Виртуализация

Виртуализа́ция — предоставление набора вычислительных ресурсов или их логического объединения, абстрагированное от аппаратной реализации, и обеспечивающее при этом логическую изоляцию друг от друга вычислительных процессов, выполняемых на одном физическом ресурсе.

Примером использования виртуализации является возможность запуска нескольких операционных систем на одном компьютере: при том каждый из экземпляров таких гостевых операционных систем работает со своим набором логических ресурсов (процессорных, оперативной памяти, устройств хранения), предоставлением которых из общего пула, доступного на уровне оборудования, управляет хостовая операционная система — гипервизор. Также могут быть подвергнуты виртуализации сети передачи данных, сети хранения данных, платформенное и прикладное программное обеспечение.

Оборудование

- Эмуляция — полная виртуализация (виртуализация всей платформы); например эмуляторы игровых консолей.

Операционные системы

- Программная виртуализация
  - Динамическая трансляция; при динамической (бинарной) трансляции проблемные команды гостевой операционной системы перехватываются гипервизором.
  - Паравиртуализация: операционная система взаимодействует с программой гипервизора, который предоставляет ей гостевой API, вместо использования напрямую таких ресурсов, как таблица страниц памяти.
    - Встроенная виртуализация
- Аппаратная виртуализация — виртуализация с поддержкой специальной процессорной архитектуры. В отличие от программной виртуализации, с помощью данной техники возможно использование изолированных гостевых систем, управляемых гипервизором напрямую.
- Виртуализация на уровне операционной системы: работа нескольких экземпляров пространства пользователя в рамках одной ОС. Примерами могут быть Docker, LXC

Программное обеспечение

- Виртуализация приложений (также виртуализация рабочего окружения): работа отдельных приложений в среде, отделённой от основной ОС. Эта концепция тесно связана с портативными приложениями. Примерами могут быть: Citrix XenApp, Microsoft App-V.
- Виртуализация сервисов: эмуляция поведения системных компонентов, необходимых для запуска приложения в целях отладки и тестирования (англ. *Application Under Test*). Вместо виртуализации компонентов целиком, эта технология виртуализует только необходимые части. Примеры: SoapUI, Parasoft Virtualize.

Память

- Виртуализация памяти (*memory virtualization*) — объединением оперативной памяти из различных ресурсов в единый массив. Реализации: Oracle Coherence, GigaSpaces XAP.
- Виртуальная память — изоляция адресного пространства приложения от всего адресного пространства. Применяется во всех современных ОС.

Хранилище данных

- Виртуализация хранения данных, представление набора физических носителей в виде единого физического носителя.
  - Блочная виртуализация
  - Файловая виртуализация
- Распределённая файловая система — любая файловая система, которая позволяет получать доступ к файлам с нескольких устройств, с помощью компьютерной сети.

- Виртуальная файловая система (***V**irtual **F**ile **S**ystem*) — уровень абстракции поверх конкретной реализации файловой системы. Целью VFS является обеспечение единообразного доступа клиентских приложений к различным типам файловых систем.
- Гипервизор хранения данных (любая файловая система, которая позволяет получать доступ к файлам с нескольких устройств, с помощью компьютерной сети.
- Виртуальная файловая система (*storage hypervisor*) — программа, которая управляет виртуализацией пространства для хранения данных и может объединять раличные физические пространства в единый логический массив.
- Виртуализация устройств хранения данных: виртуализация жёсткого (логический диск) или оптического диска (например, DAEMON Tools).

База данных

- Виртуализация данных (*data virtualization*)— представление данных в абстрактном виде, независимо от нижележащих систем управления и хранения данных, а также их структуры.
- Виртуализация баз данных.

Сеть

- Виртуализация сети (*network virtualization*) — процесс объединения аппаратных и программных сетевых ресурсов в единую виртуальную сеть.
  - Внешняя, соединяющая множество сетей в одну виртуальную.
  - Внутренняя, создающая виртуальную сеть между программными контейнерами внутри одной системы.
- Виртуальная частная сеть (*virtual private network*) — обеспечение одного или нескольких сетевых соединений поверх другой сети.

Виртуализация операционных систем

Для виртуализации операционных систем применяется серия подходов, которые по типу реализации подразделяются на программные и аппаратные.

Программная виртуализация

Динамическая трансляция

При динамической (*бинарной*) трансляции проблемные команды гостевой операционной системы перехватываются гипервизором. После того как эти команды заменяются на безопасные, происходит возврат управления гостевой системе.

Паравиртуализация

Паравиртуализация — техника виртуализации, при которой гостевые операционные системы подготавливаются для исполнения в виртуализированной среде, для чего их ядро незначительно модифицируется. Операционная система взаимодействует с программой гипервизора, который предоставляет ей гостевой API, вместо использования напрямую таких ресурсов, как таблица страниц памяти.

Метод паравиртуализации позволяет добиться более высокой производительности, чем метод динамической трансляции.

Метод паравиртуализации применим лишь в том случае, если гостевые операционные системы имеют открытые исходные коды, которые можно модифицировать согласно лицензии, или же гипервизор и гостевая операционная система разработаны одним производителем с учётом возможности паравиртуализации гостевой системы (хотя при условии того, что под гипервизором может быть запущен гипервизор более низкого уровня, то и паравиртуализации самого гипервизора).

Встроенная виртуализация

Преимущества:

- Совместное использование ресурсов несколькими гостевыми операционными системами (каталоги, принтеры и так далее).
- Удобство интерфейса для окон приложений из разных систем (перекрывающиеся окна приложений, одинаковая минимизация окон, как в хост-системе).
- При тонкой настройке на аппаратную платформу производительность мало отличается от оригинальной операционной системы. Быстрое переключение между системами (менее одной секунды).
- Простая процедура обновления гостевой операционной системы.
- Двухсторонняя виртуализация (приложения одной системы запускаются в другой и наоборот).

Реализации:

- BlueStacks Multi-OS (MOS.

Аппаратная виртуализация

- Упрощение разработки программных платформ виртуализации за счет предоставления аппаратных интерфейсов управления и поддержки виртуальных гостевых систем. Это уменьшает трудоемкость и время на разработку систем виртуализации.
- Возможность увеличения быстродействия платформ виртуализации. Управление виртуальными гостевыми системами осуществляет напрямую небольшой промежуточный слой программного обеспечения, гипервизор, что дает увеличение быстродействия.
- Улучшается защищённость, появляется возможность переключения между несколькими запущенными независимыми платформами виртуализации на аппаратном уровне. Каждая из виртуальных машин может работать независимо, в своем пространстве аппаратных ресурсов, полностью изолированно друг от друга. Это позволяет устранить потери быстродействия на поддержание хостовой платформы и увеличить защищенность.
- Гостевая система становится не привязана к архитектуре хостовой платформы и к реализации платформы виртуализации. Технология аппаратной виртуализации делает возможным запуск 64-битных гостевых систем на 32-битных хостовых системах (с 32-битными средами виртуализации на хостах).

Технологии:

- Режим виртуального 8086 (устарела)
- Intel VT (*VT-x*, *Intel Virtualization Technology for x86*)
- AMD-V

Платформы, использующие аппаратную виртуализацию:

- IBM LPAR
- VMware
- Hyper-V
- Xen
- KVM
- Bhyve

Виртуализация на уровне операционной системы

Виртуализация на уровне операционной системы позволяет запускать изолированные и безопасные виртуальные машины на одном физическом узле, но не позволяет запускать операционные системы с ядрами, отличными от типа ядра базовой операционной системы. При виртуализации на уровне операционной системы не существует отдельного слоя гипервизора. Вместо этого сама хостовая операционная система отвечает за разделение аппаратных ресурсов между несколькими виртуальными машинами и поддержку их независимости друг от друга. Среди реализаций:

- Solaris Containers/Zones

- FreeBSD Jail
- Linux-VServer[en]
- LXC (Linux Containers)
- FreeVPS[en]
- OpenVZ
- Virtuozzo
- iCore Virtual Accounts

Области применения виртуализации

Виртуальные машины

Виртуальная машина — это окружение, которое представляется для «гостевой» операционной системы, как аппаратное. Однако на самом деле это программное окружение, которое эмулируется программным обеспечением хостовой системы. Эта эмуляция должна быть достаточно надёжной, чтобы драйверы гостевой системы могли стабильно работать. При использовании паравиртуализации, виртуальная машина не эмулирует аппаратное обеспечение, а, вместо этого, предлагает использовать специальный API.

Примеры применения:

- Тестовые лаборатории и обучение: тестированию в виртуальных машинах удобно подвергать приложения, влияющие на настройки операционных систем, например инсталляционные приложения. За счёт простоты в развёртывании виртуальных машин, они часто используются для обучения новым продуктам и технологиям.
- Распространение предустановленного программного обеспечения: многие разработчики программных продуктов создают готовые образы виртуальных машин с предустановленными продуктами и предоставляют их на бесплатной или коммерческой основе. Такие услуги предоставляют Vmware VMTN или Parallels PTN.

Виртуализация ресурсов

Виртуализация ресурсов (или разделение ресурсов, англ. *partitioning*) может быть представлена как разделение одного физического узла на несколько частей, каждая из которых видна для владельца в качестве отдельного сервера. Не является технологией виртуальных машин, осуществляется на уровне ядра операционной системы.

В системах с гипервизором второго типа обе операционные системы (гостевая и гипервизора) отнимают физические ресурсы, и требуют отдельного лицензирования. Виртуальные серверы, работающие на уровне ядра ОС, почти не теряют в быстродействии, что дает возможность запускать на одном физическом сервере сотни виртуальных, не требующих дополнительных лицензий.

Дисковое пространство или пропускной канал сети разделены на некоторое количество меньших составляющих, и потому легче используемых ресурсов того же типа.

Например, к реализации разделения ресурсов можно отнести OpenSolaris Network Virtualization and Resource Control (Проект Crossbow), позволяющий создавать несколько виртуальных сетевых интерфейсов на основе одного физического.

Агрегация, распределение или добавление множества ресурсов в большие ресурсы или объединение ресурсов. Например, симметричные мультипроцессорные системы объединяют множество процессоров; RAID и дисковые менеджеры объединяют множество дисков в один большой логический диск; RAID и сетевое оборудование использует множество каналов, объединённых так, чтобы они представлялись, как единый широкополосный канал. На мета-уровне компьютерные кластеры делают все вышеперечисленное. Иногда сюда же относят сетевые файловые системы абстрагированные от хранилищ данных на которых они построены, например, Vmware VMFS, Solaris/OpenSolaris ZFS, NetApp WAFL.

Виртуализация приложений

Виртуализация приложений — процесс использования приложения, преобразованного из требующего установки в операционную систему в не требующее (требуется только запустить). Для виртуализации приложений программное обеспечение виртуализатора определяет при установке виртуализуемого приложения, какие требуются компоненты ОС, и эмулирует их. Таким образом, создаётся необходимая специализированная среда для конкретно этого виртуализируемого приложения и, тем самым, обеспечивается изолированность работы этого приложения. Для создания виртуального приложения виртуализируемое помещается в контейнер, оформленный, как правило, в виде папки. При запуске виртуального приложения запускается виртуализируемое приложение и контейнер, являющийся для него рабочей средой. Рабочая среда запускается и предоставляет локальные ранее созданные ресурсы, которое включает в себя ключи реестра, файлы и другие компоненты, необходимые для запуска и работы приложения. Такая виртуальная среда работает как прослойка между приложением и операционной системой, что позволяет избежать конфликтов между приложениями. Виртуализацию приложений обеспечивают, например, программы Citrix XenApp[5], SoftGrid[6] и VMware ThinApp.

Достоинства:

- изолированность исполнения приложений: отсутствие несовместимостей и конфликтов;
- каждый раз в первозданном виде: не загромождается реестр, нет конфигурационных файлов — необходимо для сервера;
- меньшие ресурсозатраты по сравнению с эмуляцией всей операционной системы.

Аппаратная виртуализация

Аппаратная виртуализация — виртуализация с поддержкой специальной процессорной архитектуры. В отличие от программной виртуализации, с помощью данной техники возможно использование изолированных гостевых систем, управляемых гипервизором напрямую. Гостевая система не зависит от архитектуры хостовой платформы и реализации платформы виртуализации.

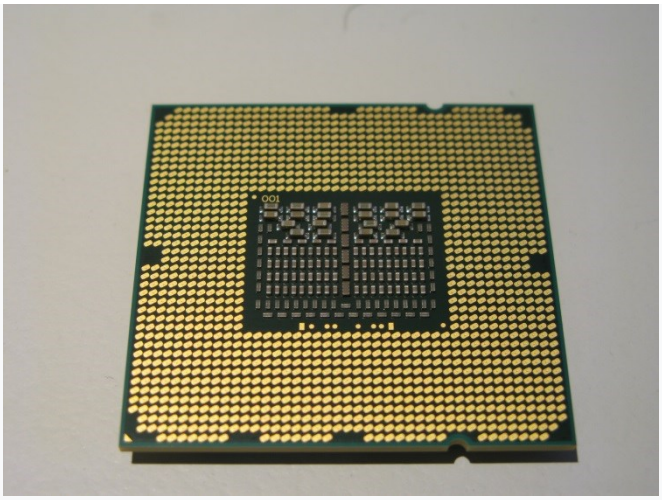
Аппаратная виртуализация обеспечивает производительность, сравнимую с производительностью невиртуализованной машины, что дает виртуализации возможность практического использования и влечет её широкое распространение. Наиболее распространены технологии виртуализации Intel-VT и AMD-V.

- В **Intel VT** (Intel Virtualization Technology) реализована виртуализация режима реальной адресации (режим совместимости с 8086). Соответствующая аппаратная виртуализация ввода-вывода — VT-d (кодовое название — **Vanderpool**). Часто обозначается аббревиатурой *VMX* (Virtual Machine eXtension).

- **AMD-V** часто обозначается аббревиатурой SVM (Secure Virtual Machines). Кодовое название — Pacifica. Соответствующая технология виртуализации ввода-вывода — IOMMU. AMD-V проще и эффективнее, чем Intel VT.[1] Поддержка AMD-V появилась в Xen 3.3.

Intel VT (Intel Virtualization Technology)

VT-x



Intel Core i7 (Bloomfield) CPU

Ранее известная под кодовым названием «Vanderpool», VT-x представляет собой технологию виртуализации Intel на платформе x86. 13 ноября 2005 года Intel выпустила две модели Pentium 4 (модели 662 и 672), которые стали первыми процессорами, поддерживающими VT-x. Флаг поддержки VT-x — «vmx»; в Linux проверяется командой `cat /proc/cpuinfo`, в Mac OS X — `sysctl machdep.cpu.features`.[2]

По состоянию на 2015 год не все процессоры Intel поддерживают VT-x, что используется компанией Intel для сегментирования своего рынка.[3] Поддержка VT-x может различаться даже между различными версиями (которые идентифицируются по *sSpec Number*) одной и той же модели.[4][5] Полный список можно посмотреть на сайте Intel.[6] Даже в мае 2011 года процессор Intel P6100, используемый в ноутбуках, не поддерживает аппаратную виртуализацию.[7]

На некоторых материнских платах пользователи должны вручную включить виртуализацию VT-x в настройках BIOS.[8]

Intel начала включать технологию виртуализации Extended Page Table (EPT)[9] для страничных таблиц[10], начиная с процессоров архитектуры Nehalem, выпущенных в 2008 году[11][12].

В 2010 году в архитектуру Westmere была добавлена технология «неограниченного гостя», заключающаяся в поддержке логического процессора в реальном режиме и требующая для работы EPT.[13][14]

Начиная с архитектуры Haswell, объявленной в 2013 году, Intel начала включать *затенение VMCS* — технологию, ускоряющую вложенную виртуализацию гипервизоров.[15]VMCS — *структура управления виртуальной машины* (virtual machine control structure) — структура данных в памяти, существующая в точности в одном экземпляре на одну виртуальную машину и управляемая гипервизором. С каждым изменением контекста выполнения между разными ВМ структура данных VMCS восстанавливается для текущей виртуальной машины, определяя состояние виртуального процессора ВМ.[16] Если используется больше гипервизора или используются вложенные гипервизоры, необходимо многократное затенение VMCS. Аппаратная поддержка затенения делает управление VMSC более эффективным.

VT-d

VT-d (Virtualization technology for directed I/O) — технология виртуализации ввода-вывода, созданная корпорацией Intel в дополнение к её технологии виртуализации вычислений (VT), известной под кодовым названием Vanderpool. Виртуализация ввода-вывода позволяет пробрасывать (pass-through) устройства на шине PCI (и более современных подобных шинах) в гостевую ОС, таким образом, что она может работать с ним с помощью своих штатных средств. Чтобы такое было возможно, в логических схемах системной платы используется специальное устройство управления

памятью ввода-вывода (IOMMU), работающее аналогично MMU центрального процессора, используя таблицы страниц и специальную таблицу отображения DMA (DMA remapping table — DMAR), которую гипервизор получает от BIOS через ACPI. Отображение DMA необходимо, поскольку гипервизор ничего не знает о специфике работы устройства с памятью по физическим адресам, которые известны лишь драйверу. С помощью DMAR он создает таблицы отображения таким образом, что драйвер гостевой ОС видит виртуальные адреса IOMMU аналогично тому, как бы он видел физические без него и гипервизора.

Intel Virtualization Technology for Directed I/O (VT-d) — это следующий важный шаг на пути к всеобъемлющей аппаратной поддержке виртуализации платформ на базе Intel. VT-d расширяет возможности технологии Virtualization Technology (VT), существующей в IA-32 (VT-x) и Itanium (VT-i), и добавляет поддержку виртуализации новых устройств ввода-вывода.

Поддержка аппаратным обеспечением

- Виртуализация ввода-вывода впервые появилась в чипсете Q35, и на сегодняшний день поддерживается всеми материнскими платами, поддерживающими технологию Intel vPro.
- Для использования Intel Virtualization Technology необходим компьютер с процессором Intel, BIOS, монитором виртуальных машин (VMM), а для некоторых моделей с определенным программным обеспечением с поддержкой этой технологии. Функциональные возможности, производительность и другие характеристики могут различаться в зависимости от аппаратного и программного обеспечения и могут потребовать обновления BIOS.
- Процессоры, поддерживающие Virtualization Technology for Directed I/O: Intel Core i7-920, Intel Core i7-940, Intel Core i7-950, Intel Core i7-870, Intel Core i7-860, Intel Core i5-650, Intel Core i5-660, Intel Core i5-670, Intel Core i5-540M, Intel Core i5-520M и т. д. [1]
- i7-920 поддерживает технологию VT-x, про VT-d на оф. сайте не указано.[17]

Поддержка программным обеспечением

- Гипервизор Xen поддерживает DMAR начиная с версии 3.3 для аппаратно-виртуализуемых доменов. Для паравиртуальных доменов отображение DMA не требуется.
- В ближайшем будущем заявлена поддержка технологии ПО Oracle VirtualBox.
- Ядро Linux экспериментально поддерживает DMAR начиная с версии 2.6.28, что позволяет встроенному гипервизору (kvm) давать доступ виртуальным машинам к PCI-устройствам.
- Поддержка Intel VT-d есть в Parallels Workstation 4.0 Extreme [2] и в Parallels Server 4 Bare Metal [3]

AMD virtualization (AMD-V)

AMD разработала свои расширения виртуализации первого поколения под кодовым названием «Pacifica», и первоначально опубликовала их как AMD Secure Virtual Machine (SVM)[18], но позже, на рынке, — под торговой маркой «AMD Virtualization», сокращенно «AMD-V».

23 мая 2006 года AMD выпустила Athlon 64 («Orleans»), Athlon 64 X2 («Windsor») и Athlon 64 FX («Windsor») в качестве первых процессоров AMD с поддержкой данной технологии.

Поддержка AMD-V также обеспечивается в семействе процессоров Athlon 64 и Athlon 64 X2 ревизий «F» или «G» на Socket AM2, Turion 64 X2, и Opteron второго поколения[19] и третьего поколения[20], а также процессорами Phenom и Phenom II. Только две модели Sempron поддерживают её: Huron and Sargas. Процессоры AMD Fusion также поддерживают AMD-V.
AMD-V не поддерживается в процессорах на Socket 939.

Процессоры Opteron, начиная с семейства 0x10 Barcelona, и процессоры Phenom II поддерживают второе поколение аппаратной виртуализации технология под названием Rapid Virtualization Indexing (ранее известная как Nested Page Tables во время его разработки), позже адаптированные Intel, как Extended Page Tables (EPT).

Наличие технологии AMD-V в процессоре определяется флагом «svm». Его можно проверить во FreeBSD через dmesg или sysctl, а в Linux — через `cat /proc/cpuinfo`.[21]

Паравиртуализация

Паравиртуализация (*Paravirtualization*) — техника виртуализации, при которой гостевые операционные системы подготавливаются для исполнения в виртуализированной среде, для чего их ядро незначительно модифицируется. Операционная система взаимодействует с программой гипервизора, который предоставляет ей гостевой API, вместо использования напрямую таких ресурсов, как таблица страниц памяти. Код, касающийся виртуализации, локализуется непосредственно в операционную систему. Паравиртуализация таким образом требует, чтобы гостевая операционная система была изменена для гипервизора, и это является недостатком метода, так как подобное изменение возможно лишь в случае, если гостевые ОС имеют открытые исходные коды, которые можно модифицировать согласно лицензии. Но зато паравиртуализация предлагает производительность почти как у реальной не виртуализированной системы. Как и при полной виртуализации, одновременно могут поддерживаться многочисленные различные операционные системы. Метод паравиртуализации позволяет добиться более высокой производительности, чем метод динамической трансляции.

Цель изменения интерфейса заключается в сокращении доли времени выполнения гостя, отведённого на выполнение операций, которые являются существенно более трудными для запуска в виртуальной среде по сравнению с не-виртуальной средой. Паравиртуализация предоставляет специально установленные обработчики прерываний, чтобы позволить гостю (гостям) и хосту принимать и опознавать эти задачи, которые иначе были бы выполнены в виртуальном домене (где производительность меньше). Таким образом, успешная паравиртуализированная платформа может позволить монитору виртуальных машин (VMM) быть проще (путём перевода выполнения критически важных задач, с виртуального домена к хосту домена) и/или уменьшить общие потери производительности машинного выполнения внутри виртуального гостя.

Впервые термин возник в проекте Denali, а после того, как это слово применили исследователи из компьютерной лаборатории Кембриджского университета в проекте Xen, оно окончательно утвердилось в терминологии. Приставка «пара» в слове паравиртуализация ничего не обозначает, просто авторам данной идеи понадобился новый термин.