

[3] Nigrini M.J. Benford's Law: applications for forensic accounting, auditing, and fraud detection. – Hoboken: John Wiley & Sons, Inc., 2012. – XX + 330 pp.

[4] Battersby S. Statistics hint at fraud in Iranian election // New Scientist. – 24 June 2009.

[5] Sambridge M., Tkalčić H., Arroucau P. Benford's Law of First Digits: from Mathematical Curiosity to Change Detector // Asia Pacific Mathematics Newsletter. – 2011. – vol. 1, No. 4. – P. 1–6.

[6] Andriotis P., Oikonomou G., Tryfonas T. JPEG steganography detection with Benford's Law // Digital Investigation. – 2013. – vol. 9, No. 3–4. – P. 246–257.

[7] Зенков А.В. Отклонения от закона Бенфорда и распознавание авторских особенностей в текстах // Компьютерные исследования и моделирование. – 2015. – Т. 7, вып. 1. – С. 197–201.

[8] Hünigerbühler N. Benfords Gesetz über führende Ziffern: wie die Mathematik Steuersündern das Fürchten lehrt // EducETH, Publikation der Eidgenössischen Technischen Hochschule Zürich. – 2007. – www.educ.ethz.ch/unt/um/mathe/ana/benford

[9] Manning C.D., Schütze H. Foundations of Statistical Natural Language Processing. – Cambridge (Mass.) – London: The MIT Press, 1999. – XXXVII + 680 p.

ИСПОЛЬЗОВАНИЕ ВИРТУАЛИЗАЦИИ В УЧЕБНОМ ПРОЦЕССЕ ДЛЯ ПОВЫШЕНИЯ КАЧЕСТВА ОБУЧЕНИЯ

УДК 004.053

Д.Г. Ермаков к. ф.-м. н., доцент
кафедра анализа систем и принятия решений
Уральский федеральный университет, ВШЭМ

А.В. Присяжный к.т.н., доцент
кафедра анализа систем и принятия решений
Уральский федеральный университет, ВШЭМ

О.Е. Хорев студент
кафедра анализа систем и принятия решений
Уральский федеральный университет, ВШЭМ

Аннотация: данная публикация посвящается решению проблемы повышения наглядности преподаваемых предметов в условиях ограниченных ресурсов или их экономии. Еще один взгляд на виртуальные машины в процессе обучения и необычное использование торрентов.

Ключевые слова: Виртуализация, VirtualPC, VmWare, VirtualBox.

Experience and perspectives of virtualization technologies in the learning process of students. Current problems and decision ways.

Key words: Virtualization, VirtualPC, VmWare, VirtualBox.

Введение и суть проблемы

Одной из наиболее важных и существенных задач в любого рода обучении, является возможность наглядной иллюстрации преподаваемого материала. Широко известно, что при использовании наглядных материалов более эффективно проходят практические занятия и, в конечном итоге, достаточно глубокое усвоение поданного материала студентам.

Исходя из своего личного опыта работы со студентами, могу сказать, что мне нередко приходится сталкиваться с проблемой усвоения, увязывания теории и практики в одно неделимое целое в таких важных предметах, как операционные системы, вычислительные сети, информационная безопасность.

Проблема банальна и проста и складывается из ограничений бюджета и ограниченных аппаратных возможностей, не позволяющих студентам самостоятельно работать на практике со всеми изучаемыми технологиями.

На данный момент практика организуется так, что все учащиеся работают на компьютерах под пользовательскими учетными записями с минимальными правами. Даже ещё хуже – на одном рабочем месте все студенты используют единственную «студенческую» учётную запись, то есть множество студентов с точки зрения ОС рабочей станции это один и тот же пользователь (единственный пользователь), в этом случае просто исключается понятие «персональный документ».

Для демонстрации процесса изменения даже сетевых настроек системы студенту требуется предоставить административные права в системе. При этом учащийся всегда может ошибиться, что в итоге в лучшем случае приведёт к неработоспособности локальной сети в пределах лаборатории.

Как же дать возможность и экспериментировать, и ошибаться, и получать результаты, не подвергая риску преждевременного обрушения инфраструктуру локальной сети?

Соответственно первая задача – как дать условия студенту для экспериментов в рамках изучаемого предмета?

Вторая задача – как в пределах одного компьютера под управлением MS Windows 7 (8, 10 и т.д.), желательно оставив его работоспособным, продемонстрировать другие, возможно перспективные, версии операционных систем и дать навыки работы с ними?

Третья задача – всегда хочется оптимизировать трудозатраты на подготовку лаборатории к работе, но при этом в начале семестра компьютеры в учебных лабораториях часто оказываются очищенными от ранее установленного программного обеспечения и документов в ходе возможных ремонтов, переустановок операционных систем и других причин. Поэтому приходится, засучив рукава, помогать лаборантам устанавливать необходимый комплект программ. Та самая типовая цепочка действий «далее, далее, далее, отказаться от установки рекламных модулей». Она отнимает, казалось бы немного времени, но если умножить затраченное время на количество компьютеров, то трудозатраты внезапно начинают занимать уже не минуты, а часы.

Эта же проблема с другой стороны – мир не стоит на месте, развивается. Выходят новые программные продукты, существующее программное обеспечение постоянно обновляется. Если не бежать вслед за прогрессом, то студенты будут получать устаревшие знания и не будут конкурентоспособны на рынке труда. Соответственно требуется менять содержимое лекций и практических заданий. Требуется устанавливать всё более современные версии программного обеспечения, а это снова выливается в дополнительные часы работы.

Варианты решения

Можно создать специализированный дистрибутив для автоматической установки, что несколько облегчает задачу, а можно организовать автоматическую установку всего программного обеспечения по сети (если она есть), но для этого надо настроить каждую рабочую станцию для загрузки по сети и подготовить соответствующий сервер, с которого рабочие станции будут загружаться и устанавливать ПО централизованно, что влечет за собой дополнительные затраты на сервер и серверное ПО.

Возможен вариант с использованием облачных технологий. Например, поднимается терминальный сервер MS Windows, на котором запускается служба APP-V, в которой публикуются нужные приложения. На рабочих станциях, на рабочих столах размещаются ярлыки для запуска приложений APP-V. С точки зрения пользователя они выглядят почти как выполняющиеся локально (хотя могут выполняться на удалённом сервере, а на стороне клиента пользователь видит только элементы интерфейса приложения). Обновить приложение достаточно только на стороне сервера.

Так же хотелось бы отметить положительный опыт коллег из Южно-Уральского государственного университета. Суть их работы заключается в том, что для каждого студента при поступлении создается персональный виртуальный компьютер с индивидуальным профилем, который постепенно наполняется необходимыми ему учебными сервисами по мере освоения им образовательной программы. После окончания вуза выпускник может продолжать использовать персональный виртуальный компьютер в рамках виртуального сообщества выпускников. Платформа развернута на базе суперкомпьютера «СКИФ-Аврора» ЮУрГУ[1].

При отсутствии суперкомпьютера или хотя бы фермы серверов, идеальный сценарий видится примерно так – вставили флэш-диск в сервер, скопировали с него заранее подготовленный в спокойной обстановке пакет программ и он самостоятельно (или при минимальном участии) установился на компьютеры учащихся и уже готов к работе.

Одному из вариантов приближения к этому идеальному сценарию и посвящается данный доклад.

Что такое виртуализация?

Для начала осветим, что же такое технология виртуализации.

Технология виртуализации – данный термин означает абстракцию вычислительных ресурсов при которой пользователь получает в использование систему, реализация которой инкапсулирована[2].

В основе виртуализации лежит возможность одного компьютера выполнять работу нескольких компьютеров благодаря распределению его ресурсов по нескольким средам. Несколько ОС и приложений возможно разместить в едином местоположении при помощи виртуальных серверов и виртуальных настольных компьютеров. При таком варианте мы можем пренебречь всеми географическими и физическими ограничениями.

Помимо того, что сократятся расходы и уменьшится энергопотребление, виртуальная инфраструктура предоставит высокий уровень доступности необходимых ресурсов, а так же более эффективные возможности для управления, возможности восстановления в критических ситуациях и повышенную безопасность. Это достигается благодаря более эффективному использованию аппаратных ресурсов[3].

Итак, к преимуществам виртуализации в общем виде можно отнести:

1. Эффективное использование вычислительных ресурсов. Сюда можно отнести как уменьшение числа компьютеров, которые необходимо обслуживать; так и экономия электроэнергии.
2. Сокращение расходов на инфраструктуру. Соответственно дает возможность уменьшить количество серверов и необходимого оборудования.
3. Снижение затрат на программное обеспечение. Уже большое количество производителей ПО предлагают схемы лицензирования, которые подготовлены специально для виртуальных сред.
4. Повышение гибкости и скорости реагирования системы. Благодаря этому появляется возможность выполнять повторяющиеся задачи за меньшее время[4].

Стоит упомянуть то, с помощью чего осуществляется данная абстракция физической аппаратной части, а именно гипервизор. Это программное обеспечение, которое позволяет на одной физической машине одновременно работать нескольким операционным системам и напрямую взаимодействовать им с физической аппаратной частью.

Сам гипервизор состоит из двух компонентов: ядро и монитор виртуальной машины, а так же он бывает двух видов:

1. Bare-metal – данный тип устанавливается на «чистое железо» и является сам операционной системой. Так же он требует сертифицированной аппаратной части.
2. Hosted – является внешним гипервизором, который устанавливается и запускается как приложение. При работе с физическими устройствами данный тип полагается на операционную систему, которая запущена на физической машине.

В итоге до использования виртуализации мы имеем только одну операционную систему в один момент времени на одной рабочей машине,

привязку программного и аппаратного обеспечения друг к другу, частое недоиспользование ресурсов рабочей машины и низкую гибкость структуры. После внедрения виртуализации мы получаем возможности для использования нескольких операционных систем на одной машине в момент времени, независимость приложений и аппаратной части, значительно улучшенное распределение ресурсов, гибкость всей инфраструктуры и, как правило, более дешевую стоимость при схожих возможностях.

ТРИЗ: ключевое противоречие задачи

Как следует из теории ТРИЗ[5], необходимо найти в задаче ключевое противоречие и уже именно его решать. Главным можно принять противоречие «один компьютер / много компьютеров». Как же, имея всего один компьютер, обеспечить одновременную работу с несколькими компьютерами?

Решение достаточно простое и оно уже прозвучало в данной публикации – виртуализация. Эта технология уже более десяти лет как «переоткрыта», стала набирать известность, активно используется и внедряется на производстве и, несколько реже, в обучении.

Небольшой экскурс в историю

В образовательном процессе, тогда ещё в УПИ (физ.-тех., каф. выч. техники) виртуализацию использовали ещё в конце 80 – начале 90 г.г. прошлого века. Тогда проводили занятия со студентами на ЭВМ ЕС-1046 работавшей под управлением операционной системы «Система виртуальных машин ЕС ЭВМ» (СВМ ЕС – «в девичестве» IBM VM/SP rel 6.0, теперь zVM). При проведении занятий использовались следующие сценарии.

Один студент – одна виртуальная машина

Каждому студенту предоставлялась собственная виртуальная машина под управлением «Подсистемы диалоговой обработки» (ПДО – «в девичестве» IBM CMS, на эту систему очень похожа была ОС MS|PC DOS что собственно и обусловило выбор DOS как системы для IBM PC). На этой виртуальной машине проходили занятия по обучению программированию (FORTRAN, PL/1, COBOL, REXX, Assembler (ЕС ЭВМ) и др.).

Одна виртуальная машина – много студентов

При этом подходе терминалы подключались к одной виртуальной машине, а студенты использовали свои учетные записи внутри гостевой ОС этой виртуальной машины. Так проводились занятия, например, по изучению UNIX-подобных ОС. На виртуальной машине запускалась UNIX – подобная ОС МОС ЕС (IBM AIX 370) студенты знакомились с работой в среде этой ОС, получали навыки программирования на языке С и изучали СУБД INGRES.

Аналогичным образом проводились занятия по ОС ЕС (IBM MVS).

Несколько студентов – несколько виртуальных машин

Когда каждый студент работает в среде ПДО своей виртуальной машины и из этой среды обращается за обслуживанием к общим сервисным виртуальным машинам. В таком подходе изучалась работа с СУБД SQL/DS и системами пакетной обработки.

К сожалению, полученный опыт был утрачен практически полностью вместе с культурой применения систем коллективного пользования и сейчас мы уже говорим не о восстановлении, а о создании с нуля.

Что можно сделать сейчас?

Процесс подготовки к работе выглядит достаточно просто – требуется создать пустую виртуальную машину и сразу задать её необходимые характеристики (размер жесткого диска, выделяемую оперативную память, устройства и проч.), а дальше внутрь неё установить необходимую операционную систему и набор необходимого программного обеспечения.

Преимущества подхода

С точки зрения внешней операционной системы виртуальная машина представляет из себя набор нескольких файлов, которые можно легко копировать по сети. Также программное обеспечение для виртуализации достаточно только один раз установить на целевую учебную машину и потом просто распространять по сети файлы собственно виртуальной машины.

Еще одно преимущество – операционная система внутри виртуальной машины оказывается заперта в «клетке» и поэтому сильно ограничена в возможностях деструктивного воздействия на внешнюю систему. Достигается очень жесткая изоляция систем. Можно позволить студентам проводить «опасные» лабораторные работы в том числе и с вредоносным ПО.

В пределах обычного компьютера можно запустить небольшую виртуальную локальную сеть и проводить серии экспериментов, при этом обладая полным доступом к любой из виртуальных систем. Можно один и тот же экземпляр виртуальной машины скопировать в несколько экземпляров и запускать одновременно. Они даже не будут при этом друг другу мешать.

Сценарии использования

Очень легко можно запустить ОС Linux в окне MS Windows и изучать методы работы или поднять в пределах одного компьютера небольшую ферму серверов и экспериментировать с методами централизованного управления.

В рамках информационной безопасности требуется изучать защиту и нападение. Слишком часто требуется изучать методы нападения, чтобы прочувствовать в полной мере необходимость адекватной защиты. Причем основная проблема в том, что многие техники виртуального нападения вполне успешно могут выводить из строя реальную локальную сеть или привести к отказу в обслуживании серверов. Иллюстрация простоты некоторых атак (напомню тот же на шумевший недавно Heartbleed) отлично закрепляет на уровне, близком к подсознательному, понимание важности эффективной защиты и укрепляет стремление к знаниям.

Или взять те же компьютерные вирусы. Не заражать же реальный компьютер для демонстрации методов лечения. И здесь тоже будет полезна виртуальная песочница, в пределах которой можно проводить виртуальные атаки и отрабатывать методы защиты.

Недостатки метода

У предлагаемого мной метода есть и недостатки. Главный из которых – каждый виртуальный компьютер требует вполне реальных ресурсов, причем и операционная система, под которой запущен гипервизор – тоже потребляет ресурсы. Следствием этого будут повышенные системные требования к рабочему компьютеру. Самым важным ресурсом, нехватка которого будет заметна в первую очередь – является оперативная память. Это следует учитывать при планировании работ. Очень полезной будет поддержка технологий аппаратной виртуализации центральным процессором.

Опять же, есть вполне успешный опыт работы с тремя виртуальными машинами одновременно под управлением ОС Windows 7, процессором Intel Core 2 Duo и 2 гигабайтами оперативной памяти.

Производительность локальной сети и время

Есть еще одно узкое место – файлы виртуальной машины занимают достаточно приличный объем (от 3 гигабайт), а при начале преподавания новой темы требуется оснащать компьютеры обновленными материалами, что способно отнимать достаточно много времени. В среднем скорость работы локальных сетей пока держится на отметке 100 мегабит, что в условиях близких к идеальным даёт пропускную способность порядка 10 мегабайт в секунду, соответственно минимальное время копирования на один компьютер по сети составит около 5 минут. Можно сжать передаваемые файлы, тогда этап копирования занимает порядка 2 минут, плюс время на распаковку на рабочей станции ученика тоже около 2 минут.

Казалось бы, требуется немного времени, но следует учесть, что при размере лаборатории в 30 компьютеров, суммарное время уже достигает порядка одного часа, поскольку при одновременном копировании с сервера по

общему каналу – пропускная способность канала делится между всеми участниками.

Получается, что идея по ускорению работы приводит к замедлению? Приходит преподаватель на занятие, копирует данные на сервер и в течение часа ждёт, пока все не смогут начать работу?

Торренты только для пиратов?

К счастью, используя опыт работы распределенных систем, а именно наработки в части протокола BitTorrent можно резко ускорить процесс. Суть идеи заключается в использовании бесплатного программного обеспечения BitTorrent Sync[6], которое устанавливается на компьютеры всех участников и настраивается для работы в режиме Master-Slave.

В этом случае компьютер преподавателя становится главным и активируется синхронизация всех новых или измененных файлов из выделенного каталога на подключенные по локальной сети компьютеры учащихся.

Ключевое отличие от простого копирования файла по сети, когда пропускная способность делится между всеми участниками, заключается в реализации протокола BitTorrent. Суть её заключается в 3 идеях:

- 1) каждый клиент может не только копировать с сервера, но и выступать в роли сервера сам, раздавая уже скачанное;
- 2) клиенты самостоятельно обмениваются между собой информацией об уже скачанных фрагментах;
- 3) клиент может начать отдавать скачанные файлы по фрагментам, скачав с сервера уже всего лишь единичный фрагмент файла.

В итоге фактически сервер отдаёт только один экземпляр необходимых файлов всем участникам, а они, уже в процессе скачивания, начинают обмениваться скачанными фрагментами между собой напрямую, снижая нагрузку на сервер. Протокол обладает некоторой способностью к самоорганизации, поэтому уже скачанные фрагменты могут повторно с сервера не запрашиваться, а будет выбираться любой другой участник. Теоретическая пиковая пропускная способность для 30 участников в итоге может достигать 3 Гигабит в секунду.

На практике весь процесс занимает примерно на 20% дольше по времени, чем копирование с сервера на одиночный компьютер.

Как всё реализовано на практике?

В качестве гипервизора на рабочих местах используется MS Virtual PC. Его выбор обусловлен тем, что он поставляется вместе с MS Windows, без дополнительной оплаты, а эмулируемое им аппаратное обеспечение поддерживается всеми основными операционными системами.

В качестве транспорта для распространения файлов используется BitTorrent Sync. Его лицензия позволяет использовать ПО в любых целях, включая коммерческие. Ограничения, налагаемые редакцией BitTorrent Sync Free незначительны при данном сценарии использования. Возможной заменой

может служить Open Source ПО Syncthing[7], распространяемое под лицензией Mozilla Public License.

В качестве вспомогательного элемента при развёртывании на конечной рабочей станции пакета с виртуальной машиной используется Open Source архиватор 7-Zip, распространяемый под лицензией GNU LGPL, которая позволяет использовать его для любых целей, включая коммерческое использование. Он обеспечивает решение 2 задач – уменьшение объема файлов, распространяемых по сети и размещение распакованных файлов в нужном месте при помощи функционала самораспаковывающихся архивов, фактически упрощенных инсталляционных пакетов.

Итог – суммарная стоимость дополнительного аппаратного и программного обеспечения 0 рублей.

Каждая рабочая станция требует одноразового проведения следующих работ:

1. Установить обновления для добавления поддержки Virtual PC (2 файла).
2. Установить ПО BitTorrent Sync (1 файл), активировать синхронизацию с сервером (33 байта, можно автоматизировать).

Процесс работы по обновлению:

0. Преподаватель готовит (или дорабатывает) образ виртуальной машины.
1. Скопировать на сервер в выделенный каталог необходимые файлы (плюс подождать пару минут).
2. Ученик запускает со своего компьютера новый файл на выполнение.

В данной статье раскрыта только одна сторона использования виртуализации, но уже ведутся работы в сторону других сценариев использования.

Перспективы использования

Обучение параллельному программированию. Каждому студенту выдается кластер на основе нескольких виртуальных машин, на котором он и тренируется, не мешая всем остальным. Так как для отладки на таких кластерах могут требоваться права администратора, такие права пользователю могут быть предоставлены, и он сможет эффективно производить отладку. Дополнительный бонус – возможность освоить администрирование такого кластера.

Моделирование IT инфраструктуры. Студенты (индивидуально или в малых группах) с начала обучения получают модель IT инфраструктуры предприятия на основе набора виртуальных машин в облачном окружении и само «виртуальное предприятие» в интересах которого далее развивают сервисы внутри этой инфраструктуры по мере прохождения профильных учебных курсов (базы данных, 1С, информационные системы, SharePoint и т.п.)

Так же следует учесть, что имеющиеся виртуальные машины относительно просто могут быть перенесены в среду облачных вычислений.

Список литературных источников

1. РСК и Intel внедрили "облачную" образовательную платформу в ЮУрГУ. [Электронный ресурс] // Режим доступа: <http://www.digit.ru/development/20110526/382169074.html>
2. Технологии виртуализации. [Электронный ресурс] // Режим доступа: <http://www.intuit.ru/studies/courses/673/529/lecture/11915>
3. Виртуализация серверов. [Электронный ресурс] // Режим доступа: <http://www.osp.ru/cw/2007/09/4038424/>
4. Технологии виртуализации. [Электронный ресурс] // Режим доступа: <http://www.sks-vm.com/?q=node/26>
5. ТРИЗ | Официальный Фонд Г.С. Альтшуллера. [Электронный ресурс] // Режим доступа: <http://www.altshuller.ru/triz/>
6. BitTorrent Sync . [Электронный ресурс] // Режим доступа: <http://www.getsync.com/>
7. Syncthing . [Электронный ресурс] // Режим доступа: <https://syncthing.net/>

ПРОВЕРКА ГИПОТЕЗЫ О ВЛИЯНИИ УРОВНЯ ЭКОНОМИЧЕСКОЙ СВОБОДЫ НА ПРЯМЫЕ ИНОСТРАННЫЕ ИНВЕСТИЦИИ С ИСПОЛЬЗОВАНИЕМ МЕТОДОВ ИНТЕЛЛЕКТУАЛЬНОГО АНАЛИЗА ДАННЫХ

УДК 004.89

Комоцкий Е.И. преподаватель
кафедра анализа систем и принятия решений
Уральский федеральный университет, ВШЭМ

Аннотация. В работе выполняется проверка либеральной теории о влиянии уровня экономических свобод на приток прямых иностранных инвестиций с помощью методов интеллектуального анализа данных.

В качестве источника данных использован Индекс Экономической Свободы (Index Of Economic Freedom) за 2013 год, кластеризация выполнена с использованием метода самоорганизующихся карт Кохонена.

Базовым положением исследования являлся тот факт, что во времена формирования этой теории пул стран, представляющих инвестиционную привлекательность, был ограничен, и в него входили только высокоразвитые страны Запада и их сателлиты. Однако, в конце XX – начале XXI века мир изменился, и его политическая и экономическая структура претерпела значительные изменения, вследствие чего данная теория утратила адекватность.

Полученные результаты позволяют утверждать, что теория о решающем влиянии уровня экономических свобод на приток иностранных инвестиций утратила адекватность в процессе формирования многополярного мира и развития экономик и финансовых институтов бывших колониальных или отсталых стран, а также осколков СССР.