

МОДЕЛЬ ПРОЦЕССА ДИНАМИЧЕСКОЙ МИГРАЦИИ С КОПИРОВАНИЕМ ДАННЫХ ПОСЛЕ ОСТАНОВКИ ВИРТУАЛЬНЫХ МАШИН

С. М. АЛЕКСАНКОВ

Университет ИТМО, 197101, Санкт-Петербург, Россия

НИИ „Масштаб“, 194100, Санкт-Петербург, Россия

E-mail: aleksankov.sergey@gmail.com

Исследуется процесс динамической миграции с копированием данных после остановки виртуальных машин без использования общего хранилища. Алгоритм динамической миграции рассматривается с позиции исследования времени переноса и недоступности сервисов. Предложена аналитическая модель, позволяющая оценить ожидаемое время переноса виртуальных машин и время недоступности сервисов. Модель может быть использована для анализа и оптимизации функционирования современных центров обработки данных кластерной архитектуры, где перемещение виртуальных машин может осуществляться между физическими серверами при профилактических и аварийных работах.

Ключевые слова: динамическая миграция, виртуализация, виртуальные машины, высокая доступность, кластеры, моделирование, копирование после остановки

Введение. Высокая надежность, отказоустойчивость и безопасность систем критического применения обеспечивается при резервировании основных подсистем [1—7], в том числе средств хранения [8], обработки и передачи данных [9] с консолидацией вычислительных ресурсов, а также при кластеризации [4—6], виртуализации [1—6] и динамическом распределении запросов.

Для создания условий непрерывного доступа клиентов к приложениям и сервисам используются кластеры высокой доступности. Высокая доступность ресурсов кластерной системы в случае сбоя одного из серверов или при проведении профилактических работ обеспечивается в результате перемещения служб или приложений между узлами [10—15]. Кластер (для пользователя) представляет собой единое целое, и его неполадки могут быть выражены в кратковременном снижении производительности или недоступности какого-либо ресурса в течение определенного времени (от нескольких секунд до нескольких минут) в зависимости от организации системы [1, 2].

Неотъемлемой частью технологии виртуализации является миграция виртуальных машин (ВМ), что позволяет обеспечить прозрачное перемещение приложений вместе с их средами выполнения между физическими машинами [3].

Динамическая миграция (Live Migration) [1—6] направлена на повышение эффективности центров обработки данных на основе балансировки загрузки, оптимизации размещения виртуальных машин на физических узлах, обеспечения высокой готовности и отказоустойчивости серверов кластерной системы. Технология динамической миграции включена в пакеты известных систем виртуализации (KVM, Microsoft HyperV, VMWare VSphere, Xen).

С помощью технологии динамической миграции виртуальные машины могут быть переданы с одного физического узла на другой без отключения системы, что позволяет существенно повысить ее доступность, сократить задержки обслуживания запросов и обеспечить непрерывность вычислительного процесса в следующих случаях:

— неравномерная нагрузка узлов кластера, во избежание которой требуется балансировка нагрузки кластера, т.е. виртуальные машины на перегруженном физическом узле могут быть перенесены на менее загруженные узлы;

— аварийные работы, в том числе связанные с восстановлением работоспособности отдельных отказавших узлов кластера;

— профилактические работы.

В настоящее время технология динамической миграции интенсивно внедряется при построении центров обработки данных и отказоустойчивых компьютерных систем кластерной архитектуры высокой готовности. Эффективность этой технологии подтверждается результатами испытаний и опытом эксплуатации реальных систем.

Вместе с тем отсутствие аналитических моделей систем с динамической миграцией виртуальных машин затрудняет процесс системотехнического проектирования таких систем, основанного на анализе проектных решений и их оптимизации. В связи с этим в настоящей статье предложены варианты построения аналитических моделей оценивания времени переноса виртуальных машин и времени недоступности сервисов при динамической миграции. Также рассматриваются возможности применения предлагаемых моделей для оценки эффективности проектных решений построения вычислительных кластеров на основе технологии динамической миграции виртуальных машин.

Задачи исследования. Процесс динамической миграции, позволяющий переносить виртуальные машины практически не прерывая выполнение приложений, можно разделить на два шага:

— передача управления узлу назначения;

— передача данных (память, диск) на узел назначения.

Возможны два способа реализации процесса динамической миграции, основанные на различной последовательности указанных шагов:

1) с копированием данных до остановки ВМ (pre-copy): сначала происходит передача данных, а затем — передача управления;

2) с копированием данных после остановки ВМ (post-copy): сначала происходит передача управления узлу назначения, а затем — передача данных.

Построению моделей для расчета времени переноса и времени недоступности сервисов при первом способе посвящена работа [7], в которой показано, что данный подход позволяет сократить время прерывания процесса выполнения приложений, но только при малых значениях скорости чтения.

В настоящей статье рассматривается второй способ — с копированием данных после остановки ВМ или с копированием данных в фоновом режиме. В общем этот процесс можно разделить на следующие этапы.

1. Остановка виртуальной машины на исходном узле.

2. Копирование содержимого регистров виртуальных процессоров и данных о состоянии устройств на узел назначения (Working Set).

3. Возобновление функционирования виртуальной машины на узле назначения.

4. Копирование данных памяти и виртуальных дисков на узел назначения в фоновом режиме. Если виртуальная машина „пытается“ получить доступ к отсутствующей странице, то происходит ошибка доступа, остановка выполнения приложения, и требуемая страница передается узлу назначения (передача по запросу).

На рис. 1 представлена временная диаграмма динамической миграции с копированием данных после остановки ВМ; здесь T_{WS} — время передачи “Working Set”; $(T_{b_1}, T_{b_2}, \dots, T_{b_N}) \in T_b$ — время этапов передачи данных в фоновом режиме; $(T_{a_1}, T_{a_2}, \dots, T_{a_N}) \in T_a$ — время передачи страниц по запросу; vm_1, vm_2, \dots, vm_N — передаваемые страницы памяти в фоновом режиме.

Способ копирования данных после остановки ВМ хорошо себя зарекомендовал при высокой скорости записи, но в случае высокой скорости чтения большое количество передач страниц по запросу приводит к длительному простоя системы [3].

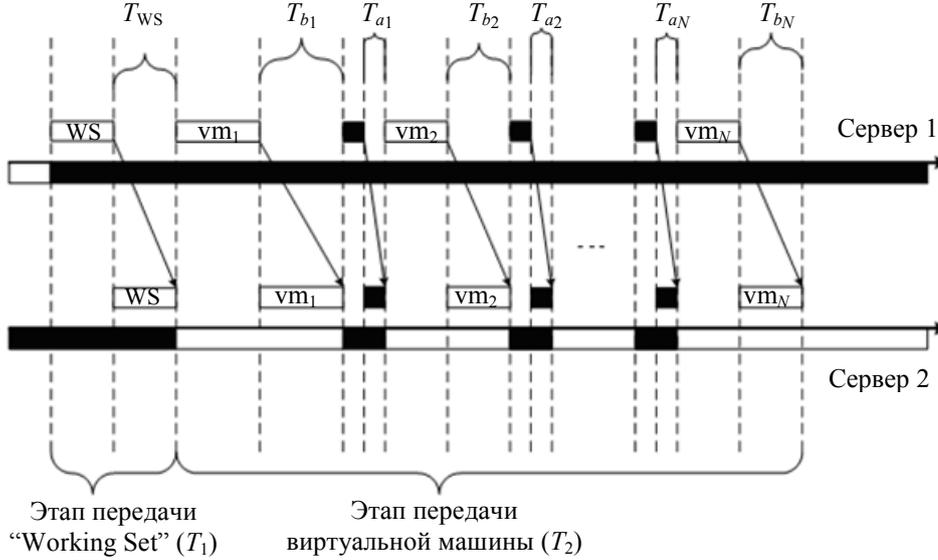


Рис. 1

При построении моделей предполагается, что виртуальные машины работают на кластере из двух физических серверов без общего хранилища (Live Migration Nothing Shared). Такой вид динамической миграции реализован в системах VMWare VSphere и Microsoft HyperV.

Оценка времени динамической миграции. Будем считать известными скорость чтения данных — λ_r , скорость передачи данных — μ , объем данных виртуальной машины (диск, память) — V .

Время переноса ВМ T_m определяется как сумма времени “Working Set” (T_1) и времени передачи данных памяти и дисков (T_2):

$$T_m = T_1 + T_2. \tag{1}$$

Время T_1 рассчитывается как

$$T_1 = \frac{V_1}{\mu} = \frac{V_{WS}}{\mu},$$

где V_1 — объем данных, передаваемых при копировании содержимого регистров виртуальных процессоров и состояний устройств (также называемый “Working Set”) со скоростью μ , а время T_2 — как

$$T_2 = T_a + T_b,$$

при этом $T_a = V_a / \mu$, $T_b = V_b / \mu\alpha$, где V_a — объем данных, переданных по запросу за время T_a со скоростью μ ; V_b — объем данных, переданных в фоновом режиме за время T_b со скоростью μ ; α — коэффициент накладных расходов на другие (вычислительные) задачи.

Таким образом,

$$T_2 = \frac{V_a}{\mu} + \frac{V_b}{\mu\alpha}. \tag{2}$$

Объем данных, переданных по запросу, равен

$$V_a = \lambda_r T_2, \tag{3}$$

т.е. количеству данных, переданных при обращении к ним ВМ для чтения.

Объем данных, переданных в фоновом режиме, определяется как

$$V_b = V - V_a = V - \lambda_r T_2. \tag{4}$$

Подставив выражения (3) и (4) в формулу (2), получим

$$T_2 = \frac{\lambda_r T_2}{\mu} + \frac{V - \lambda_r T_2}{\mu \alpha},$$

откуда

$$T_2 = \frac{V}{\mu \alpha + \lambda_r (1 - \alpha)}.$$

Тогда общее время передачи содержимого памяти, относящейся к переносимой виртуальной машине, вычисляется следующим образом:

$$T_m = \frac{V_{WS}}{\mu} + \frac{V}{\mu \alpha + \lambda_r (1 - \alpha)},$$

Время недоступности сервиса при миграции виртуальной машины определяется как

$$T_d = T_a = \frac{V - \lambda_r T_2}{\mu \alpha}.$$

Результаты расчетов. На рис. 2 представлен график зависимости времени переноса виртуальной машины и времени недоступности сервисов от скорости чтения данных. Время миграции T_m сначала с ростом скорости чтения уменьшается, а затем устанавливается в некотором значении. Это связано с тем, что при малых величинах λ_r значительное количество данных передается в фоновом режиме с меньшей скоростью μ , соответственно на это затрачивается больше времени. Чем выше значения λ_r , тем больше происходит ошибок доступа к страницам, вследствие чего больше данных передается по запросу с максимально возможной скоростью μ , поэтому время T_m уменьшается. Время недоступности сервиса T_d с увеличением скорости чтения вначале возрастает, а затем устанавливается в некотором значении. При малых значениях λ_r почти не происходит ошибок доступа к страницам, и функционирование ВМ не прерывается. С ростом λ_r вероятность ошибок доступа к сервису ВМ увеличивается.

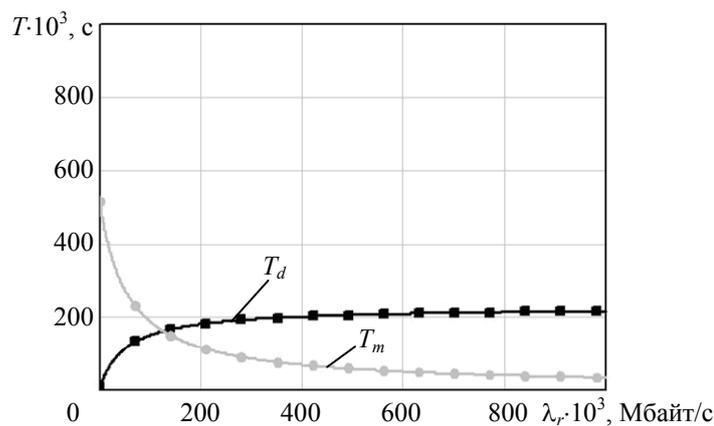


Рис. 2

Анализ зависимости времени T_m и времени T_d от скорости передачи данных μ (рис. 3) показывает, что время недоступности сервисов уменьшается уже вначале и устанавливается в малых значениях, так как содержимое памяти и диска передается быстро и ошибок доступа не возникает. Аналогично уменьшается и время переноса, устанавливаясь в некотором значении, после того как почти перестают возникать ошибки доступа к страницам.

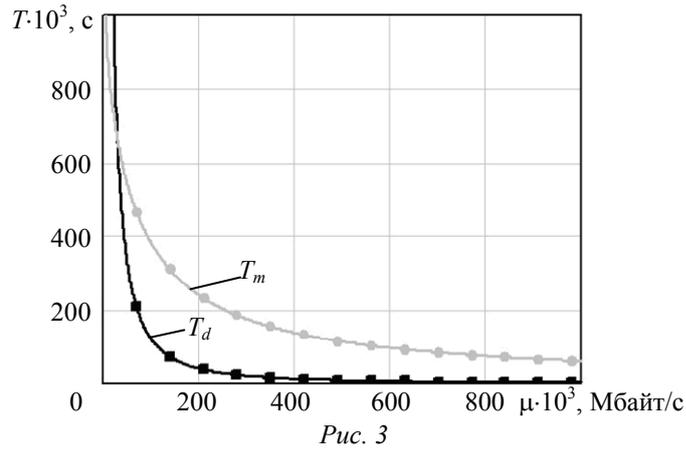
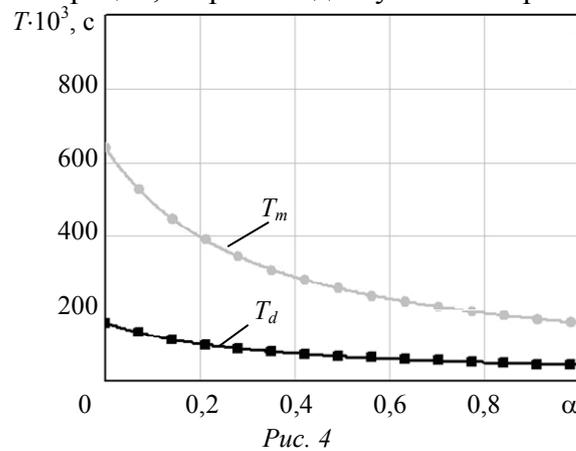
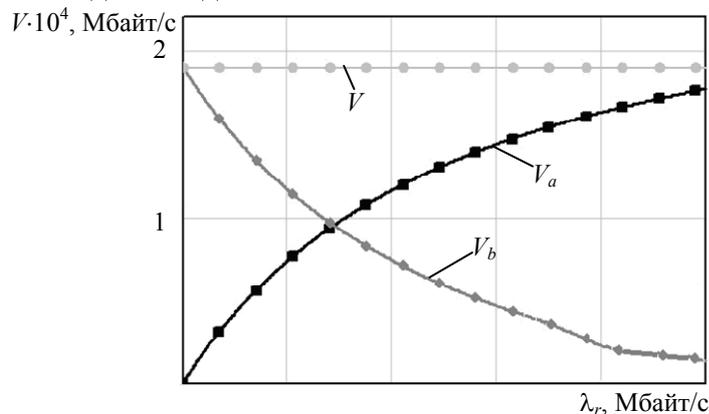


График зависимости времени T_m и времени T_d от коэффициента накладных расходов α представлен на рис. 4. С ростом значения α увеличивается скорость передачи данных μ , при этом уменьшается и время миграции, и время недоступности сервисов.



Анализ зависимостей объема данных (при передаче по запросу, в фоновом режиме, суммарного количества) от скорости чтения (рис. 5) показывает, что с увеличением значения λ_r уменьшается объем данных V_b , передаваемых в фоновом режиме, так как увеличивается количество ошибок доступа и соответственно объем данных V_a , передаваемых по запросу. Суммарный объем данных V не изменяется, в отличие, например, от метода динамической миграции с копированием данных до остановки ВМ.



Все расчеты были проведены для $V=20\,000$ Мбайт, $WS=1000$ Мбайт, $\lambda_r = 30$ Мбайт/с, $\mu=125$ Мбайт/с, $\alpha=0,3$ с использованием пакета компьютерной математики MathCad 15.

Расчеты подтверждают, что использование динамической миграции с копированием данных после остановки виртуальной машины позволяет существенно сократить время недоступности сервисов при больших значениях скорости записи, в отличие от первого подхода, который при этих значениях не эффективен. Следует, однако, отметить, что при больших значениях скорости чтения λ_r не эффективен второй способ, так как увеличивается время недоступности сервисов.

Предлагаемые модели могут быть использованы для анализа и оптимизации функционирования современных центров обработки данных кластерной архитектуры, где перемещение виртуальных машин может осуществляться между физическими серверами при профилактических и аварийных работах для обеспечения балансировки нагрузки. Функции виртуализации могут эффективно использоваться при высокой доступности кластерных систем и совмещении с динамическим распределением запросов. При миграции виртуальных машин желательно минимизировать недоступность сервисов, чего и позволяет добиться динамическая миграция.

Заключение. Предложена модель, позволяющая оценить время переноса виртуальных машин и время недоступности сервисов при динамической миграции с копированием данных после остановки ВМ. Рассмотренная аналитическая модель показала, что данный способ динамической миграции рекомендуется применять во время профилактических работ, когда скорость чтения данных виртуальной машиной невысока, но при этом ее можно использовать при высокой скорости записи.

СПИСОК ЛИТЕРАТУРЫ

1. *Hai Jin, Li Deng, Song Wu, Xuanhua Shi, Xiaodong Pan.* Live virtual machine migration with adaptive memory compression // Cluster Computing and Workshops (CLUSTER '09): IEEE Intern. Conf., 31 Aug. — 4 Sept. 2009. P. 1—10.
2. *Bing Wei.* A novel energy optimized and workload adaptive modeling for live migration // Intern. Journal of Machine Learning and Computing. 2012. Vol. 2, N 2. P. 162—167.
3. *Sahni S., Varma V.* A hybrid approach to live migration of virtual machines // Cloud Computing in Emerging Markets (CCEM): IEEE Intern. Conf., 11—12 Oct. 2012. P. 1—5.
4. *Косивченко А.* Повышение надежности систем на базе виртуализации. 2009 [Электронный ресурс]: <<http://technet.microsoft.com/ru-ru/library/ee922690.aspx>>, (дата обращения 30.05.2015).
5. Обзор миграции виртуальных машин и хранилища. 2013 [Электронный ресурс]: <<https://technet.microsoft.com/ru-ru/library/jj628158.aspx>>, (дата обращения 30.05.2015).
6. *Yangyang Wu, MingZhao.* Performance modeling of virtual machine live migration // Cloud Computing (CLOUD): IEEE Intern. Conf., 4—9 July 2011. P. 492—499.
7. *Алексанков С. М.* Модели динамической миграции с итеративным подходом и сетевой миграции виртуальных машин // Научно-технический вестник информационных технологий, механики и оптики. 2015. № 6 (15). С. 1098—1104.
8. *Богатырев В. А., Алексанков С. М., Демидов Д. В., Беззубов В. Ф.* Надежность резервированного вычислительного комплекса при ограниченном восстановлении // Научно-технический вестник информационных технологий, механики и оптики. 2013. № 3 (85). С. 67—72.
9. *Богатырев В. А., Богатырев С. В., Богатырев А. В.* Оптимизация кластера с ограниченной доступностью кластерных групп // Научно-технический вестник СПбГУ ИТМО. 2011. № 1 (71). С. 63—67.
10. *Богатырев В. А.* К распределению функциональных ресурсов в отказоустойчивых многомашинных вычислительных системах // Приборы и системы. Управление, контроль, диагностика. 2001. № 12. С. 1—5.
11. *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, N 1. P. 46—56. DOI: 10.3103/S0146411615010022.

12. Богатырев В. А. Оценка надежности и оптимальное резервирование кластерных компьютерных систем // Приборы и системы. Управление, контроль, диагностика. 2006. № 10. С. 18—21.
13. Богатырев В. А., Богатырев А. В., Богатырев С. В. Оценка надежности выполнения кластерами запросов реального времени // Изв. вузов. Приборостроение. 2014. Т. 57, № 4. С. 46—48.
14. Богатырев В. А., Богатырев А. В., Богатырев С. В. Перераспределение запросов между вычислительными кластерами при их деградации // Изв. вузов. Приборостроение. 2014. Т. 57, № 9. С. 54—58.
15. Богатырев В. А., Богатырев А. В. Оптимизация резервированного распределения запросов в кластерных системах реального времени // Информационные технологии. 2015. Т. 21, № 7. С. 495—502.

Сведения об авторе

Сергей Михайлович Алексанков

— аспирант; Университет ИТМО, кафедра вычислительной техники; НИИ „Масштаб“; инженер; E-mail: alexsankov.sergey@gmail.com

Рекомендована кафедрой
вычислительной техники

Поступила в редакцию
08.02.16 г.

Ссылка для цитирования: Алексанков С. М. Модель процесса динамической миграции с копированием данных после остановки виртуальных машин // Изв. вузов. Приборостроение. 2016. Т. 59, № 5. С. 348—354.

**MODEL OF LIVE MIGRATION PROCESS
WITH DATA COPYING AFTER VIRTUAL MACHINES STOPPING**

S. M. Aleksankov

ITMO University, 197101, St. Petersburg, Russia

Masshtab Research Institute, 194100, St. Petersburg, Russia

E-mail: alexsankov.sergey@gmail.com

The process of live migration without shared storage with post-copy approach is studied. An algorithm of live migration of virtual machines is considered from the view point of migration time investigation and unavailability of services. An analytical model is proposed for predicting the migration time and time of unavailability of services, for example, at management implementation of preventive and emergency works on the physical nodes in data centers.

Keywords: live migration, virtualization, virtual machines, high availability, clusters, modeling, post-copy

Data on author

Sergey M. Aleksankov — Post-Graduate Student; ITMO University, Department of Computation Technologies; Masshtab Research Institute, Engineer; E-mail: alexsankov.sergey@gmail.com

For citation: Aleksankov S. M. Model of live migration process with data copying after virtual machines stopping // Izv. vuzov. Priborostroyeniye. 2016. Vol. 59, N 5. P. 348—354 (in Russian).

DOI: 10.17586/0021-3454-2016-59-5-348-354