

Literatura

1. Gorban' A.N., Rossiev D.A. Neironnye seti na personal'nom komp'yutere. Novosibirsk: Nauka, 1996. 276 s.
2. Zolotarev O.V., Sharnin M.M., Eromasova A., Tezadova F.M. Sovremennye podkhody k obrabotke mnogoyazychnykh tekstov, osnovannye na metodakh distributivnoj semantiki // Sbornik trudov mezhdunarodnoj nauchnoj konferentsii po fiziko-tehnicheskoy informatike – SRT2018 (Pushchino, 28–31 maya 2018 g.). Protvino, 2018. S. 43–47.
3. Camacho-Collados J., Mohammad T.P. On the Role of Text Preprocessing in Neural Network Architectures: An Evaluation Study on Text Categorization and Sentiment Analysis // Cornell University. URL: <https://arxiv.org/abs/1707.01780> (date of the application: 23.08.2018).
4. Mikolov T., Chen K., Corrado G., Dean J. Efficient Estimation of Word Representations in Vector Space // Cornell University. URL: <https://arxiv.org/abs/1301.3781> (date of the application: 16.01.2013).
5. Mikolov T., Sutskever I., Chen K., Corrado G., Dean J. Distributed Representations of Words and Phrases and Their Compositionality // Cornell University. URL: <https://arxiv.org/abs/1310.4546> (date of the application: 16.10.2013).
6. Molino P., Wang Y., Zhang J. Parallax: Visualizing and Understanding the Semantics of Embedding Spaces via Algebraic Formulae // Cornell University. URL: <https://arxiv.org/pdf/1905.12099.pdf> (date of the application: 27.01.2020).

DOI: 10.25586/RNU.V9I187.20.01.P.099

УДК 519.64+004.75+004.94

Р.С. Хабаров, Ю.С. Фоменко

МЕТОДИКА ОРГАНИЗАЦИИ ОБРАБОТКИ ДАННЫХ В ЕДИНОЙ
ТЕРРИТОРИАЛЬНО-РАСПРЕДЕЛЕННОЙ ИНФОРМАЦИОННОЙ
СИСТЕМЕ ДИСТАНЦИОННОГО ЗОНДИРОВАНИЯ ЗЕМЛИ

Представлена методика организации обработки данных с космических аппаратов дистанционного зондирования Земли в Единой территориально-распределенной информационной системе, основанная на оптимизации назначения смешанных приоритетов в сети массового обслуживания с учетом специфики обработки данных дистанционного зондирования Земли. Приведен пример реализации предложенной методики для обработки данных с космического аппарата «Ресурс-П». *Ключевые слова:* обработка данных дистанционного зондирования Земли, сети массового обслуживания, оптимизация сетей массового обслуживания, многоканальные системы, приоритетное обслуживание, смешанные приоритеты.

R.S. Khabarov, Yu.S. Fomenko

DATA PROCESSING ORGANIZATION METHODOLOGY
IN THE UNIFIED TERRITORIAALLY DISTRIBUTED EARTH
REMOTE SENSING INFORMATION SYSTEM

A methodology for data processing organizing from Earth remote sensing spacecraft in the Unified Territorial Distributed Information System is presented. It is based on optimization of the mixed priorities

assignment in queuing network taking into account the specifics of processing of Earth remote sensing data. An example of the implementation of the proposed processing data methodology from the "Resurs-P" spacecraft is given.

Keywords: remote sensing data processing, queuing networks, queuing network optimization, multi-channel systems, priority service, mixed priorities.

Введение

В настоящее время происходит завершение развертывания наземной инфраструктуры сбора, приема, обработки, хранения и распространения данных Единой территориально-распределенной информационной системы (ЕТРИС) дистанционного зондирования Земли (ДЗЗ) [5]. В [1] отмечается, что, в связи с увеличением орбитальной группировки космических аппаратов (КА) ДЗЗ и соответствующим ростом объемов принимаемой информации [2; 6], в ближайшие годы значительно увеличится нагрузка на средства обработки данных ДЗЗ. Вследствие этого наиболее актуальным становится вопрос повышения оперативности [3] предоставления продуктов ДЗЗ потребителям [8].

В данной работе предлагается методика организации обработки данных на основе гибкого назначения приоритетов в вычислительных узлах центра обработки данных (ЦОД) ДЗЗ. Математической основой методики является подход к оптимизации назначения смешанных приоритетов в сети массового обслуживания (СеМО), предложенный в [4].

Этапы методики

1. Формализация исходных данных. Обработываемые изображения с КА ДЗЗ представляются в виде типов заявок СеМО. Для уменьшения размерности задачи маршруты с близкими характеристиками (временами обработки этапов, требованиями по оперативности) объединяются в один тип. Этапы обработки представляются в виде узлов СеМО. На каждый этап задано определенное количество виртуальных машин (ВМ) с установленным специальным программным обеспечением (СПО) [7].

На основе собранной в ходе работы ЦОД ДЗЗ статистики происходит определение следующих параметров, используемых в дальнейшем алгоритмом оптимизации:

- значения интенсивностей поступающих потоков изображения от каждого типа КА ДЗЗ L^k , $k = \overline{1, K}$;
- маршрутные матрицы $R^k = \{r_{ij}^k\}$, задающие порядок обработки для изображения k -го типа и характеризующие вероятности переходов из i -го в j -й этап;
- начальные моменты распределения времени обработки на каждой из виртуальных машин (ВМ) во всех этапах $\{b_{j,l}^k\}$, $l = \overline{1, L}$, $k = \overline{1, K}$, $j = \overline{1, M}$.

Заданы ограничения на средние времена обработки по каждому из типов изображений ДЗЗ в виде вероятностей $P_{\text{троб}}^k (V^k < V_{\text{троб}}^k)$ и комплексный показатель оперативности

$$P_{\text{дц}} = \prod_{i=1}^N P^k (V^k < V_{\text{троб}}^k), \quad (1)$$

где N – количество типов изображений ДЗЗ.

Обозначим вариант назначения смешанного приоритета в узле коротежем из двух чисел

$$\{X_i^k\} = \langle PR_i^k, NP_i^k \rangle, \quad k = \overline{1, K}, \quad i = \overline{1, M},$$

где PR_i^k , NP_i^k – значения абсолютного и относительного для k -го типа в i -м узле приоритета соответственно.

Значения приоритетов нумеруются с единицы: чем ниже приоритет типа заявки в узле, тем больше значение.

Требуется найти множество вариантов назначения смешанных приоритетов по узлам сети $\{X_i^k\}$, при котором обеспечивается максимум комплексного показателя оперативности $\{X_i^k\} = Arg \max \mathbf{P}(P_{\text{треб}}^k, \lambda^k, \{b_{j,l}^k\}, R^k), l = \overline{1, L}, k = \overline{1, K}, j = \overline{1, M}$ при заданных ограничениях на вероятности непревышения времени обработки $P_{\text{треб}}^k (V^k < V_{\text{треб}}^k)$.

2. Осуществляется предварительный расчет на основе модели СеМО с многоканальными узлами и учетом процессов Split-Join с помощью методов, представленных в [6; 7]. Узлы СеМО Split-Join описывают заданные технологией обработки процессы разделения изображений на отдельные фрагменты, параллельную обработку фрагментов и финальную «сшивку» в готовый продукт. Выявляются типы изображений ДЗЗ, для которых не удовлетворяются ограничения по своевременности обработки.

3. Оптимизация назначения приоритетов в СеМО с помощью метода, представленного в [4].

4. В случае успешной работы на этапе 3 происходит назначение приоритетов для планируемых к обработке типов данных ДЗЗ. При невозможности удовлетворения требований вырабатываются рекомендации, которые могут содержать необходимое количество дополнительных унифицированных серверов.

Пример применения методики

Рассмотрим пример применения методики для организации обработки с трех КА ДЗЗ «Ресурс-П» по трем технологиям. Итоговое количество поступающих типов маршрутов $K = 9$. Обработка осуществляется в комплексе вычислительных ресурсов, где под каждый этап выделено определенное количество ВМ с соответствующим программным обеспечением. В таблице 1 указаны условные данные – начальные моменты времени обработки, рассчитанные для ВМ с 16 GB ОЗУ и 4 ядрами процессоров.

Таблица 1

Исходные данные

Тип обработки	Количество ВМ, шт.	Номер операции	Начальные моменты времени обработки, мин								
			КА № 1			КА № 2			КА № 3		
			Технологии обработки								
			1	2	3	1	2	3	1	2	3
Восстановление информации	6	1	4.224	4.226	2.926	4.244	4.244	4.244	2.542	2.542	2.542
			42.90	42.90	42.90	42.24	42.24	42.24	5.046	5.046	5.046
			725.2	725.2	725.2	582.5	582.5	582.5	42.95	42.95	42.95
Геодезическая привязка	6	2	4.442	4.442	4.442	2.486	2.486	2.486	2.654	2.654	2.654
			25.92	25.92	25.92	25.69	25.69	25.69	27.88	27.88	27.88
			242.8	242.8	242.8	259.0	259.0	259.0	294.5	294.5	294.5
Радиометрическая коррекция и «Сшив-ка» спектрональ-ных изображений	10	3	4.222	4.222	4.222	4.244	4.244	4.244	4.222	4.222	4.222
			24.74	24.74	24.74	26.55	26.55	26.55	24.57	24.57	24.57
			424.6	424.6	424.6	450.0	450.0	450.0	422.6	422.6	422.6
Радиометрическая коррекция и «Сшив-ка» панхроматиче-ских изображений	10	4	2.886	2.886	2.886	2.988	2.988	2.988	4.008	4.008	4.008
			22.25	22.25	22.25	22.67	22.67	22.67	22.96	22.96	22.96
			248.7	248.7	248.7	276.2	276.2	276.2	282.6	282.6	282.6

Тип обработки	Количество ВМ, шт.	Номер операции	Начальные моменты времени обработки, мин								
			КА № 1			КА № 2			КА № 3		
			Технологии обработки								
			1	2	3	1	2	3	1	2	3
Геодезическое ориентирование детального снимка	12	5	–	22.42	–	–	22.24	–	–	8.440	–
			–	484.7	–	–	479.7	–	–	276.2	–
			–	2.94e5	–	–	2.89e5	–	–	5.98e4	–
Ортогональное преобразование и pansharpning	14	6	–	20.42	–	–	20.78	–	–	22.92	–
			–	270.4522	–	–	294.9	–	–	424.7	–
			–	2.24e5	–	–	2.29e5	–	–	2.24e5	–
Расчет карты диспарантности	16	7	–	–	22.72	–	–	20.92	–	–	22.20
			–	–	448.7	–	–	402.7	–	–	477.8
			–	–	2.67e5	–	–	2.44e5	–	–	2.88e5
Формирование цифровых моделей местности	10	8	–	–	22.44	–	–	22.24	–	–	24.22
			–	–	492.8	–	–	420.2	–	–	446.4
			–	–	2.99e5	–	–	2.46e5	–	–	2.44e5

В таблице 2 приведены представленные в [5] интенсивности поступающих данных по типам маршрутов.

Таблица 2

Интенсивности поступления задач на обработку

Интенсивность $L_k, \text{мин}^{-1}$	КА № 1			КА № 2			КА № 3		
	1	2	3	1	2	3	1	2	3
$L_k, \text{мин}^{-1}$	0.21	0.21	0.22	0.15	0.18	0.17	0.15	0.11	0.14

На рисунках 1–3 представлены схемы посещения узлов в СеМО.

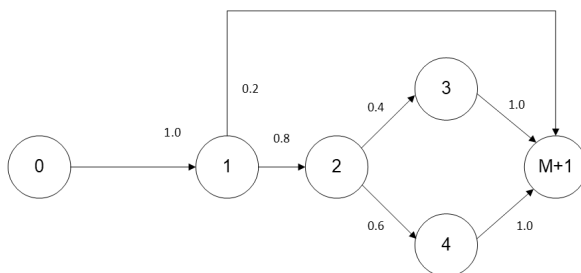


Рис. 1. СеМО для маршрутов с первой технологией обработки

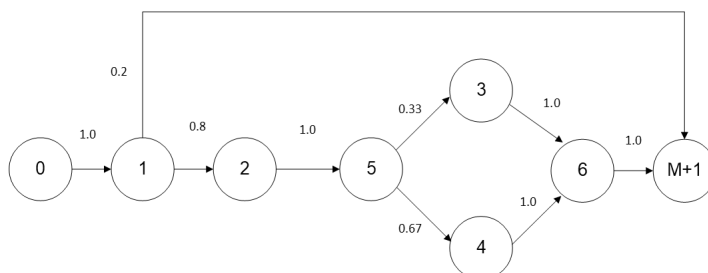


Рис. 2. СеМО для маршрутов со второй технологией обработки

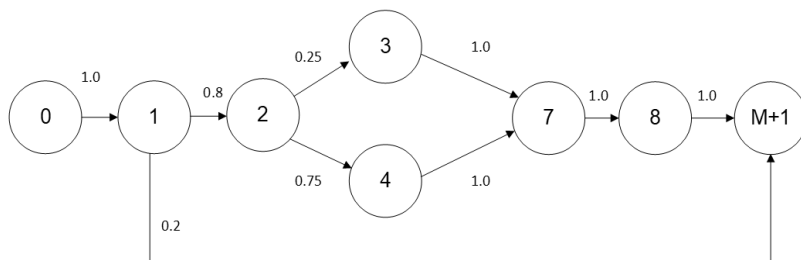


Рис. 3. СеМО для маршрутов с третьей технологией обработки

В таблице 3 приведены ограничения и расчеты до и после применения оптимизации.

Таблица 3

Результаты расчетов до и после применения методики

Параметр	Номер типа заявки								
	1	2	3	4	5	6	7	8	9
$V_{\text{треб}}^k$	37.00	190.0	250.0	45.00	185.0	250.0	55.00	160.0	265.0
$P_{\text{треб}}^k$	0.950	0.950	0.950	0.950	0.950	0.950	0.950	0.950	0.950
$P_{\text{до опт}}^k$	0.831	0.934	0.992	0.822	0.917	0.981	0.842	0.925	0.993
$P_{\text{после опт}}^k$	0.962	0.973	0.978	0.963	0.976	0.979	0.961	0.988	0.982

Как видно из таблицы 3, до применения методики для изображений ДЗЗ типов 1, 4 и 7 вероятности $P_{\text{до опт}}^k$ превышают требуемые $P_{\text{треб}}^k$. Рассчитанный по формуле (1) комплексный показатель оперативности до применения методики составил 0.44, после – 0.785. Прирост показателя около 34,5%.

Проводились расчеты при различных вариантах исходных данных (конфигурации ВМ, увеличении интенсивностей поступления данных ДЗЗ, изменении тактико-технических характеристик перспективных КА), в ходе которых диапазон прироста комплексного показателя оперативности составил 12–42%.

Заключение

Предложенная методика позволяет повысить оперативности обработки данных в ЕТРИС ДЗЗ в диапазоне от 12 до 42% в случаях существенного повышения интенсивности потоков поступающей информации и отказов вычислительных узлов ЦОД.

Литература

1. Лошкарёв П.А. и др. Развитие ЕТРИС ДЗЗ с применением облачных технологий // Геоматика. URL: <http://geomatica.ru/clauses/137/> (дата обращения: 27.01.2020).
2. Нечай А.А. Формирование безопасной информационной среды // Актуальные проблемы современности: наука и общество. 2019. № 4. С. 43–44.

3. *Нечай А.А., Борисов А.А., Борисова Ю.И.* Точечный анализ данных дистанционного зондирования Земли средствами языка программирования Python // Вестник Российского нового университета. Серия «Сложные системы: модели, анализ и управление». 2019. Вып. 1. С. 49–55.
4. Программа оптимизации назначения приоритетов в сети массового обслуживания с многоканальными узлами / *Хабаров Р.С.* Свидетельство о государственной регистрации программы для ЭВМ № 2019663860, Российская Федерация; заявл. 29.08.19; зарегистр. 24.10.19.
5. Роскосмос завершает создание ЕТРИС ДЗЗ // Совзона. URL: <https://sovzond.ru/press-center/news/dzz/5624/> (дата обращения: 27.01.2020).
6. *Рыжиков Ю.И., Лохвицкий В.А., Хабаров Р.С.* Метод расчета длительности обработки задач в системе массового обслуживания с учетом процессов Split-Join // Известия вузов. Приборостроение. 2019. № 5. С. 419–422.
7. *Хабаров Р.С., Лохвицкий В.А.* Модель оценивания оперативности многопоточной обработки задач в распределенной вычислительной среде с учетом процессов Split-Join // Вестник Российского нового университета. Серия «Сложные системы: модели, анализ и управление». 2019. Вып. 1. С. 26–34.
8. *Шаймарданов А.М., Нечай А.А., Лепехин С.В.* Математическая модель систем автоматического управления с широтно-импульсной модуляцией // Вестник Российского нового университета. Серия «Сложные системы: модели, анализ и управление». 2019. Вып. 2. С. 27–39.

Literatura

1. *Loshkaryov P.A. i dr.* Razvitiye ETRIS DZZ s primeneniem oblachnykh tekhnologij // Geomatika. URL: <http://geomatika.ru/clauses/137/> (data obrashcheniya: 27.01.2020).
2. *Nechaj A.A.* Formirovaniye bezopasnoy informatsionnoy sredy // Aktual'nye problemy sovremenosti: nauka i obshchestvo. 2019. № 4. S. 43–44.
3. *Nechaj A.A., Borisov A.A., Borisova Yu.I.* Tochechnyj analiz dannykh distantsionnogo zondirovaniya Zemli sredstvami yazyka programmirovaniya Python // Vestnik Rossijskogo novogo universiteta. Seriya "Slozhnye sistemy: modeli, analiz i upravlenie". 2019. Vyp. 1. S. 49–55.
4. Programma optimizatsii naznacheniya prioritetov v seti massovogo obsluzhivaniya s mnogokanal'nymi uzlami / *Khabarov R.S.* Svidetel'stvo o gosudarstvennoj registratsii programmy dlya EVM № 2019663860, Rossijskaya Federatsiya; zayavl. 29.08.19; zaregistr. 24.10.19.
5. Roskosmos zavershaet sozdanie ETRIS DZZ // Sovzond. URL: <https://sovzond.ru/press-center/news/dzz/5624/> (data obrashcheniya: 27.01.2020).
6. *Ryzhikov Yu.I., Lohvitskij V.A., Khabarov R.S.* Metod rascheta dlitel'nosti obrabotki zadach v sisteme massovogo obsluzhivaniya s uchedom protsessov Split-Join // Izvestiya vuzov. Pribo-rostroenie. 2019. № 5. S. 419–422.
7. *Khabarov R.S., Lohvitskij V.A.* Model' otsenivaniya operativnosti mnogopotochnoj obrabotki zadach v raspredelennoj vychislitel'noj srede s uchedom protsessov Split-Join // Vestnik Rossijskogo novogo universiteta. Seriya "Slozhnye sistemy: modeli, analiz i upravlenie". 2019. Vyp. 1. S. 26–34.
8. *Shajmardanov A.M., Nechaj A.A., Lepexin S.V.* Matematicheskaya model' sistem avtomaticheskogo upravleniya s shirotno-impul'snoj modulyatsiej // Vestnik Rossijskogo novogo universiteta. Seriya "Slozhnye sistemy: modeli, analiz i upravlenie". 2019. Vyp. 2. S. 27–39.

А.Г. Певнева, А.И. Зимовец, Г.А. Санжаревский

СИСТЕМНЫЙ АНАЛИЗ И МОДЕЛИРОВАНИЕ КРИТИЧЕСКОЙ НАГРУЗКИ НА ИНФОРМАЦИОННЫЕ РЕСУРСЫ

Приводятся результаты предварительного анализа ситуации критической нагрузки, дается описание методик нагрузочного тестирования информационных ресурсов в ситуации переполнения буфера запросов по различным протоколам.

Ключевые слова: нагрузочное тестирование, системный анализ, критическая нагрузка.

A.G. Pevneva, A.I. Zimovets, G.A. Sanzharevskij

SYSTEM ANALYSIS AND MODELING OF CRITICAL LOAD ON INFORMATION RESOURCES

The results of a preliminary analysis of the critical load situation are given, and methods for load testing information resources in situations of request buffer overflows using various protocols are described.

Keywords: load testing, system analysis, critical load.

Введение

Системное исследование любого объекта или процесса есть последовательность взаимно влияющих этапов: сбор и анализ данных, математическое моделирование, оптимальное проектирование. Целью настоящего исследования служит выявление проблематики *первого этапа*, представляющего собой программное моделирование ситуации критической нагрузки на информационные ресурсы и предмодельную обработку полученной статистической информации.

Критическая нагрузка на информационный ресурс возникает в том числе в результате распределенной сетевой атаки типа «отказ в обслуживании» (DDoS-атаки). Как объект системного исследования она интересна для изучения, потому что, во-первых, угроза подобных атак остается актуальной на протяжении многих лет, даже такие мощные сетевые ресурсы, как eBay или Yahoo, страдают от DDoS-атак [4]. Известно три различных метода организации атак такого типа: превышение пропускной способности полосы, превышение возможности сервера, использование ошибок ресурса. Первый вид атаки предполагает, что на ресурс сети направляется большое количество запросов по различным протоколам, превышая их пропускную способность. Атаки второго типа вызывают «падение» конкретной службы сервера. Третий тип атаки требует кропотливого исследования программной организации ресурса для выявления уязвимости. Примером уязвимостей считаются скрипты, выполняющие огромное количество запросов к базе данных и использующие процессорное время.

По этому направлению существует достаточно открытых источников, содержащих статистические данные. Эти данные, несмотря на их разнородность и даже в некоторых