

УДК 004.6

МНОГОЭТАПНОЕ ОБСЛУЖИВАНИЕ ЗАПРОСОВ, КРИТИЧНЫХ К ЗАДЕРЖКАМ ОЖИДАНИЯ, В МНОГОУРОВНЕВЫХ СИСТЕМАХ

В.А. Богатырев^a, С.В. Богатырев^b

^a Университет ИТМО, Санкт-Петербург, 197101, Российская Федерация

^b Самсунг Электроникс, Сеул, 137-965, Южная Корея

Адрес для переписки: Vladimir.bogatyrev@gmail.com

Информация о статье

Поступила в редакцию 21.07.17, принята к печати 22.08.17

doi: 10.17586/2226-1494-2017-17-5-872-878

Язык статьи – русский

Ссылка для цитирования: Богатырев В.А., Богатырев С.В. Многоэтапное обслуживание запросов, критичных к задержкам ожидания, в многоуровневых системах // Научно-технический вестник информационных технологий, механики и оптики. 2017. Т. 17. № 5. С. 872–878. doi: 10.17586/2226-1494-2017-17-5-872-878

Аннотация

Предмет исследования. Проведен анализ эффективности резервированного обслуживания запросов, критичных к времени их суммарного поэтапного ожидания, в очередях узлов многоуровневого кластера. Выполнен поиск проектных решений по организации многоэтапного резервированного обслуживания копий запросов в многоуровневом кластере, позволяющей повысить вероятность своевременного обслуживания запросов, критичных к суммарному времени ожидания на всех уровнях (этапов обслуживания) системы. **Метод исследования.** Построены модели многоэтапного резервированного обслуживания копий запросов в многоуровневом кластере. Эффективность резервирования оценивается по вероятности непревышения суммарного поэтапного времени ожидания в очередях узлов всех уровней заданного предельного допустимого времени. Проведен выбор наилучших вариантов организации резервированного обслуживания запросов, при этом запрос считается выполненным успешно, если хотя бы для одной безошибочно выполненной его копии суммарное время этапов ожидания в очередях узлов всех уровней кластера не превосходит предельно допустимое значение ожидания. **Полученные результаты.** Выполненные на основе расчетов эксперименты с использованием предложенных моделей позволили на примере трехуровневого кластера установить зависимости вероятности непревышения допустимого времени суммарного поэтапного ожидания запроса от интенсивности запросов, кратности резервирования и значения предельно допустимого суммарного времени ожидания запросов в очередях узлов всех уровней кластера. Показано существование области эффективного резервированного обслуживания запросов в многоуровневых кластерах и оптимальной кратности резервирования копий запросов в зависимости от загрузки системы и ограничений на допустимое суммарное время поэтапного ожидания в очередях узлов кластеров всех уровней. **Практическая значимость.** Предложенные модели могут быть использованы при оценке надежности и своевременности резервированных процессов обслуживания, критичных к ожиданию запросов, а также при обосновании выбора дисциплин и параметров обслуживания в многоуровневых компьютерных системах, в том числе реального времени.

Ключевые слова

многоуровневый кластер, вероятность своевременного обслуживания, резервированное обслуживание, кратность резервирования, система массового обслуживания, предельно допустимое время ожидания

MULTI-STEP SERVICE OF REQUESTS CRITICAL TO QUEUEING DELAYS IN MULTI-TIER SYSTEMS

V.A. Bogatyrev^a, S.V. Bogatyrev^b

^a ITMO University, Saint Petersburg, 197101, Russian Federation

^b Samsung Electronics, Seoul, 137-965, South Korea

Corresponding author: yvsemyashkina@mail.ru

Article info

Received 27.07.17, accepted 22.08.17

doi: 10.17586/2226-1494-2017-17-5-872-878

Article in Russian

For citation: Bogatyrev V.A., Bogatyrev S.V. Multi-step service of requests critical to queueing delays in multi-tier systems. *Scientific and Technical Journal of Information Technologies, Mechanics and Optics*, 2017, vol. 17, no. 5, pp. 872–878 (in Russian). doi: 10.17586/2226-1494-2017-17-5-872-878

Abstract

Subject of Research. The paper deals with effectiveness analysis of redundant service of requests time-critical to total gradual waiting in the queues of multi-tier clustering nodes. We search for design solutions on the organization of multi-stage redundant service of copies of requests in multi-tier cluster that makes it possible to increase the probability of timely servicing of requests critical to the total waiting time for all system levels (stages of maintenance). **Method.** We modeled the multi-stage redundant service of copies of requests in multi-tier cluster. The redundancy effectiveness is estimated by the probability of not exceeding the total phased waiting time in queues of nodes for all levels specified for the maximum allowable time. We selected the best options for the organization of redundant requests; the request is considered successful if at least for one accurately executed copy, total waiting time of the stages in the queues of the nodes at all levels of the cluster does not exceed the maximum allowable waiting value. **Main Results.** The experiments based on calculations carried out with the use of the proposed models on the three-tier cluster example gave the possibility to elucidate the dependence of probability of not exceeding the allowable total request time on the intensity of requests, the multiplicity of redundancy and the maximum allowable cumulative waiting time of requests in the queues of the nodes at all cluster levels. We have shown the existence of efficient redundant service of requests in multi-tier clusters, and the optimal multiplicity of redundancy of request copies depending on the system load and constraints on the allowable total time of staged waiting in the queues of the nodes in the clusters at all levels. **Practical Relevance.** The proposed models can be used to assess the reliability and timeliness of the service processes that are critical to request waiting time, as well as in the justification of selection of disciplines and service parameters in multi-tier computer systems, including real-time ones.

Keywords

multi-tier cluster, timely service probability, redundant service, multiplicity of redundancy, queuing system, maximum allowable waiting time

Введение

Современные центры хранения и обработки данных строятся на основе многоуровневой архитектуры при резервировании и консолидации ресурсов каждого уровня, что позволяет обеспечить высокую надежность [1–3], безопасность [4–6], доступность и производительность структуры системы [7–11], а также организуемых в ней вычислительных процессов, в том числе процессов реального времени, критичных к задержке выдачи результатов. Для систем, критичных к задержкам выдачи результатов, надежность вычислительного процесса характеризуется не только вероятностью работоспособности структуры во время выполнения требуемых функций, но и вероятностью их реализации к заданному моменту времени при обеспечении, в ряде случаев, непрерывности вычислительного процесса.

Оценка вероятности не превышения времени ожидания заданного предельно допустимого значения для систем массового обслуживания (СМО) типа М/М/1 (классификация Кендалла [12]) может быть получена на основе известной формулы для стационарного распределения времени ожидания [12]. Модификация этой формулы для одноуровневого кластера, представляемого группой одноканальных СМО типа М/М/1, предложена в [13]. Оценка вероятности не превышения времени ожидания предельно допустимого значения в двухуровневых компьютерных системах кластерной архитектуры, узлы которых представляются одноканальными СМО, предложена в [14].

Для систем, критичных к задержкам ожидания запросов, вероятность своевременного получения результата может быть повышена при резервированном обслуживании копий поступающих запросов [15]. Однако при резервировании вычислений возникает техническое противоречие, связанное с увеличением загрузки СМО из-за копирования запросов, что обуславливает необходимость поиска области эффективности резервированного обслуживания [15], определяемой в зависимости от дисциплины обслуживания с учетом длины очереди [15–17] или интенсивности входного потока запросов. Так, для дисциплины, названной в [16, 17] «широковещательное обслуживание с копированием запроса», при поступлении запроса в многоканальную СМО резервированное обслуживание предлагается проводить только для запросов, поступающих при пустой общей очереди, что, как установлено в [16, 17], позволяет сократить среднее время ожидания запросов.

Для повышения надежности и своевременности обслуживания всех критичных к задержкам ожидания запросов в одноуровневом компьютерном кластере, представляемом группой одноканальных СМО типа М/М/1 с бесконечными очередями, в [15] предлагается дисциплина с копированием и резервированным обслуживанием всех запросов, вне зависимости от длины очереди в момент их поступления. Эффективность резервированного обслуживания копий запросов, критичных к времени ожидания, как показано в [15, 18–20], зависит от дисциплины уничтожения неактуальных копий запросов, т.е. копий, находящихся в очередях после обслуживания хотя бы одной из них в других СМО или при истечении допустимого времени ожидания [18–20]. Резервирование процессов обслуживания запросов, критичных к времени ожидания, может использоваться не только в кластерных системах компьютеров, но и в системах резервированной передачи данных [21–24].

В предлагаемой работе впервые ставится задача построения модели для оценки вероятности своевременного обслуживания в многоуровневых кластерных системах (начиная с трехуровневых), в том числе с резервированием ресурсов на каждом уровне. Модель ориентирована на исследование эффективности резервированного обслуживания копий запросов, критичных к времени их суммарного поэтапного

ожидания, в очередях узлов всех уровней. Отличие предлагаемой модели заключается в возможности учета при оценке вероятности своевременного выполнения запросов, многоэтапности их обслуживания и поэтапного накопления времени ожидания в очередях одноканальных СМО, реализующих этапы обслуживания на разных уровнях кластера.

Цель исследования – поиск проектных решений по организации многоэтапного обслуживания запросов (в том числе, возможно, резервированного) в многоуровневом кластере, позволяющих повысить вероятность своевременного обслуживания запросов, критичных к суммарному времени ожидания, на всех уровнях (этапов) системы.

Объект исследования

В качестве объекта исследования рассмотрим многоуровневый компьютерный кластер, имеющий три или большее число уровней, причем узлы каждого уровня могут резервироваться с различной кратностью. В общем случае будем считать, что в кластере выделяется h уровней. В качестве уровней анализируемой системы могут рассматриваться кластеры, объединяющие WEB-серверы, серверы приложений, серверы баз данных, системы хранения и др. Поступающие на каждый уровень запросы сбалансированно распределяются между узлами кластера соответствующего уровня. Для каждого поступающего в многоуровневый кластер запроса при необходимости могут создаваться копии, каждая из которых обслуживается сначала одним из n_1 узлов первого уровня, затем одним из n_2 узлов второго уровня и т.д. Различные копии на каждом уровне кластера обслуживаются разными узлами.

В качестве критерия эффективности функционирования многоуровневого кластера, критичного к времени ожидания запросов, будем рассматривать вероятность непревышения суммарного поэтапного времени ожидания в очередях узлов заданного предельного допустимого времени t_0 .

При резервированном обслуживании запрос считается выполненным успешно, если хотя бы для одной безошибочно выполненной его копии суммарное время этапов ожидания в очередях узлов всех уровней кластера не превосходит предельно допустимое значение ожидания t_0 .

Модель трехуровневой системы

При построении модели будем предполагать, что при поступлении запроса формируются k его копий, которые поэтапно обслуживаются сначала k узлами кластера первого, затем – второго уровня и так далее до верхнего уровня. Будем считать, что в кластерах первого, второго и третьего уровней число узлов равно n_1, n_2, n_3 , а средние времена выполнения в них копии запроса соответственно равны v_1, v_2, v_3 .

Каждый узел кластера представим системой массового обслуживания типа М/М/1 с бесконечной очередью [12].

Предельно допустимое суммарное время ожидания запросов в очередях узлов всех уровней кластера t_0 разобьем на N одинаковых интервалов, длительности которых равны $t=t_0/N$. Выделим моменты начала отсчета интервалов ожидания $i=0, 1, \dots, N-1$. В кластере первого уровня отсчитываются интервалы $i_1=0, 1, \dots, N-1$, второго уровня – интервалы $i_2=0, 1, \dots, N-1-i_1$, третьего уровня – интервалы $i_3=0, 1, \dots, N-1-i_1-i_2$ и т.д. Остаточное время допустимого ожидания в очереди узлов h -го уровня кластера, таким образом, вычисляется как $t_h = t_0 - (i_1 + i_2 + \dots + i_{h-1}) (t_0/N)$.

Для трехуровневого кластера ($h=3$) вероятность непревышения суммарного времени поэтапного ожидания в очередях всех уровней предельно допустимого ожидания t_0 вычислим как

$$P_3 = \sum_{i_1=0}^{N-1} \sum_{i_2=0}^{N-1-i_1} P_{i_1} P_{i_2} P_{i_1, i_2},$$

где вероятность того, что время ожидания запросов в узле кластера первого уровня соответствует i_1 -му ($i_1=1, 2, \dots, N-1$) интервалу ожидания, отсчитываемому от момента времени $(i_1-1)(t_0/N)$ до момента времени $i_1(t_0/N)$, определяется как

$$P_{i_1} = 1 - \frac{\Lambda}{n_1} v_1 \exp\left(-\frac{i_1 t_0}{N} \left(v_1^{-1} - \frac{\Lambda}{n_1}\right) - b_{i_1}\right),$$

$$b_{i_1} = \begin{cases} 1 - \frac{\Lambda}{n_1} v_1 \exp\left(-\frac{t_0 (i_1 - 1)}{N} \left(v_1^{-1} - \frac{\Lambda}{n_1}\right)\right), & \text{if } i_1 \geq 1, \\ 0, & \text{if } i_1 = 0 \end{cases}$$

где Λ – интенсивность суммарного потока поступающих в кластер запросов.

Вероятность того, что время ожидания запросов в узле кластера второго уровня принадлежит i_2 -му ($i_2=1, 2, \dots, N-1-i_1$) интервалу ожидания, отсчитываемому от момента времени $(i_2-1)(t_0/N)$ до $i_2(t_0/N)$, вычисляется как

$$P_{i_2} = 1 - \frac{\Lambda}{n_2} v_2 \exp\left(-\frac{i_2 t_0}{N} \left(v_2^{-1} - \frac{\Lambda}{n_2}\right)\right) - b_{i_2}, \quad b_{i_2} = \begin{cases} 1 - \frac{\Lambda}{n_2} v_2 \exp\left(-\frac{t_0 (i_2 - 1)}{N} \left(v_2^{-1} - \frac{\Lambda}{n_2}\right)\right), & \text{if } i_2 \geq 1, \\ 0, & \text{if } i_2 = 0 \end{cases}$$

Случаи, когда $i_1=0$ или $i_2=0$, отражают поступление запроса (или его копии при резервированном обслуживании) соответственно в узел кластера первого или второго уровня при его незанятости, вероятность указанных событий определяется как

$$P_{i_1} = 1 - \frac{\Lambda}{n_1} v_1, \quad P_{i_2} = 1 - \frac{\Lambda}{n_2} v_2.$$

Вероятность того, что время ожидания запросов в узле кластера третьего уровня не превосходит остаточное время ожидания, вычисляется как

$$P_{i_1, i_2} = \left(1 - \frac{\Lambda}{n_3} v_3 \exp\left(-\left(t_0 - (i_1 + i_2) \frac{t_0}{N}\right) \left(\frac{1}{v_3} - \frac{\Lambda}{n_3}\right)\right)\right).$$

При интенсивности потока запросов Λ_0 и создании k копий каждого запроса (k – кратность резервирования запросов) при его резервированном обслуживании суммарная интенсивность копий запросов, поступающих в кластер (до распределения по узлам) каждого уровня, определяется как $\Lambda = k \Lambda_0$, в частности при нерезервированном обслуживании $\Lambda = \Lambda_0$. При независимом выполнении каждой копии в узлах кластера вероятность R не превышения суммарного ожидания на трех этапах обслуживания хотя бы для одной из k копий допустимого предела t_0 вычислим как

$$R_3 = 1 - (1 - P_3)^k.$$

Таким образом, определена вероятность своевременного обслуживания в трехуровневом кластере хотя бы одной из k создаваемых копий в предположении безотказности и безошибочности ресурсов, задействованных при выполнении копии запроса.

Модель h -уровневой системы

Для h -уровневого кластера вероятность не превышения суммарного времени поэтапного ожидания в очередях всех уровней установленного предела t_0 определим как

$$P_h = \sum_{i_1=0}^{N-1} \sum_{i_2=0}^{N-1-i_1} \dots \sum_{i_j=0}^{N-1-i_1-i_2-\dots-i_{j-1}} \dots \sum_{i_{h-1}=0}^{N-1-i_1-i_2-\dots-i_{j-1}-\dots-i_{h-1}} P_{i_1} P_{i_2} \dots P_{i_j} \dots P_{i_{h-1}} P_{i_1, i_2, \dots, i_{h-1}},$$

где вероятности $P_{i_1}, P_{i_2}, \dots, P_{i_j}, \dots, P_{i_{h-1}}$ вычисляются как

$$P_{i_j} = 1 - \frac{\Lambda}{n_j} v_j \exp\left(-\frac{i_j t_0}{N} \left(v_j^{-1} - \frac{\Lambda}{n_j}\right)\right) - b_{i_j}, \quad b_{i_j} = \begin{cases} 1 - \frac{\Lambda}{n_j} v_j \exp\left(-\frac{t_0 (i_j - 1)}{N} \left(v_j^{-1} - \frac{\Lambda}{n_j}\right)\right), & \text{if } i_j \geq 1, \\ 0, & \text{if } i_j = 0 \end{cases}$$

при $j=1, 2, \dots, (h-1)$. Вероятность того, что время ожидания запросов в узле кластера h уровня не превосходит остаточное время ожидания после прохождения запроса с первого по $(h-1)$ -й этапов обслуживания вычисляется как

$$P_{i_1, i_2, \dots, i_{h-1}} = \left(1 - \frac{\Lambda}{n_h} v_h \exp\left(-\left(t_0 - \frac{t_0}{N} \sum_{s=1}^{h-1} i_s\right) \left(\frac{1}{v_h} - \frac{\Lambda}{n_h}\right)\right)\right).$$

При независимом выполнении каждой копии в узлах кластера внутри каждого уровня вероятность R_h не превышения суммарного ожидания на h этапах обслуживания хотя бы для одной из k копий допустимого предела t_0 в предположении безотказности узлов системы и безошибочности в них вычислений определим как:

$$R_h = 1 - (1 - P_h)^k.$$

Если интенсивность отказов узлов и ошибок, возникающих во время ожидания, равна λ_0 , а во время вычислений – λ_v , то вероятность своевременного и безошибочного h -этапного обслуживания хотя бы одной из k резервных копии запроса в h -уровневом кластере определим как

$$R_h = 1 - \left(1 - P_h e^{-(\lambda_0 t_0 + \lambda_v \sum_{i=1}^h v_i)}\right)^k.$$

Результаты расчетов

При расчетах будем считать, что число уровней кластера $h=3$, число узлов на уровнях $n_1 = n_2 = n_3=10$, а средние времена поэтапного выполнения запросов $v_1 = v_2 = v_3 = 0,3$ с. Расчет проведем для идеального случая безошибочности вычислений при безотказности узлов.

На рис. 1 представлена зависимость вероятности P не превышения допустимого времени t_0 суммарного поэтапного ожидания запроса в трехуровневом кластере от интенсивности запросов Λ_0 при поэтапном нерезервированном (рис. 1, а) и резервированном (рис. 1, б) обслуживании запросов. На рис. 1, а, кривые 1–3 отражают допустимое время ожидания $t_0= 0,5; 1; 2$ с соответственно.

На рис. 1, б, кривые 1–3 соответствуют вероятности не превышения допустимого времени суммарного ожидания запроса $t_0=1$ с при кратности резервированного обслуживания запросов $k=1, 2, 3$; кривые 4–6 – вероятности не превышения допустимого времени суммарного ожидания запроса при кратности резервированного обслуживания $k=3$ для $t_0= 0,6; 0,8; 1,2$ с.

Из представленных зависимостей видно, что эффективность резервированного обслуживания при различной его кратности зависит от ограничений на предельно допустимое ожидание t_0 и от загруженности узлов, при этом чем меньше загруженность (интенсивность запросов), тем большая кратность резервирования целесообразна. Таким образом, существует зависимость оптимальной кратности резервирования от загруженности узлов системы и от предельно допустимого суммарного времени поэтапного ожидания на всех этапах обслуживания.

На рис. 2 приведена зависимость вероятности своевременности трехэтапного обслуживания R от его кратности резервирования (числа формируемых копий запросов) k для различных значениях интенсивностей потока запросов (рис. 2, а) и допустимого суммарного времени ожидания t_0 (рис. 2, б). На рис. 2, а, кривые 1–4 соответствуют значениям $\Lambda = 5; 4; 3, 2$ 1/с при $t_0=1$ с, а на рис. 2, б, кривые 1–5 – значениям $t_0 = 0,4; 0,6; 0,8; 1; 1,5$ с при $\Lambda = 5$ 1/с.

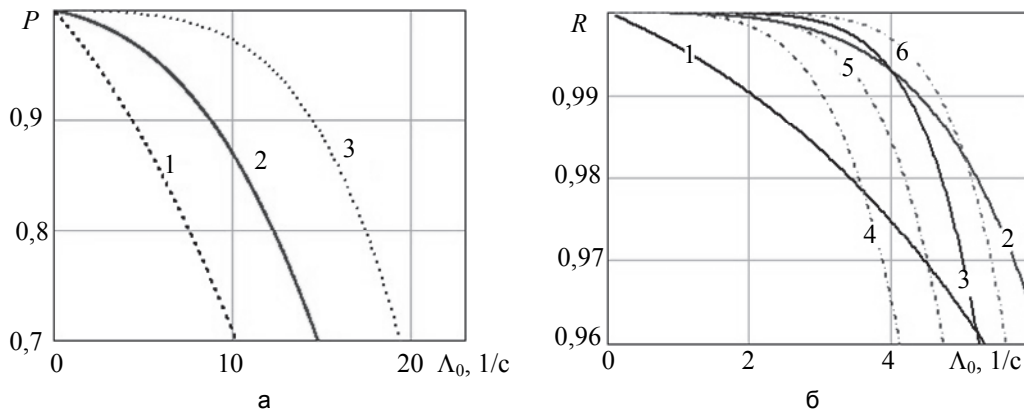


Рис. 1. Вероятность P не превышения допустимого времени t_0 суммарного ожидания запроса в трехуровневом кластере при: нерезервированном обслуживании (а); резервированном обслуживании запросов (б)

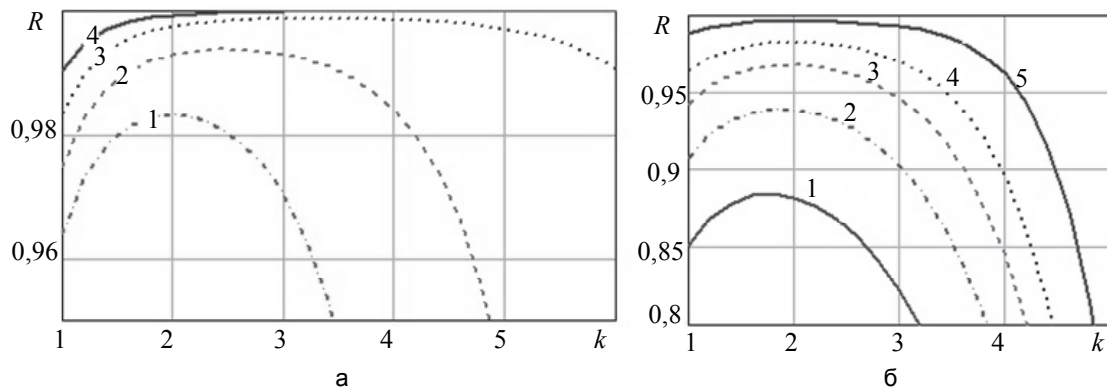


Рис. 2. Зависимость вероятности своевременности обслуживания R от кратности его резервирования k для различных значениях интенсивностей потока запросов (а) и допустимого суммарного времени ожидания t_0 (б)

Приведенные графики позволяют сделать вывод о существовании оптимальной кратности резервированного обслуживания, причем, чем меньше загруженность системы (интенсивность потока запро-

сов) и меньше допустимое суммарное время ожидания на всех этапах обслуживания, тем больше оптимальная кратность резервирования, при которой достигается максимум вероятности своевременности многоэтапного обслуживания. Следует заметить, что оптимальная кратность резервированного обслуживания запросов может быть не обязательно целой – ее нецелое значение соответствует смешанной стратегии обслуживания, при которой с некоторыми вероятностями выбираются целые значения кратностей из заданного ряда.

Заключение

Для многоуровневых компьютерных систем, предполагающих поэтапное выполнение запросов в узлах кластеров всех уровней, предложена аналитическая модель и определена эффективность резервированного обслуживания запросов, критичных к суммарной задержке поэтапного ожидания, в очередях узлов кластеров разных уровней.

Проанализировано влияние кратности резервированного обслуживания в трехуровневой вычислительной системе кластерной архитектуры на вероятность своевременного выполнения запросов с учетом их поэтапного резервированного выполнения в кластерах всех уровней.

Показано существование области эффективного резервированного обслуживания копий запросов в многоуровневых кластерах и оптимальной кратности резервирования копий запросов в зависимости от загрузки системы и ограничений на допустимое суммарное время поэтапного ожидания в очередях узлов кластеров разных уровней.

Предложенные модели могут быть использованы при оценке надежности и своевременности функционирования резервированных процессов обслуживания критичных к ожиданию запросов и при обосновании выбора дисциплин и параметров обслуживания в многоуровневых компьютерных системах, в том числе реального времени.

Литература

1. Sorin D.J. *Fault Tolerant Computer Architecture*. Morgan & Claypool, 2009. 103 p.
2. Koren I., Krishna C.M. *Fault Tolerant Systems*. San Francisco: Morgan Kaufmann Publishers, 2009. 378 p.
3. Перегуда А.И., Тимашов Д.А. Математическая модель надежности локальной вычислительной сети // Информационные технологии. 2008. № 10. С. 7–15.
4. Новиков С.Н. Методология защиты пользовательской информации на основе технологий сетевого уровня мультисервисных сетей связи / под ред. В.П. Шувалова. М.: Горячая линия - Телеком, 2015. 128 с.
5. Советов Б.Я., Колбанёв М.О., Татарникова Т.М. Технологии инфокоммуникации и их роль в обеспечении информационной безопасности // Геополитика и безопасность. 2014. №1(25). С. 69–77.
6. Верзун Н.А., Колбанев М.О., Татарникова Т.М. Технологическая платформа четвертой промышленной революции // Геополитика и безопасность. 2016. № 2 (34). С. 73–78.
7. Aliev T.I. The synthesis of service discipline in systems with limits // *Communications in Computer and Information Science*. 2016. V. 601. P. 151–156. doi: 10.1007/978-3-319-30843-2_16
8. Aliev T.I., Rebezova M.I., Russ A.A. Statistical methods for monitoring travel agencies // *Automatic Control and Computer Sciences*. 2015. V. 49. N 6. P. 321–327. doi: 10.3103/S0146411615060024
9. Aleksanin S.A., Zharinov I.O., Korobeynikov A.G., Perezyabov O.A., Zharinov O.O. Evaluation of chromaticity coordinate shifts for visually perceived image in terms of exposure to external illuminance // *ARPN Journal of Engineering and Applied Sciences*. 2015. V.10. N 17. P. 7494–7501.
10. Bogatyrev V.A. Protocols for dynamic distribution of requests through a bus with variable logic ring for reception authority transfer // *Automatic Control and Computer Sciences*. V. 33. N 1. P. 57–63.
11. Korobeynikov A.G., Fedosovsky M.E., Maltseva N.K., Baranova O.V., Zharinov I.O., Gurjanov A.V., Zharinov O.O. Use of information technologies in design and production activities of instrument-making plants // *Indian Journal of Science and Technology*. 2016. V. 9. N 44. doi: 10.17485/ijst/2016/v9i44/104708

References

1. Sorin D.J. *Fault Tolerant Computer Architecture*. Morgan & Claypool, 2009, 103 p.
2. Koren I., Krishna C.M. *Fault Tolerant Systems*. San Francisco, Morgan Kaufmann Publishers, 2009, 378 p.
3. Pereguda A.I., Timashov D.A. The mathematical model of reliability of local area network (LAN). *Informacionnye Tehnologii*, 2008, no. 10, pp. 7–15. (In Russian)
4. Novikov S.N. *Methodology of Protection of the User Information on the Basis of Technologies of Network Level of Multiservice Communication Networks*. Ed. V.P. Shuvalov. Moscow, Goryachaya Liniya – Telekom Publ., 2015, 128 p. (In Russian)
5. Sovetov B.Ya., Kolbanev M.O., Tatarnikova T.M. Infocommunication technologies and their role in information security. *Geopolitics and Security*, 2014, no. 1, pp. 69–77. (In Russian)
6. Verzun N.A., Kolbanev M.O., Tatarnikova T.M. The technological platform of the fourth industrial revolution. *Geopolitics and Security*, 2016, no. 2, pp. 73–78. (In Russian)
7. Aliev T.I. The synthesis of service discipline in systems with limits. *Communications in Computer and Information Science*, 2016, vol. 601, pp. 151–156. doi: 10.1007/978-3-319-30843-2_16
8. Aliev T.I., Rebezova M.I., Russ A.A. Statistical methods for monitoring travel agencies. *Automatic Control and Computer Sciences*, 2015, vol. 49, no. 6, pp. 321–327. doi: 10.3103/S0146411615060024
9. Aleksanin S.A., Zharinov I.O., Korobeynikov A.G., Perezyabov O.A., Zharinov O.O. Evaluation of chromaticity coordinate shifts for visually perceived image in terms of exposure to external illuminance. *ARPN Journal of Engineering and Applied Sciences*, 2015, vol. 10, no. 17, pp. 7494–7501.
10. Bogatyrev V.A. Protocols for dynamic distribution of requests through a bus with variable logic ring for reception authority transfer. *Automatic Control and Computer Sciences*, 1999, vol. 33, no. 1, pp. 57–63.
11. Korobeynikov A.G., Fedosovsky M.E., Maltseva N.K., Baranova O.V., Zharinov I.O., Gurjanov A.V., Zharinov O.O. Use of information technologies in design and production activities of instrument-making plants. *Indian Journal of Science and Technology*, 2016, vol. 9, no. 44.

12. Вишнеvский В.М. Теоретические основы проектирования компьютерных сетей. М.: Техносфера, 2003. 512 с.
13. Богатырев В.А., Богатырев А.В. Функциональная надежность систем реального времени // Научно-технический вестник информационных технологий, механики и оптики. 2013. № 4 (86). С. 150–151.
14. Богатырев В.А., Богатырев А.В., Богатырев С.В. Оценка своевременности выполнения критических запросов в двухуровневых кластерах // Научно-технический вестник информационных технологий, механики и оптики. 2014. № 2 (90). С. 177–179.
15. Bogatyrev V.A., Bogatyrev A.V. Functional reliability of a real-time redundant computational process in cluster architecture systems // Automatic Control and Computer Sciences. 2015. V. 49. N 1. P. 46–56. doi: 10.3103/S0146411615010022
16. Dudin A.N., Sun' B. A multiserver MAP/PH/N system with controlled broadcasting by unreliable servers // Automatic Control and Computer Sciences. 2009. V. 43. N 5. P. 247–256. doi: 10.3103/S0146411609050046
17. Lee M.H., Dudin A.N., Klimenok V.I. The SM/V/N queueing system with broadcasting service // Mathematical Problem in Engineering. 2006. V. 2006. Art. 98171. doi: 10.1155/MPE/2006/98171
18. Богатырев В.А., Богатырев А.В. Модель резервированного обслуживания запросов реального времени в компьютерном кластере // Информационные технологии. 2016. Т. 22. № 5. С. 348–355.
19. Богатырев В.А., Богатырев А.В. Надежность функционирования кластерных систем реального времени с фрагментацией и резервированным обслуживанием запросов // Информационные технологии. 2016. Т. 22. № 6. С. 409–416.
20. Богатырев В.А., Богатырев С.В. Резервированное обслуживание в кластерах с уничтожением неактуальных запросов // Вестник компьютерных и информационных технологий. 2017. № 1(151). С. 21–28.
21. Богатырев В.А., Богатырев С.В. Эффективность резервирования и фрагментации пакетов при передаче по агрегированному каналу // Известия высших учебных заведений. Приборостроение. 2017. Т. 60. № 2. С. 165–170. doi: 10.17586/0021-3454-2017-60-2-165-170
22. Богатырев В.А., Богатырев С.В. Резервированная передача данных через агрегированные каналы в сети реального времени // Известия высших учебных заведений. Приборостроение. 2016. Т. 59. № 9. С. 735–740. doi: 10.17586/0021-3454-2016-59-9-735-740
23. Bogatyrev V.A., Parshutina S.A., Poptcova N.A., Bogatyrev A.V. Efficiency of redundant service with destruction of expired and irrelevant request copies in real-time clusters // Communications in Computer and Information Science. 2016. V. 678. P. 337–348. doi: 10.1007/978-3-319-51917-3_30
24. Bogatyrev V.A., Parshutina S.A. Redundant distribution of requests through the network by transferring them over multiple paths // Communications in Computer and Information Science. 2016. V. 601. P. 199–207. doi: 10.1007/978-3-319-30843-2_21
- doi: 10.17485/ijst/2016/v9i44/104708
12. Vishnevskii V.M. *Theoretical Foundations of Computer Network Design*. Moscow, Tekhnosfera Publ., 2003, 512 p.
13. Bogatyrev V.A., Bogatyrev A.V. Functional reliability of real-time systems. *Scientific and Technical Journal of Information Technologies, Mechanics and Optics*, 2013, no. 4, pp. 150–151. (In Russian)
14. Bogatyrev V.A., Bogatyrev A.V., Bogatyrev S.V. Timeliness assessment for implementation of the critical requests in two-level clusters. *Scientific and Technical Journal of Information Technologies, Mechanics and Optics*, 2014, no. 2, pp. 177–179. (In Russian)
15. Bogatyrev V.A., Bogatyrev A.V. Functional reliability of a real-time redundant computational process in cluster architecture systems. *Automatic Control and Computer Sciences*, 2015, vol. 49, no. 1, pp. 46–56. doi: 10.3103/S0146411615010022
16. Dudin A.N., Sun' B. A multiserver MAP/PH/N system with controlled broadcasting by unreliable servers. *Automatic Control and Computer Sciences*, 2009, vol. 43, no. 5, pp. 247–256. doi: 10.3103/S0146411609050046
17. Lee M.H., Dudin A.N., Klimenok V.I. The SM/V/N queueing system with broadcasting service. *Mathematical Problem in Engineering*, 2006, vol. 2006, art. 98171. doi: 10.1155/MPE/2006/98171
18. Bogatyrev V.A., Bogatyrev A.V. The model of redundant service requests real-time in a computer cluster. *Informacionnye Tehnologii*, 2016, vol. 22, no. 5, pp. 348–355. (In Russian)
19. Bogatyrev V.A., Bogatyrev A.V. The reliability of the cluster real-time systems with fragmentation and redundant service requests. *Informacionnye Tehnologii*, 2016, vol. 22, no. 6, pp. 409–416. (In Russian)
20. Bogatyrev V.A., Bogatyrev S.V. Redundant service clusters with the destruction of irrelevant queries. *Herald of Computer and Information Technologies*, 2017, no. 1, pp. 21–28. (In Russian)
21. Bogatyrev V.A., Bogatyrev S.V. Effectiveness of redundancy and packet fragmentation in transmission via aggregated channels. *Journal of Instrument Engineering*, 2017, vol. 60, no. 2, pp. 165–170. (In Russian) doi: 10.17586/0021-3454-2017-60-2-165-170
22. Bogatyrev V.A., Bogatyrev S.V. Redundant data transmission using aggregated channels in real-time network. *Journal of Instrument Engineering*, 2016, vol. 59, no. 9, pp. 735–740. (In Russian) doi: 10.17586/0021-3454-2016-59-9-735-740
23. Bogatyrev V.A., Parshutina S.A., Poptcova N.A., Bogatyrev A.V. Efficiency of redundant service with destruction of expired and irrelevant request copies in real-time clusters. *Communications in Computer and Information Science*, 2016, vol. 678, pp. 337–348. doi: 10.1007/978-3-319-51917-3_30
24. Bogatyrev V.A., Parshutina S.A. Redundant distribution of requests through the network by transferring them over multiple paths. *Communications in Computer and Information Science*, 2016, vol. 601, pp. 199–207. doi: 10.1007/978-3-319-30843-2_21

Авторы

Богатырев Владимир Анатольевич – доктор технических наук, профессор, Университет ИТМО, Санкт-Петербург, 197101, Российская Федерация. Vladimir.bogatyrev@gmail.com

Богатырев Станислав Владимирович – старший инженер, Самсунг Электроникс, Сеул, 137-965, Южная Корея. realloc@gmail.com

Authors

Vladimir A. Bogatyrev – D.Sc., Professor, ITMO University, Saint Petersburg, 197101, Russian Federation. Vladimir.bogatyrev@gmail.com

Stanislav V. Bogatyrev – senior engineer, Samsung Electronics, Seoul, 137-965, South Korea. realloc@gmail.com