

doi: 10.17586/2226-1494-2025-25-5-988-995

УДК 004. 942

Оценка надежности восстанавливаемого кластера контейнерной виртуализацией

Владимир Анатольевич Богатырев^{1✉}, Ван Кю Фунг²

¹ Санкт-Петербургский государственный университет аэрокосмического приборостроения, Санкт-Петербург, 190000, Российская Федерация

^{1,2} Университет ИТМО, Санкт-Петербург, 197101, Российская Федерация

¹ vladimir.bogatyrev@gmail.com✉, <https://orcid.org/0000-0003-0213-0223>

² phungvanquy97@gmail.com, <https://orcid.org/0009-0006-3278-1106>

Аннотация

Введение. При построении отказоустойчивых кластеров высокой готовности с малыми задержками обслуживания потоков запросов все большее применение находит технология контейнерной виртуализации. При проектировании высоконадежных кластеров важной задачей является их структурно-параметрический модельно-ориентированный синтез с учетом влияния числа развертываемых контейнеров на производительность, задержки обслуживания потоков запросов и надежность системы. **Метод.** Обоснование решений по обеспечению высокой готовности кластера основывается на разработке моделей восстанавливаемого кластера при реконфигурации с учетом миграции виртуальных контейнеров. Новизна предлагаемых марковских моделей кластера состоит в учете двухэтапного восстановления его работоспособности с определением влияния на надежность кластера числа контейнеров, подлежащих миграции в процессе реконфигурации — до и после физического восстановления отказавших серверов. Рассмотрены два варианта миграции контейнеров при восстановлении кластера. В первом варианте, на этапе физического восстановления отказавшего сервера, миграция контейнеров на исправный сервер не происходит, а во втором — происходит. На втором этапе реконфигурации после физического восстановления отказавшего сервера осуществляется миграция контейнеров, при которой возможно как увеличение, так и уменьшение числа развернутых в них контейнеров. **Основные результаты.** На основе предлагаемых марковских моделей надежности кластера с контейнерной виртуализацией дана оценка его коэффициента готовности и определено влияние числа контейнеров, загружаемых при миграции на двух этапах реконфигурации, на надежность системы. **Обсуждение.** Предложенные марковские модели надежности кластера с контейнерной виртуализацией направлены на обоснование выбора проектных решений по организации и восстановлению работоспособности кластера после отказов серверов с учетом влияния вариантов реализации миграции виртуальных контейнеров на готовность системы. В дальнейших исследованиях предполагается анализ влияния вариантов миграции контейнеров как на готовность кластера, так и на задержки обслуживания запросов на двух рассматриваемых этапах реконфигурации.

Ключевые слова

отказоустойчивость, коэффициент готовности, контейнерная виртуализация, кластер, миграция контейнеров, марковская модель, надежность

Ссылка для цитирования: Богатырев В.А., Фунг В.К. Оценка надежности восстанавливаемого кластера контейнерной виртуализацией // Научно-технический вестник информационных технологий, механики и оптики. 2025. Т. 25, № 5. С. 988–995. doi: 10.17586/2226-1494-2025-25-5-988-995

Assessment of the reliability of a recoverable container virtualization cluster

Vladimir A. Bogatyrev¹✉, Van Quy Phung²

¹ Saint Petersburg State University of Aerospace Instrumentation (SUAI), Saint Petersburg, 190000, Russian Federation

^{1,2} ITMO University, Saint Petersburg, 197101, Russian Federation

¹ vladimir.bogatyrev@gmail.com✉, <https://orcid.org/0000-0003-0213-0223>

² phungvanquy97@gmail.com, <https://orcid.org/0009-0006-3278-1106>

Abstract

Container virtualization technology is increasingly being used in the development of fault-tolerant clusters with high availability and low request processing latency. In designing highly reliable clusters, a key task is the structural-parametric model-oriented synthesis which takes into account the impact of the number of deployed containers on performance, request processing latency, and system reliability. Justifying the choice of solutions to ensure high cluster reliability currently requires the development of reliability models for recoverable container virtualization clusters during reconfiguration, considering the migration of virtual containers. The basis for decisions to ensure high cluster availability is the development of models for a recoverable cluster during reconfiguration, taking into account the migration of virtual containers. The novelty of the proposed Markov model of a cluster lies in considering a two-stage recovery of its operability, determining the impact of the number of containers to be migrated during reconfiguration — both before and after the physical recovery of failed servers — on cluster reliability. Two options for container migration during cluster recovery are considered. In the first scenario, during the physical recovery phase of a failed server, container migration to a functional server does not occur, while in the second scenario it does. In the second stage of reconfiguration, following the physical recovery of a failed server, container migration takes place, allowing for either an increase or decrease in the number of containers deployed on them. Based on the proposed Markov models of cluster reliability with container virtualization, an evaluation of its readiness coefficient is provided, and the influence of the number of containers loaded during migration at the two reconfiguration stages on system reliability is determined. The proposed Markov models of cluster reliability with container virtualization are aimed at justifying design decisions for organizing and restoring cluster operability after server failures, considering the impact of container migration implementation options on system availability. Future research will analyze the impact of container migration options on both cluster availability and request processing latency at the two considered reconfiguration stages.

Keywords

fault tolerance, availability factor, container virtualization, cluster, container migration, Markov model, reliability

For citation: Bogatyrev V.A., Phung V.Q. Assessment of the reliability of a recoverable container virtualization cluster. *Scientific and Technical Journal of Information Technologies, Mechanics and Optics*, 2025, vol. 25, no. 5, pp. 988–995 (in Russian). doi: 10.17586/2226-1494-2025-25-5-988-995

Введение

Требования высокой надежности, доступности, отказоустойчивости и бесперебойности распределенных вычислений являются ключевыми для современных инфокоммуникационных систем [1–3].

Для распределенных компьютерных систем [4, 5] высокая доступность, надежность и отказоустойчивость при непрерывности вычислительных процессов и малых задержках обслуживания потока запросов достигаются при консолидации вычислительных ресурсов в кластерах, что предполагает их реконфигурацию после отказов с использованием репликации и динамической миграции виртуальных машин [6, 7] и контейнеров [8–10].

Построение отказоустойчивых кластеров требует обоснования и оптимизации решений, обеспечивающих надежность, производительность и высокоскоростной надежный доступ [11]. Эти решения должны быть взаимоувязаны на всех уровнях распределенной системы, включая резервированные средства передачи [12], обработки и хранения данных [13].

Выбор решений по организации высокоскоростного доступа и надежности сети может опираться на методы многопутевой маршрутизации и резервированных передач [14–16].

Обоснование и структурно-параметрическая оптимизация построения и эксплуатации кластера с целью

обеспечения наибольшей надежности при малых задержках запросов в очередях должны опираться на расчеты надежности и задержек обслуживания запросов с учетом реконфигурации после отказов или изменений трафика. Известные марковские модели восстанавливаемых резервированных систем параллельной структуры не в полной мере отражают особенности функционирования кластеров, в том числе комплексное влияние на надежность и производительность кластера, организацию контроля, а также устройств обработки, хранения и внутрикластерного обмена [17–21].

Влияние организации тестового и оперативного контролей двухузлового кластера учитывается моделью, предложенной в работе [22]. Модель кластера, компонуемого из дублированных компьютерных узлов [23], отражает влияние на надежность и задержки обслуживания, связанное с организацией обмена внутри кластера, но не учитывает влияние различных дисциплин восстановления после отказов. Указанные модели не учитывают двухэтапность восстановления кластера после отказов, включающей сначала физическое, а затем информационное восстановление. Модель, отражающая двухэтапность восстановления узлов в дублированном кластере, объединяющем два компьютерных узла и два узла двухвходовой памяти при их полной связанности, предложена в [24] для различных вариантов разделения ресурсов процессорных узлов на решение функциональных задач и информационное восстановление памяти.

Представленные в работах [6, 7, 25] модели кластерных систем не отражают особенности технологии виртуальных машин, в том числе вопросы репликации и динамической миграции виртуальных машин при восстановлении физических серверов после их отказов [9, 26].

Модель двухузловой кластерной системы при репликации и динамической миграции виртуальных машин [27], а также при реконфигурации кластера после отказов его физических серверов, описанная в [28], не раскрывает особенностей реконфигурации кластера при контейнерной виртуализации.

В настоящее время существует потребность в разработке моделей кластеров контейнерной виртуализации, показывающих влияние реконфигурации системы при отказе серверов с их поэтапным физическим и информационным восстановлением, предполагающим перераспределение трафика и миграцию виртуальных контейнеров. Эта потребность обусловлена необходимостью обоснования и оптимизации решений по построению кластера и его реконфигурации после отказов.

Цель данной работы — разработка моделей оценки надежности восстанавливаемого кластера при реконфигурации с учетом миграции виртуальных контейнеров, ориентированных на обоснование выбора решений по обеспечению высокой надежности кластера.

Новизна предлагаемой модели состоит в учете двухэтапного восстановления работоспособности кластера с оценкой влияния на надежность системы числа контейнеров, подлежащих миграции на этапах во время и после физического восстановления отказавших серверов.

Структура кластера с контейнерной виртуализацией

Рассмотрим простейшую реализацию кластера, включающую два сервера с контейнерной виртуализацией и узел — балансировщик нагрузки, осуществляющий распределение запросов между серверами и участвующий в развертывании и миграции виртуальных контейнеров, в том числе при реконфигурации после

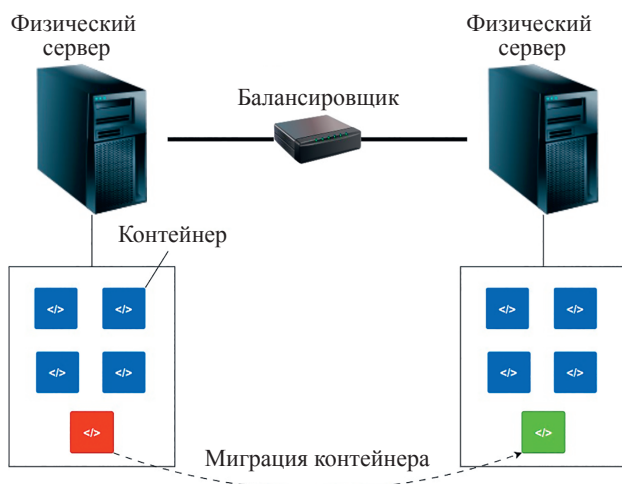


Рис. 1. Структура кластера
Fig. 1. Cluster Structure

отказов. Балансировщик, распределяющий запросы, представляется одноканальной системой массового обслуживания с неограниченной очередью (М/М/1) [29].

Физический сервер, реализующий обслуживание потока запросов, моделируется многоканальной системой массового обслуживания с неограниченной очередью, при этом каждый канал соответствует контейнеру. Интенсивность обслуживания запросов в каждом канале (контейнере) зависит от числа загруженных и активных контейнеров, т. е. задействованных в выполнении поступающих запросов. Такой подход позволяет учесть разделение ограниченных вычислительных ресурсов сервера между контейнерами, участвующими и не участвующими в обслуживании запросов. Особенностью рассматриваемых систем с контейнерной виртуализацией является существование оптимального числа контейнеров, развернутых на узле, при котором достигается минимум задержки обслуживания запросов [30]. Причем оптимальное число развернутых контейнеров зависит от интенсивности потока запросов. Это позволяет предположить, что при отказе одного из серверов и перераспределении потока запросов на исправный сервер число развернутых на нем контейнеров до восстановления отказавшего сервера должно изменяться таким образом, чтобы минимизировать задержки обслуживания потока запросов при одновременном увеличении коэффициента готовности кластера.

Такой подход отличается от традиционного тем, что позволяет учесть разделение ограниченных вычислительных ресурсов физического сервера между активными контейнерами.

Модель надежности кластера при реконфигурации с миграцией контейнеров

Построим марковскую модель надежности восстанавливаемого кластера с учетом этапов физического и информационного восстановления серверов, связанных с загрузкой (миграцией) виртуальных контейнеров на сервер после его физического восстановления. Информационное восстановление происходит под управлением узла-распределителя (балансировщика нагрузки) и может включать изменение числа развернутых в узлах контейнеров (миграцию контейнеров). Миграция виртуальных контейнеров осуществляется на двух этапах реконфигурации.

Этап 1. После отказа сервера до его физического восстановления.

Этап 2. После физического восстановления отказавшего сервера до его информационного восстановления, включающего загрузку заданного числа контейнеров.

Диаграммы состояний и переходов для двух рассматриваемых вариантов реконфигурации кластера представлены на рис. 2 и 3. Состояние кластера характеризуем двумя строками, верхняя из которых отображает состояние серверов, а нижняя — распределителя запросов. Распределитель рассматривается в двух состояниях: «1» — исправен, «X» — отказал. Для каждого сервера выделены состояния: исправность с указанием числа развернутых в нем контейнеров (n_0

или n_1); завершение физического восстановления до начала загрузки контейнеров — «0». При этом n_0 — число контейнеров в исходном состоянии (при работоспособности двух серверов), а n_1 — число контейнеров, загружаемых на сервер, сохранявший работоспособность после отказа второго сервера. Временем, связанным с уменьшением числа загруженных контейнеров, пренебрежем.

При этом считаем, что время загрузки контейнеров пропорционально их числу.

Рассмотрим два варианта дисциплины реконфигурации системы.

При варианте 1 реконфигурации, после отказа одного из двух серверов, вся нагрузка распределяется на исправный сервер, при этом не происходит изменение числа развернутых в нем контейнеров, т. е. на этапе 1 реконфигурации миграция контейнеров не реализуется (рис. 2). После физического восстановления отказавшего сервера осуществляется развертывание в нем первоначально заданного числа контейнеров, таким образом

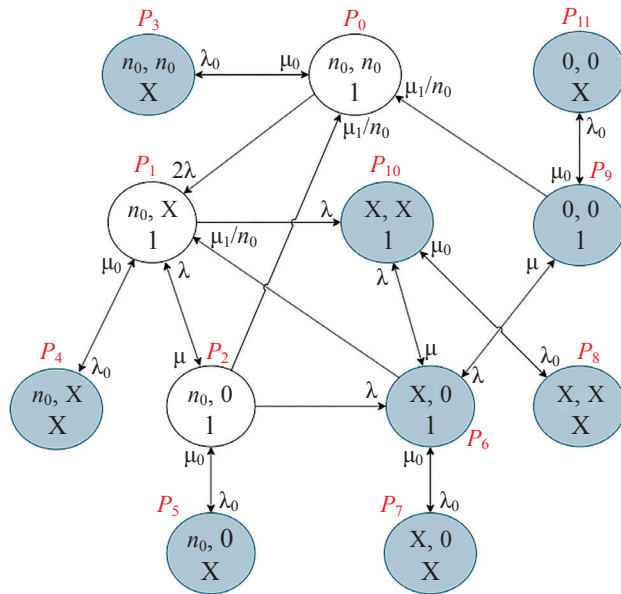


Рис. 2. Диаграмма состояний и переходов кластера без миграции контейнеров до этапа физического восстановления отказавшего сервера.

λ и λ_0 — интенсивности отказов физического сервера и балансировщика; μ_0 — интенсивность восстановления балансировщика; μ и μ_1 — интенсивности физического и информационного восстановлений серверов;

$P_0, P_1, P_2, \dots, P_{11}$ — вероятности состояний системы.

Узлы, отмеченные серым фоном — состояния системы, при которых процесс выполнения требуемых функций невозможен из-за неработоспособности необходимых для этого ресурсов

Fig. 2. State and transition diagram of a cluster without migration prior to the physical recovery stage of the failed server.

λ and λ_0 — failure rates of the physical server and balancer; μ_0 — balancer recovery rate; μ and μ_1 — physical and information recovery rates of servers; $P_0, P_1, P_2, \dots, P_{11}$ — system state probabilities.

Nodes marked with a gray background — system states in which the process of performing the required functions is impossible due to the inoperability of the resources required for this

на этапе 2 реконфигурации миграция контейнеров происходит. В результате система восстанавливается до исходного состояния.

При варианте 2 реконфигурации, после отказа одного из двух серверов, вся нагрузка распределяется на исправный сервер, но во время физического восстановления отказавшего сервера на работоспособном сервере под управлением балансировщика-распределителя нагрузки реализуется изменение числа (миграция) развернутых в нем контейнеров (рис. 3). На этапе 2 реконфигурации происходит изменение числа контейнеров на обоих серверах кластера, причем возможно как увеличение, так и уменьшение числа развернутых в них контейнеров.

Отметим, что при увеличении числа контейнеров на этапе 1 система может снижать [31] производительность сохранившего работоспособность сервера. Это связано с тем, что даже не принимающие участие в обслуживании запросов контейнеры при малой загрузке все равно потребляют определенную долю общих ресурсов, которые разделяются между активными и пассивными (простаивающими) контейнерами. Таким образом, чтобы минимизировать время пребывания запросов в системе, необходимо перенести лишь определенное число контейнеров, зависящее от интенсивности трафика. Анализ влияния числа развертываемых контейнеров на среднее время пребывания запросов в системе в зависимости от интенсивности потока трафика, определяющего число активных контейнеров, выполнен в работе [30].

При построении диаграмм состояний и переходов марковской модели предположим, что процесс физического восстановления узлов выполняется операто-

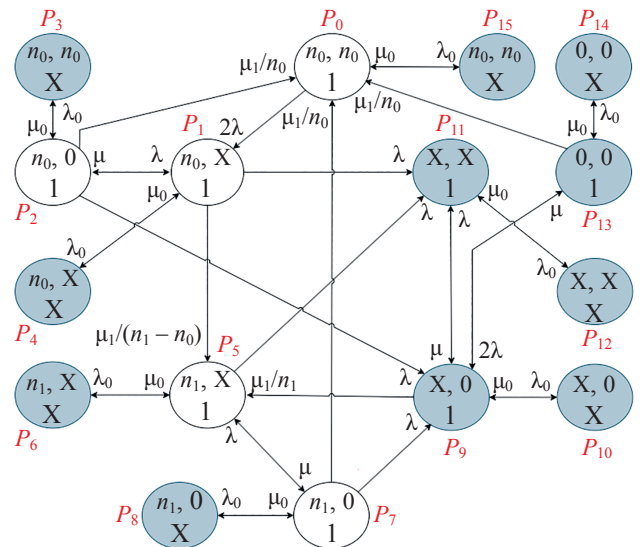


Рис. 3. Диаграмма состояний и переходов кластера с миграцией контейнеров на исправный сервер во время физического восстановления отказавшего сервера.

$P_0, P_1, P_2, \dots, P_{15}$ — вероятности состояний системы.

Fig. 3. State and transition diagram of a cluster without migration prior to the physical recovery stage of the failed server.

$P_0, P_1, P_2, \dots, P_{15}$ — system state probabilities

ром, при этом его работа не влияет на вычислительные процессы в работоспособном сервере и в том числе на миграцию контейнеров в нем, т. е. эти процессы могут быть совмещены по времени.

Для диспетчеризации и выполнения функциональных запросов требуется работоспособность балансировщика и хотя бы одного сервера. Миграция контейнеров требует физической работоспособности сервера и работоспособности балансировщика (диспетчера). Процесс загрузки контейнеров одновременно в два сервера может быть совмещен. Загрузка нескольких контейнеров в сервер производится последовательно.

При построении модели учтено, что накопления отказов серверов не происходит при отказе двух серверов или отказе одного из них и неработоспособности диспетчера, так как в этом случае выполнение запросов в системе невозможно и поэтому предполагаются отключения от серверов электропитания, что исключает отказы еще не отказавших серверов. При неработоспособности двух серверов отключение электропитания у диспетчера-балансировщика (несмотря на невозможность выполнения запросов в серверах) не происходит, ввиду того что в это время он должен выполнять оповещение для поступающих функциональных запросов о невозможности их выполнения.

Условием работоспособности кластера является работоспособность балансировщика нагрузки и хотя бы одного сервера с размещенными на нем контейнерами.

Для варианта 2 (рис. 3) переход в исходное состояние возможен при загрузке n_0 контейнеров параллельно в два либо в один сервер. Возможен также случай, когда в один сервер загружается n_0 контейнеров, а в другой меньшее их число.

Система уравнений, соответствующая диаграмме состояний и переходов на рис. 2, имеет следующий вид:

$$\begin{cases} 0 = -(2\lambda + \lambda_0)P_0 + \frac{\mu_1}{n_0}P_2 + \frac{\mu_1}{n_0}P_9, \\ 0 = -(\lambda_0 + \mu + \lambda)P_1 + 2\lambda P_0 + \frac{\mu_1}{n_0}P_6 + \lambda P_2 + \mu P_4, \\ 0 = -\left(\lambda + \lambda_0 + \lambda + \frac{\mu_1}{n_0}\right)P_2 + \mu P_1 + \mu P_5, \\ 0 = -\mu_0 P_3 + \lambda_0 P_0, \\ 0 = -\mu_0 P_4 + \lambda_0 P_1, \\ 0 = -\mu_0 P_5 + \lambda_0 P_2, \\ 0 = -\left(\lambda_0 + \frac{\mu_1}{n_0} + \lambda + \mu\right)P_6 + \mu_0 P_7 + \lambda P_2 + \mu P_{10} + \lambda P_9, \\ 0 = -\mu_0 P_7 + \lambda_0 P_6, \\ 0 = -\mu_0 P_8 + \lambda_0 P_{10}, \\ 0 = -\left(\lambda_0 + \lambda + \frac{\mu_1}{n_0}\right)P_9 + \mu_0 P_{11} + \mu P_6, \\ 0 = -(\mu + \lambda_0)P_{10} + \lambda P_1 + \lambda P_6 + \mu_0 P_8, \\ 1 = \sum_{i=0}^{11} P_i. \end{cases}$$

Коэффициент готовности кластера в этом случае:

$$K = P_0 + P_1 + P_2.$$

Заметим, что при расчете коэффициента готовности учитываются все состояния, при которых возможно обслуживание потока запросов без их дифференциации по задержкам, независимо от работоспособности одного или двух серверов. При этом предполагается, что физическое и информационное восстановление отказавшего сервера не приводит к замедлению работы исправного сервера, а балансировщик не замедляет процессы в серверах.

Диаграмма состояний и переходов, представленных на рис. 3, приводит к системе уравнений:

$$\begin{cases} 0 = -(2\lambda + \lambda_0)P_0 + \frac{\mu_1}{n_0}P_2 + \mu_0 P_{15} + \frac{\mu_1}{n_0}P_{13} + \frac{\mu_1}{n_0}P_7, \\ 0 = -\left(\lambda_0 + \mu + \frac{\mu_1}{n_1 - n_0} + \lambda\right)P_1 + \lambda P_2 + 2\lambda P_0 + \mu_0 P_4, \\ 0 = -\left(\lambda + \lambda_0 + \lambda + \frac{\mu_1}{n_0}\right)P_2 + \mu P_1 + \mu_0 P_3, \\ 0 = -\mu_0 P_3 + \lambda_0 P_2, \\ 0 = -\mu_0 P_4 + \lambda_0 P_1, \\ 0 = -\mu_0 P_5 + \lambda_0 P_2, \\ 0 = -(\lambda_0 + \lambda + \mu)P_5 + \frac{\mu_1}{n_1 - n_0}P_1 + \mu_0 P_6 + \frac{\mu_1}{n_1}P_9 + \lambda P_7, \\ 0 = -\mu_0 P_6 + \lambda_0 P_5, \\ 0 = -\left(\lambda_0 + \mu + \frac{\mu_1}{n_0} + \mu\right)P_7 + \mu P_5 + \mu P_8, \\ 0 = -\mu_0 P_8 + \lambda_0 P_7, \\ 0 = -\left(\lambda_0 + \mu + \lambda + \frac{\mu_1}{n_1}\right)P_9 + \mu_0 P_{10} + \lambda P_{13} + \mu P_{11} + \\ + \lambda P_2 + \lambda P_7, \\ 0 = -\mu_0 P_{10} + \lambda_0 P_9, \\ 0 = -(\lambda_0 + \mu)P_{11} + \mu_0 P_{12} + \lambda P_9 + \lambda P_5 + \lambda P_1, \\ 0 = -\mu_0 P_{12} + \lambda_0 P_{11}, \\ 0 = -\left(\lambda_0 + \frac{\mu_1}{n_0} + \lambda\right)P_{13} + \mu_0 P_{14} + \mu P_9, \\ 0 = -\mu_0 P_{14} + \lambda_0 P_{13}, \\ 1 = \sum_{i=0}^{15} P_i. \end{cases}$$

Отметим, что в состоянии, вероятность нахождения в котором равна P_1 , выполнение запросов не происходит (так как в одном сервере вычислительный процесс и загрузка контейнеров не совмещены). Исходя из этого, стационарный коэффициент готовности кластера можно вычислить следующим образом:

$$K = P_0 + P_2 + P_5 + P_7.$$

Таблица. Зависимость коэффициента готовности кластера от числа контейнеров n_1 , развернутых на этапе 1 реконфигурации в сохранившем работоспособность сервере

Table. Dependence of the cluster availability coefficient on the number of n_1 containers deployed at the first stage of reconfiguration in a functioning server

Число развернутых контейнеров n_1 , шт.	Коэффициент готовности	Примечание
2	0,998995	На этапе 1 реконфигурации миграции контейнеров нет
3	0,998331	Число загруженных на этапе 1 реконфигурации контейнеров меньше числа контейнеров в отказавшем сервере
4	0,997999	Полная миграция всех контейнеров с отказавшего сервера на исправный
5	0,997799	Число контейнеров, дополнительно загруженных в исправный сервер, больше их числа в отказавшем сервере

Определим влияние числа контейнеров, загружаемых при миграции на двух этапах реконфигурации, на надежность кластера, определяемую по стационарному коэффициенту готовности. При расчетах примем, что в исходном состоянии в каждом из двух серверов загружено по $n_0 = 2$ контейнера, интенсивности отказов балансировщика и сервера $\lambda = \lambda_0 = 10^{-4}$ 1/ч, а интенсивности их восстановлений $\mu_0 = 0,1$ 1/ч и $\mu = 0,1$ 1/ч при интенсивности миграции контейнера $\mu_1 = 0,1$ 1/ч. Результаты расчета коэффициента готовности кластера в зависимости от числа контейнеров n_1 , развернутых на этапе 1 реконфигурации в сохранившем работоспособность сервере, представлены в таблице.

Результаты расчета показывают отрицательное влияние на коэффициент готовности кластера увеличения числа загружаемых контейнеров при миграции на сервер, сохранивший работоспособность после отказа. Это связано с тем, что при миграции контейнеров на этапе 1 реконфигурации выполнение запросов на сервере, сохранившем работоспособность, не происходит. Миграция снижает коэффициент готовности, так как замедляет процесс восстановления: помимо времени на физическое восстановление добавляется время на информационное восстановление, что затягивает двухэтапный процесс реконфигурации и приводит к рискам перехода в неработоспособные состояния, требующие больше усилий (времени) для восстановления исправного состояния системы. Подчеркнем, что миграция контейнеров позволяет в состояниях с частичной потерей вычислительных ресурсов сохранить производительность системы, благодаря накоплению отказов, в том числе при обеспечении оптимального числа контейнеров, что приводит к максимальной производительности кластера.

Обсуждение результатов и направления развития

Предложенные модели надежности кластера направлены на обоснование выбора проектных решений по организации реконфигурации и восстановления работоспособности кластера после отказов, с учетом влияния различных вариантов реализации миграции виртуальных контейнеров на готовность системы.

Отличие предлагаемых марковских моделей заключается в учете двухэтапного восстановления работоспособности кластера при его реконфигурации, с определением влияния на надежность системы числа загружаемых при миграции контейнеров на двух выделенных этапах реконфигурации — до и после физического восстановления отказавших серверов.

В качестве дальнейшего развития работы предполагается исследование комплексного влияния различных вариантов миграции контейнеров как на готовность кластера, так и на задержки обслуживания запросов на двух рассматриваемых этапах реконфигурации.

Заключение

Предложены марковские модели надежности кластеров контейнерной виртуализации с двухэтапным восстановлением работоспособности отказавших серверов при их физическом и информационном восстановлении, с учетом влияния на готовность системы числа загружаемых при миграции контейнеров на каждом этапе восстановления.

Предложенные модели ориентированы на обоснование проектных решений по обеспечению надежности восстанавливаемых кластеров с контейнерной виртуализацией, включая дисциплины их восстановления и реконфигурации.

Литература

References

- Goyal P., Deora S.S. Reliability of Trust Management Systems in Cloud Computing // *Indian Journal of Cryptography and Network Security*. 2022. V. 2. N 1. P. 1–5. <https://doi.org/10.54105/ijcns.C1417.051322>
- Chen G., Guan N., Huang K., Yi W. Fault-tolerant real-time tasks scheduling with dynamic fault handling // *Journal of Systems Architecture*. 2020. V. 102. P. 101688. <https://doi.org/10.1016/j.sysarc.2019.101688>
- Shubinsky I.B., Rozenberg I.N., Papic L. Adaptive fault tolerance in real-time information systems // *Reliability: Theory and Applications*. 2017. V. 12. N 1 (44). P. 18–25.
- Chinnaiah N.R., Niranjana N. Fault tolerant software systems using software configurations for cloud computing // *Journal of Cloud Computing*. 2018. V. 7. P. 3. <https://doi.org/10.1186/s13677-018-0104-9>
- Srivastava A., Kumar N. Queueing model based dynamic scalability for containerized cloud // *International Journal of Advanced Computer Science and Applications*. 2023. V. 14. N 1. P. 465–472. <https://doi.org/10.14569/IJACSA.2023.0140150>
- Shukur H.M., Zeebaree S.R.M., Zebari R.R., Zeebaree D.Q., Ahmed O.M., Salih A.A. Cloud computing virtualization of resources allocation for distributed systems // *Journal of Applied Science and Technology Trends*. 2020. V. 1. N 2. P. 98–105. <https://doi.org/10.38094/jastt1331>
- Alam I., Sharif K., Li F., Latif Z., Karim M.M., Biswas S., Nour B., Wang Y. A survey of network virtualization techniques for Internet of things using SDN and NFV // *ACM Computing Surveys*. 2020. V. 53. N 2. P. 1–40. <https://doi.org/10.1145/3379444>
- Chen H., Qin W., Wang L. Task partitioning and offloading in IoT cloud-edge collaborative computing framework: a survey // *Journal of Cloud Computing*. 2022. V. 11. P. 86. <https://doi.org/10.1186/s13677-022-00365-8>
- Kushchazli A., Safargalieva A., Kochetkova I., Gorshenin A. Queueing model with customer class movement across server groups for analyzing virtual machine migration in cloud computing // *Mathematics*. 2024. V. 12. N 3. P. 468. <https://doi.org/10.3390/math12030468>
- Kumari P., Kaur P. A survey of fault tolerance in cloud computing // *Journal of King Saud University — Computer and Information Sciences*. 2021. V. 33. N 10. P. 1159–1176. <https://doi.org/10.1016/j.jksuci.2018.09.021>
- Tatarnikova T.M., Arkhipov E.D. Designing fault-tolerant systems with micro-service architecture // *Proc. of the 27th International Conference on Soft Computing and Measurements (SCM)*. 2024. P. 348–351. <https://doi.org/10.1109/SCM62608.2024.10554143>
- 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. V. 33. N 1. P. 57–63.
- Sovetov B.Ya., Tatarnikova T.M., Poymanova E.D. Storage scaling management model // *Information and Control Systems*. 2020. N 5 (108). P. 43–49. <https://doi.org/10.31799/1684-8853-2020-5-43-49>
- Bogatyrev A.V., Bogatyrev V.A., Bogatyrev S.V. The probability of timeliness of a fully connected exchange in a redundant real-time communication system // *Proc. of the Wave Electronics and its Application in Information and Telecommunication Systems (WECONF)*. 2020. P. 1–4. <https://doi.org/10.1109/WECONF48837.2020.9131517>
- Bogatyrev V.A., Bogatyrev S.V., Bogatyrev A.V. Control of multipath transmissions in the nodes of switching segments of reserved paths // *Proc. of the International Conference on Information, Control, and Communication Technologies (ICCT)*. 2022. P. 1–5. <https://doi.org/10.1109/ICCT56057.2022.9976839>
- Terskov V., Sakash I. The reliability evaluation of local computer networks using markov model of multiple heterogeneous groups of switches // *E3S Web of Conferences*. 2024. V. 592. P. 3036. <https://doi.org/10.1051/e3sconf/202459203036>
- Половко А.М., Гуров С.В. Основы теории надежности. СПб.: БХВ-Петербург, 2006. 702 с.
- Koren I. Fault-Tolerant Systems. Morgan Kaufmann, 2007. 400 p.
- Aysan H. Fault-tolerance strategies and probabilistic guarantees for real-time systems. Doctoral dissertation, Mälardalen University, 2012. 109 p.
- Рахман П.А., Шарипов М.И. Модель надежности двухузлового кластера приложений высокой готовности в системах управления
- Goyal P., Deora S.S. Reliability of Trust Management Systems in Cloud Computing. *Indian Journal of Cryptography and Network Security*, 2022, vol. 2, no. 1, pp. 1–5. <https://doi.org/10.54105/ijcns.C1417.051322>
- Chen G., Guan N., Huang K., Yi W. Fault-tolerant real-time tasks scheduling with dynamic fault handling. *Journal of Systems Architecture*, 2020, vol. 102, pp. 101688. <https://doi.org/10.1016/j.sysarc.2019.101688>
- Shubinsky I.B., Rozenberg I.N., Papic L. Adaptive fault tolerance in real-time information systems. *Reliability: Theory and Applications*, 2017, vol. 12, no. 1 (44), pp. 18–25.
- Chinnaiah N.R., Niranjana N. Fault tolerant software systems using software configurations for cloud computing. *Journal of Cloud Computing*, 2018, vol. 7, pp. 3. <https://doi.org/10.1186/s13677-018-0104-9>
- Srivastava A., Kumar N. Queueing model based dynamic scalability for containerized cloud. *International Journal of Advanced Computer Science and Applications*, 2023, vol. 14, no. 1, pp. 465–472. <https://doi.org/10.14569/IJACSA.2023.0140150>
- Shukur H.M., Zeebaree S.R.M., Zebari R.R., Zeebaree D.Q., Ahmed O.M., Salih A.A. Cloud computing virtualization of resources allocation for distributed systems. *Journal of Applied Science and Technology Trends*, 2020, vol. 1, no. 2, pp. 98–105. <https://doi.org/10.38094/jastt1331>
- Alam I., Sharif K., Li F., Latif Z., Karim M.M., Biswas S., Nour B., Wang Y. A survey of network virtualization techniques for Internet of things using SDN and NFV. *ACM Computing Surveys*, 2020, vol. 53, no. 2, pp. 1–40. <https://doi.org/10.1145/3379444>
- Chen H., Qin W., Wang L. Task partitioning and offloading in IoT cloud-edge collaborative computing framework: a survey. *Journal of Cloud Computing*, 2022, vol. 11, pp. 86. <https://doi.org/10.1186/s13677-022-00365-8>
- Kushchazli A., Safargalieva A., Kochetkova I., Gorshenin A. Queueing model with customer class movement across server groups for analyzing virtual machine migration in cloud computing. *Mathematics*, 2024, vol. 12, no. 3, pp. 468. <https://doi.org/10.3390/math12030468>
- Kumari P., Kaur P. A survey of fault tolerance in cloud computing. *Journal of King Saud University — Computer and Information Sciences*, 2021, vol. 33, no. 10, pp. 1159–1176. <https://doi.org/10.1016/j.jksuci.2018.09.021>
- Tatarnikova T.M., Arkhipov E.D. Designing fault-tolerant systems with micro-service architecture. *Proc. of the 27th International Conference on Soft Computing and Measurements (SCM)*, 2024, pp. 348–351. <https://doi.org/10.1109/SCM62608.2024.10554143>
- 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.
- Sovetov B.Ya., Tatarnikova T.M., Poymanova E.D. Storage scaling management model. *Information and Control Systems*, 2020, no. 5 (108), pp. 43–49. <https://doi.org/10.31799/1684-8853-2020-5-43-49>
- Bogatyrev A.V., Bogatyrev V.A., Bogatyrev S.V. The probability of timeliness of a fully connected exchange in a redundant real-time communication system. *Proc. of the Wave Electronics and its Application in Information and Telecommunication Systems (WECONF)*, 2020, pp. 1–4. <https://doi.org/10.1109/WECONF48837.2020.9131517>
- Bogatyrev V.A., Bogatyrev S.V., Bogatyrev A.V. Control of multipath transmissions in the nodes of switching segments of reserved paths. *Proc. of the International Conference on Information, Control, and Communication Technologies (ICCT)*, 2022, pp. 1–5. <https://doi.org/10.1109/ICCT56057.2022.9976839>
- Terskov V., Sakash I. The reliability evaluation of local computer networks using markov model of multiple heterogeneous groups of switches. *E3S Web of Conferences*, 2024, vol. 592, pp. 3036. <https://doi.org/10.1051/e3sconf/202459203036>
- Polovko A.M., Gurov S.V. *Fundamentals of Reliability Theory*. St. Petersburg, BHV-Petersburg Publ., 2006, 702 p. (in Russian)
- Koren I. *Fault-Tolerant Systems*. Morgan Kaufmann, 2007, 400 p.
- Aysan H. *Fault-tolerance strategies and probabilistic guarantees for real-time systems*. Doctoral dissertation, Mälardalen University, 2012, 109 p.

- предприятием // Экономика и менеджмент систем управления. 2015. № 3 (17). С. 85–102.
21. Хомоненко А.Д., Благовещенская Е.А., Проурзин О.В., Андрук А.А. Прогноз надежности кластерной вычислительной системы с помощью полумарковской модели альтернирующих процессов и мониторинга // Наукоемкие технологии в космических исследованиях Земли. 2018. Т. 10. № 4. С. 72–82. <https://doi.org/10.24411/2409-5419-2018-10099>
 22. Bogatyrev V.A., Vinokurova M.S. Control and safety of operation of duplicated computer systems // Communications in Computer and Information Science. 2017. V. 700. P. 331–342. https://doi.org/10.1007/978-3-319-66836-9_28
 23. Bogatyrev V.A. Exchange of duplicated computing complexes in fault-tolerant systems // Automatic Control and Computer Sciences. 2011. V. 45. N 5. P. 268–276. <https://doi.org/10.3103/S014641161105004X>
 24. Богатырев В.А., Богатырев С.В., Богатырев А.В. Оценка готовности компьютерной системы к своевременному обслуживанию запросов при его совмещении с информационным восстановлением памяти после отказов // Научно-технический вестник информационных технологий, механики и оптики. 2023. Т. 23. № 3. С. 608–617. <https://doi.org/10.17586/2226-1494-2023-23-3-608-617>
 25. Compastie M., Badonnel R., Festor O., He R. From virtualization security issues to cloud protection opportunities: An in-depth analysis of system virtualization models // Computers & Security. 2020. V. 97. P. 101905. <https://doi.org/10.1016/j.cose.2020.101905>
 26. Choudhary A., Govil M.C., Singh G., Awasthi L.K., Pilli E.S., Kapil D. A critical survey of live virtual machine migration techniques // Journal of Cloud Computing. 2017. V. 6. P. 23. <https://doi.org/10.1186/s13677-017-0092-1>
 27. Алексанков С.М. Модели динамической миграции с итеративным подходом и сетевой миграции виртуальных машин // Научно-технический вестник информационных технологий, механики и оптики. 2015. Т. 15. № 6. С. 1098–1104. <https://doi.org/10.17586/2226-1494-2015-15-6-1098-1104>
 28. Bogatyrev V.A., Derkach A.N. Evaluation of a cyber-physical computing system with migration of virtual machines during continuous computing // Computers. 2020. V. 9. N 2. P. 42. <https://doi.org/10.3390/computers9020042>
 29. Клейнрок Л. Теория массового обслуживания. М.: Машиностроение, 1979. 432 с.
 30. Фунг В., Богатырев В.А., Кармановский Н.С., Лэ В.Х. Оценка вероятностно-временных характеристик компьютерной системы с контейнерной виртуализацией // Научно-технический вестник информационных технологий, механики и оптики. 2024. Т. 24. № 2. С. 249–255. <https://doi.org/10.17586/2226-1494-2024-24-2-249-255>
 31. Nguyen T.A., Kim D.S., Park J.S. A comprehensive availability modeling and analysis of a virtualized servers system using stochastic reward nets // The Scientific World Journal. 2014. V. 2014. P. 165316. <https://doi.org/10.1155/2014/165316>
 20. Rakhman P.A., Sharipov M.I. Reliability model of a two-node cluster of high-availability applications in enterprise management systems. *Ekonomika i menedzhment sistem upravleniya*, 2015, no. 3 (17), pp. 85–102. (in Russian)
 21. Khomonenko A.D., Blagoveshchenskaya E.A., Prouzrin O.V., Andruk A.A. Forecasting the reliability of a cluster computing system using a semi-Markov model of alternating processes and monitoring. *High Technologies in Earth Space Research. H&ES Research*, 2018, vol. 10, no. 4, pp. 72–82. (in Russian). <https://doi.org/10.24411/2409-5419-2018-10099>
 22. Bogatyrev V.A., Vinokurova M.S. Control and safety of operation of duplicated computer systems. *Communications in Computer and Information Science*, 2017, vol. 700, pp. 331–342. https://doi.org/10.1007/978-3-319-66836-9_28
 23. Bogatyrev V.A. Exchange of duplicated computing complexes in fault-tolerant systems. *Automatic Control and Computer Sciences*, 2011, vol. 45, no. 5, pp. 268–276. <https://doi.org/10.3103/S014641161105004X>
 24. Bogatyrev V.A., Bogatyrev S.V., Bogatyrev A.V. Assessment of the readiness of a computer system for timely servicing of requests when combined with information recovery of memory after failures. *Scientific and Technical Journal of Information Technologies, Mechanics and Optics*, 2023, vol. 23, no. 3, pp. 608–617. (in Russian). <https://doi.org/10.17586/2226-1494-2023-23-3-608-617>
 25. Compastie M., Badonnel R., Festor O., He R. From virtualization security issues to cloud protection opportunities: An in-depth analysis of system virtualization models. *Computers & Security*, 2020, vol. 97, pp. 101905. <https://doi.org/10.1016/j.cose.2020.101905>
 26. Choudhary A., Govil M.C., Singh G., Awasthi L.K., Pilli E.S., Kapil D. A critical survey of live virtual machine migration techniques. *Journal of Cloud Computing*, 2017, vol. 6, pp. 23. <https://doi.org/10.1186/s13677-017-0092-1>
 27. Aleksankov S.M. Models of live migration with iterative approach and move of virtual machines. *Scientific and Technical Journal of Information Technologies, Mechanics and Optics*, 2015, vol. 15, no. 6, pp. 1098–1104. (in Russian). <https://doi.org/10.17586/2226-1494-2015-15-6-1098-1104>
 28. Bogatyrev V.A., Derkach A.N. Evaluation of a cyber-physical computing system with migration of virtual machines during continuous computing. *Computers*, 2020, vol. 9, no. 2, pp. 42. <https://doi.org/10.3390/computers9020042>
 29. Kleinrock L. *Queueing Systems*. Volume 1: Theory. Wiley-Interscience, 1975, 417 p.
 30. Phung V.Q., Bogatyrev V.F., Karmanovskiy N.S., Le V.H. Evaluation of probabilistic-temporal characteristics of a computer system with container virtualization. *Scientific and Technical Journal of Information Technologies, Mechanics and Optics*, 2024, vol. 24, no. 2, pp. 249–255. (in Russian). <https://doi.org/10.17586/2226-1494-2024-24-2-249-255>
 31. Nguyen T.A., Kim D.S., Park J.S. A comprehensive availability modeling and analysis of a virtualized servers system using stochastic reward nets. *The Scientific World Journal*. 2014. V. 2014. P. 165316. <https://doi.org/10.1155/2014/165316>

Авторы

Богатырев Владимир Анатольевич — доктор технических наук, профессор, Санкт-Петербургский государственный университет аэрокосмического приборостроения, Санкт-Петербург, 190000, Российская Федерация; профессор, Университет ИТМО, Санкт-Петербург, 197101, Российская Федерация, [sc 7006571069](https://orcid.org/0000-0003-0213-0223), <https://orcid.org/0000-0003-0213-0223>, vladimir.bogatyrev@gmail.com

Фунг Ван Кю — аспирант, Университет ИТМО, Санкт-Петербург, 197101, Российская Федерация, <https://orcid.org/0009-0006-3278-1106>, phungvanquy97@gmail.com

Authors

Vladimir A. Bogatyrev — D.Sc., Professor, Saint Petersburg State University of Aerospace Instrumentation (SUAI), Saint Petersburg, 190000, Russian Federation; Professor, ITMO University, Saint Petersburg, 197101, Russian Federation, [sc 7006571069](https://orcid.org/0000-0003-0213-0223), <https://orcid.org/0000-0003-0213-0223>, vladimir.bogatyrev@gmail.com

Van Quy Phung — PhD Student, ITMO University, Saint Petersburg, 197101, Russian Federation, <https://orcid.org/0009-0006-3278-1106>, phungvanquy97@gmail.com

Статья поступила в редакцию 08.05.2025
Одобрена после рецензирования 24.08.2025
Принята к печати 21.09.2025

Received 08.05.2025
Approved after reviewing 24.08.2025
Accepted 21.09.2025



Работа доступна по лицензии
Creative Commons
«Attribution-NonCommercial»