памяти каждого узла и вероятности обращения к каждой странице, а также технико-экономические характеристики модулей мультимедийного вычислительного комплекса. Необходимо определить количество узлов (серверов, или рабочих станций), уровень мультипрограммирования, размер одной зоны оперативной памяти, количество блоков внешней памяти каждого типа и размещение страниц по блокам памяти таким образом, чтобы выполнялись ограничения на общую производительность всего

комплекса в целом, а приведенные итоговые затраты на модули комплекса были минимальными. Пусть V_0 – объем одного блока оперативной памяти, N – максимально допустимое число блоков оперативной памяти в мобильном вычислительном комплексе; G^* – стоимость мультимедийного сервера; a – объем, занимаемый оркестратором, включая балансировщик нагрузки и служебный блок; $\pi = \pi(a)$ неубывающая функция, представляющая зависимость среднего числа операций, выполняемых на сервере, от объема ресурсов балансировщика; G_0 – стоимость одного блока оперативной памяти; δ – среднее количество операций на узле, для обслуживания запроса, определяется с учетом «системных затрат», $T^{(1)}$ – среднее время подготовки к обслуживанию запроса; A – объем оперативной памяти, выделенной под область управляющей программы операционной системы; A_{jk} – объем, занимаемый транзитами операционной системы.

Далее, X_{ijk} – количество страниц с вероятностью обращения p_j , которые размещаются в блоках памяти j -го типа, подключенных к k -му каналу, m – количество узлов обслуживания; R – максимально возможное число одновременно обрабатываемых запросов на одном узле; λ – размер зоны основной памяти для обслуживания запроса.

Параметры модели функционирования целевой информационной системы определяются следующими соотношениями:

$$\sigma_0 = 1/T^{(1)}; \quad \sigma_{n+1} = f/\pi(a);$$

$$\sigma_k = \left\{ \sum_{j=1}^{j=N} E(X_{ijk}) \left(T_j + \frac{b}{v_j} \right) \right\}^{-1}, \quad k = 1, \dots, n;$$

$$E(x : x \geq 0) = \begin{cases} 1, & \text{если } x > 0, \\ 0, & \text{если } x = 0; \end{cases}$$

$$\lambda_k = \sum_{i=1}^{i=M} \sum_{j=1}^{j=N} p_i X_{ijk}, \quad k = 1, \dots, n;$$

$$\alpha_{n+1} = \pi(a)\delta^{-1}; \quad \alpha_0 = 1 - \alpha_{n+1}. \quad (1)$$

Используя эти соотношения, а также алгоритм для оценки производительности $\mu(R)$, коэффициент загрузки процессоров η_{n+1} и коэффициенты загрузки внешних устройств (η_k) можно рассчитать как функции вида

$$\mu = \mu(z, \pi(a), R, \{X_{ijk}\}),$$

$$\eta_s = \eta_s(z, \pi(a), R, \{X_{ijk}\}), \quad s = 1, \dots, (n+1). \quad (2)$$

Характеристики обработки информации телерепортажа

Номер типа информации i	Характеристика типа обрабатываемых данных	Интенсивность поступления λ_i	Среднее время отклика системы β_{il}
1	Исходные данные от источников систем хранения данных	1 раз/ч	2 с
2	Команды управления средствами коммутации	3 раза/ч	3 с
3	Другие исходные данные от источников	2 раза/мин	10 с
4	Получение обобщенных справок по запросам	5 раз/мин	5 с
5	Запросы на получение подробных справок	0,3 раза/ч	10 с
6	Запросы на проведение математических расчетов	1 раз/сут	0,5 мин
7	Запросы на предоставление специальной информации	100 раз/сут	30 с
8	Запросы на пересылку файлов	0,5 раз/час	0,015 ч

Введем переменные:

$$y_{jk} = \left[V_j^{-1} \left(b \sum_{i=1}^{i=M} (X_{ijk} + A_{jk}) \right) \right] - \text{число устройств с памятью } j\text{-го типа, подключенных}$$

к k -му каналу;

$$Z_j^{(1)} = \sum_{j=1}^{j=N} B_j^{-1} y_{jk} \left[- \text{количество узлов, подключенных к } k\text{-му каналу.} \right]$$

При решении задачи синтеза системы значения параметров системы должны удовлетворять следующим перечисленным ограничениям:

– производительность системы должна быть оптимизирована исходя из критериев количества запросов и блоков памяти, необходимых для обработки этих запросов;

– к одному и тому же каналу не подключаются различные устройства внешней памяти;
– количество мультимедийных серверов, подключенных к каналу и объем ресурсов, потребляемых балансировщиком ограничены.

Таким образом необходимо решить задачу многокритериальной оптимизации методом сведения к однокритериальной, то есть поиска минимума функции затрат:

$$G = \min \left\{ \sum_{i=1}^{i=M} \sum_{k=1}^{k=n} (G_k y_{jk} + G^{(1)} Z_k^{(1)}) + \sum_{k=1}^{k=n} G^{(2)} E(Z_k^{(1)}) \right\}.$$

В [13] описано применение метода «ветвей и границ» для поиска наилучших значений переменных. При построении дерева ветвлений выделяются четыре уровня. На первом уровне ветвление начинается из корневой вершины Ω . Из корневой вершины проводят дугу, которой соответствует значение $z = z_{\min} (z_{\min} > 1, \text{ целое})$.

На каждом следующем уровне ветвление продолжается из последней вершины, полученной на предыдущем уровне. Из этой вершины проводится дуга, которой соответствует значение $a = a_{\min} (a_{\min} \geq 1)$ [14].

В итоге обследования объектов автоматизации в телестудиях Санкт-Петербурга, анализа возможных сценариев использования автоматизированной системы передачи сообщений в эфир определены интенсив-

ности информационных потоков в ведомственной корпоративной сети, а по результатам экспериментов, натурных испытаний, компьютерного моделирования и сравнения с существующими аналогами оценены средние времена проведения транзакций (таблица) [15].

Для выполнения требований по своевременности сбора информации о происходящих интересных событиях в мире необходимо создание информационно-вычислительной системы с автоматическим сбором информации с «умных» вещей, в результате чего производительность возрастает в среднем на 30%, а вероятность своевременной обработки достигает 0,93, что является приемлемым для решения поставленной задачи.

Заключение

Представленный подход к реализации систем сбора информации в реальном времени может быть без ограничения общности применен к расчетам минимального времени, требуемого на подготовку к выходу информации в эфир и оптимизации системы передачи новостной информации. Он основан на обобщенной законченной модели сети массового обслуживания, алгоритмы расчета параметров которой известны и неоднократно опробованы в различных областях применения. Представленные предварительные расчеты временных показателей на основе данных новостного потока дали положительные результаты. Решение может быть применено для проектирования и дальнейшей оптимизации сложных информационно-телекоммуникационных систем, реализованных с применением аппаратных устройств, технически ограниченных по уровню оперативной и постоянной памяти, а также систем, хранящих информацию в распределенных базах данных.

Список литературы

1. Бабич А.В., Хаханов В.И., Мурад А.А. Исследование процессов передачи данных в реальном режиме времени // АСУ и приборы автоматики. 2009. № 147. URL: <https://cyberleninka.ru/article/n/issledovanie-protsessov-peredachi-dannyh-v-realnom-rezhime-vremeni> (дата обращения: 05.07.2025).
2. Майсурадзе А.И., Козлов В.Д. Моделирование задержек передачи информации в вычислительном кластере для мониторинга коммуникационной среды // Московский государственный университет имени М.В. Ломоносова. Труды XIX Международной конференции «Аналитика и управление данными в областях с интенсивным использованием данных» (DAMDID/ RCDL'2017) (Москва, 10–13 октября 2017 г.). EDN: XQUYQH.
3. Певнева А.Г., Обухов А.В., Швеиц Т.С. Методика интеллектуального планирования операций с прерываниями работы автоматизированной мобильной системы сбора данных // Вестник Российского нового университета. Серия: Сложные системы: модели, анализ и управление. 2023. № 4. С. 16–22. DOI: 10.18137/RNU.V9187.23.04.P.16.
4. Савельев М.Ф., Обухов А.В., Малахова Н.С., Воробьев Е.Г. Метод решения задачи безопасного движения транспорта и пешеходов на основе квантовых и нейрокompьютерных систем // Интеллектуальные технологии на транспорте. 2024. № 4 (40). С. 43–49. DOI: 10.20295/2413-2527-2024-440-43-49.
5. Лохвицкий В.А., Певнева А.Г., Левчик Э.С. Модель процесса распределенной обработки задач с учетом временных затрат на управление ресурсами // Авиакосмическое приборостроение. 2024. № 2. С. 40–46. DOI: 10.25791/aviakosmos.2.2024.1392.
6. Лазарев Е.А. Бикритериальная модель и алгоритмы оптимизации сети передачи данных: автореф. дис. ... канд. техн. наук. Нижний Новгород: НГТУ имени Р.Е. Алексеева, 2013. 23 с.
7. Пономарев Д.Ю. Оценка среднего времени задержки в инфокоммуникационных сетях с использованием методов тензорного анализа // Моделирование, оптимизация и информационные технологии. 2018. Т. 6. № 4. DOI: 10.26102/2310-6018/2018.23.4.013.
8. Лохвицкий В.А., Краснов С.А., Свиначук А.А. Метод повышения защищенности цифрового медиаконтента на основе использования идентификационных номеров // Вопросы оборонной техники. Серия 16: Технические средства противодействия терроризму. 2022. № 3–4 (165–166). С. 18–23. EDN: RBHNFZ.
9. Неретина К.А., Лохвицкий В.А., Захаров А.И., Брякалов Г.А. Исследование возможностей повышения оперативности выполнения последовательно-параллельных алгоритмов на основе использования технологии CUDA // Вопросы оборонной техники. Серия 16: Технические средства противодействия терроризму. 2022. № 9–10 (171–172). С. 151–156. DOI: 10.53816/23061456_2022_9-10_151. EDN: RVQTXE.
10. Лохвицкий В.А., Гончаренко В.А., Левчик Э.С. Модель масштабируемого микросервиса на основе системы массового обслуживания с «охлаждением» // Интеллектуальные технологии на транспорте. 2022. № 1 (29). С. 39–44. DOI: 10.24412/2413-2527-2022-129-39-44. EDN: ZALQRA.
11. Лохвицкий В.А., Захаров А.И., Брякалов Г.А., Неретина К.А. Исследование ускорения обработки данных в многоядерных вычислительных системах на основе имитационного моделирования // Вестник Российского нового университета. Серия: Сложные системы: модели, анализ и управление. 2022. № 1. С. 87–94. DOI: 10.18137/RNU.V9187.22.01.P.087. EDN: QYNESC.
12. Певнева А.Г., Зимовец А.И., Санжаревский Г.А. Системный анализ и моделирование критической нагрузки на информационные ресурсы // Вестник Российского нового университета. Серия: Сложные системы: модели, анализ и управление. 2020. № 1. С. 105–111. DOI: 10.25586/RNU.V9187.20.01.P.105.
13. Краснов С.А., Лохвицкий В.А., Хабаров Р.С. Модель представления данных при автоматической рубрикации разнородной информации в защищенных информационно-аналитических системах // Вопросы оборонной техники. Серия 16: Технические средства противодействия терроризму. 2022. № 3–4 (165–166). С. 24–31. EDN: ZNUDTN.
14. Певнева А.Г., Обухов А.В., Зимовец А.И. Модель нечеткого байесовского классификатора для обработки информации // Современные наукоемкие технологии. 2021. № 12–1. С. 78–83. DOI: 10.17513/snt.38958. EDN: BESTLU.
15. Гладышев А.И., Зимовец А.И., Нечай А.А., Обухов А.В. Применение Big Data для анализа околоземного космического пространства // Вестник Российского нового университета. Серия: Сложные системы: модели, анализ и управление. 2020. № 4. С. 127–134. DOI: 10.25586/RNU.V9187.20.04.P.127. EDN: NDTTXW.