



Московский государственный университет имени М.В. Ломоносова
Факультет вычислительной математики и кибернетики
Кафедра автоматизации систем вычислительных комплексов

Михеев Павел Алексеевич

**Исследование эффективности использования
физических ресурсов
легковесными контейнерами в облачных системах.**

Научный руководитель:
к.ф.-м.н.
В.А. Антоненко

Москва, 2018

Аннотация

Контейнер - изолированная группа процессов в операционной системе, имеющая доступ к ограниченному количеству ресурсов системы. Данная изоляция достигается за счет механизмов ядра ОС, таким образом, контейнеры используют это ядро, не тратя ресурсы на поддержку гостевого ядра.

Контейнеры обладают небольшим временем запуска, большой пропускной способностью, и показатель производительности вычислений в них ближе к показателю вычислений на физических ресурсах [1], [2]. Однако не существует систем, позволяющих исследовать производительность и деградацию производительности контейнеров при масштабировании и сравнить ее с производительностью и деградацией для виртуальных машин.

В данной работе предлагается использовать существующие методики измерения производительности виртуальных машин и применить их к легковесным контейнерам [3], [4].

Целью работы является разработка системы измерения производительности и ее деградации у виртуальных машин и контейнеров, запущенных на физическом сервере.

Данная система позволит численно сравнить производительность контейнеров и виртуальных машин.

Введение

Облачные вычисления в современной информационной среде занимают все больше и больше пространства. Развитие открытого программного обеспечения по управлению облаками увеличивает количество тех, кто использует облачные системы как для использования внутри компании, так и для предоставления услуг клиентам. Одним из немногих сдерживающих факторов такого роста вовлеченности использования облачных вычислений является цена физического оборудования.

Традиционно облачные вычисления используют виртуальные машины как сущности, которые так или иначе предоставляются пользователю. Так, в случае облака, работающего по модели Infrastructure-as-a-Service (IaaS), пользователю предоставляется полный доступ к заказанной виртуальной машине, которую он настраивает под собственные требования. В облаках типа Platform-as-a-Service предоставляется доступ к интерфейсам некоторых приложений, запущенных в виртуальной машине. В облаках Software-as-a-Service пользователь запрашивает услугу, обработка которой происходит с помощью виртуальных машин, доступа к которым у пользователя нет.

Количество виртуальных машин, которое может быть запущено в облаке, напрямую зависит от количества физических ресурсов. Под количеством физических ресурсов в данной работе будет пониматься совокупное (по всем серверам) количество ядер процессоров, совокупное количество оперативной памяти и совокупный объем жестких дисков системы.

Пусть X - количество доступных физических ресурсов в облаке, а χ_i - количество ресурсов, требующихся для i -ой виртуальной машины. Тогда n - максимальное число виртуальных машин с такими требованиями, которое может быть запущено на данных ресурсах, если $\sum_{i=1}^n \chi_i = X$.

Однако может возникать ситуация, когда количество запрашиваемых

виртуальных машин превосходит максимальное число n , определенное выше. При этом у владельца облака нет возможности или необходимости докупить оборудование, так как, к примеру, такая ситуация возникает редко. Важный момент, что ресурсы виртуальной машины могут быть использованы ее процессами не на 100% процентов. Тогда, если предоставить доступ к этим же физическим ресурсам или их части той виртуальной машине, номер которой превосходит n , то обе эти виртуальные машины смогут осуществлять свои функции. В таком случае $\sum_{i=1}^n \chi_i > X$, но при этом все виртуальные машины размещены.

Введем коэффициент $\theta = \frac{\sum_i \chi_i}{X}$ - коэффициент перекрытия (overlap), являющийся отношением суммы ресурсов, требуемых виртуальным машинам системы, к физическим ресурсам системы.

Разумеется, при увеличении коэффициента перекрытия производительность процессов в виртуальных машинах падает, так как при использовании одних физических ресурсов (в первую очередь, ядер процессора) разными виртуальными машинами исполнение машин одновременно будет невозможно. Из-за этого будет наблюдаться деградация производительности виртуальных машин при увеличении коэффициента перекрытия.

Еще одной быстро занявшей рынок технологией стала легковесная виртуализация. В этой технологии виртуализации гостевая операционная система отсутствует, а все процессы запускаются в рамках ядра хостовой операционной системы. Изоляция подобных процессов друг от друга и ограничение доступных им ресурсов достигается за счет специальных механизмов ядра. Такие процессы и их потомки с наложенными на них ограничениями называются контейнерами. Так же как и процессы в виртуальной машине изолированы от процессов другой виртуальной машины, процессы в контейнере изолированы от процессов других контейнеров. Отличие заключается в том, что системные вызовы в вирту-

альной машине идут в операционную систему машины, тогда как системные вызовы контейнеров идут напрямую к ядру операционной системы, в которой данный контейнер запущен.

Отсутствие необходимости поддерживать гостевую операционную систему обладает как достоинствами, так и недостатками. Подробнее о них будет сказано ниже.

Контейнеры могут быть использованы в облачных системах так же, как и виртуальные машины. Они потребляют некоторое количество ресурсов, ресурсы могут разделять между несколькими контейнерами, а при увеличении коэффициента перекрытия будет наблюдаться деградация производительности.

В данной работе предлагается исследовать и сравнить производительность виртуальных машин и контейнеров, а так же деградацию производительности машин и контейнеров при увеличении коэффициента перекрытия. Для этого была разработана система, способная запускать виртуальные машины или контейнеры и запускать в них приложения, позволяющие измерять производительность, а так же собирать результаты работы этих приложений.

Основной гипотезой данной работы является следующее: контейнеры обладают более высокой производительностью по сравнению с виртуальными машинами, деградация производительности виртуальных машин при увеличении коэффициента перекрытия происходит быстрее, чем у контейнеров.

Глава 1.

Постановка задачи

Цель работы

Разработать и реализовать систему, позволяющую сравнить производительность виртуальных машин и контейнеров, а так же деградацию производительности виртуальных машин и контейнеров при увеличении коэффициента перекрытия.

С помощью разработанной системы провести эксперименты, позволяющие проверить следующую гипотезу: **контейнеры обладают более высокой производительностью по сравнению с виртуальными машинами, деградация производительности виртуальных машин при увеличении коэффициента перекрытия происходит быстрее, чем у контейнеров.**

План решения задачи

1. Сравнить различные технологии виртуализации.
2. Составить обзор методик оценки производительности виртуальных машин.
3. Выбрать из предыдущего обзора методику оценки производительности

сти или разработать свою, которую возможно применить для оценки производительности контейнеров.

4. Разработать систему, реализующую методику оценки производительности из предыдущего пункта, позволяющий численно оценить производительность и ее деградацию при увеличении коэффициента перекрытия в случае виртуальных машин и контейнеров.
5. С помощью разработанной системы провести эксперименты, позволяющие проверить следующую гипотезу: **контейнеры обладают более высокой производительностью по сравнению с виртуальными машинами, деградация производительности виртуальных машин при увеличении коэффициента перекрытия происходит быстрее, чем у контейнеров.**

Глава 2.

Технологии виртуализации

Глава 3.

Обзор существующих решений

Привет, git. Почему ты не работаешь?

фывфыв

Литература

- [1] Jason Anderson, Udit Agarwal, Hongda Li. Performance Considerations of Network Functions Virtualization using Containers. //2016International Conference on Computing, Networking and Communications (ICNC), 2016.
- [2] Wes Felter, Alexandre Ferreira [и др.]. An Updated Performance Comparison of Virtual Machines and Linux Containers. // IBM Research Report, 2014.
- [3] George Kousiouris, Tommaso Cucinotta, Theodora Varvarigoua. The effects of scheduling, workload type and consolidation scenarios on virtual machine performance and their prediction through optimized artificial neural networks. // The Journal of Systems and Software, 2011.
- [4] Anton Beloglazov, Rajkumar Buyya. Optimal online deterministic algorithms and adaptive heuristics for energy and performance efficient dynamic consolidation of virtual machines in Cloud data centers. // Wiley Online Library, 2011.
- [5] Haikun Liu, Hai Jin, Cheng-Zhong Xu, Xiaofei Liao. Performance and energy modeling for live migration of virtual machines. // Springer US, 2013.
- [6] Homepage of VirtualBox. <https://www.virtualbox.org> (дата обращения 01.09.2017)

- [7] Homepage of Xen. <https://www.xenproject.org> (дата обращения 01.09.2017)
- [8] Homepage of VMWare ESX. <http://www.vmware.com/ru/products/esxi-and-esx.html> (дата обращения 01.09.2017)
- [9] Homepage of Qemu. <http://www.qemu.org> (дата обращения 01.09.2017)
- [10] Homepage of KVM. <https://www.linux-kvm.org> (дата обращения 01.09.2017)
- [11] Quan Jia, Zhaohui Wang, Angelos Stavrou. The Heisenberg Measuring Uncertainty in Lightweight Virtualization Testbeds. //Department of Computer Science, George Mason University, Fairfax, VA, 2009.
- [12] Homepage of LXC. URL: <https://linuxcontainers.org/ru/> (дата обращения 01.09.2017)
- [13] Homepage of OpenVZ. URL: <https://openvz.org/> (дата обращения 31.10.2017)
- [14] Homepage of Rocket. URL: <https://coreos.com/rkt/docs/latest/> (дата обращения 31.10.2017)
- [15] Homepage of DrawBridge. URL: <http://research.microsoft.com/en-us/projects/drawbridge/> (дата обращения 31.10.2017)
- [16] Homepage of runC. URL: <https://runc.io/> (дата обращения 30.04.2017)
- [17] Homepage of Docker. URL: <https://docs.docker.com/> (дата обращения 01.09.2017)