УДК 004.318

Кожин А.С. [1,2]

[1] Московский физико-технический институт (государственный университет)

[2] ЗАО “МЦСТ”

Аппаратная поддержка виртуализации вычислительных ресурсов (аналитический обзор)

Технология виртуализации процессоров с различными операционными системами на одном мейнфрейме впервые была реализована в 60-х годах компанией IBM и нашла достаточное применение. К 1980-ому году интерес к виртуализации постепенно ослаб ввиду появления и быстрого распространения персональных компьютеров и снова возник только в начале нового тысячелетия.

Причиной возросшего внимания стало стремление к более эффективному использованию вычислительных мощностей серверов. Постоянное расширение инфраструктуры ИТ привело к значительному увеличению затрат на ее поддержку при том, что средний коэффициент загруженности серверов к тому времени составлял всего 25%. Внедрение программной виртуализации позволяло запускать несколько пользовательских операционных систем одновременно на одном сервере, или формировать серверный кластер из множества компьютеров для решения одной задачи. Такая консолидация вычислительных ресурсов заметно увеличила эффективность вычислительных комплексов.

Повсеместное оснащение серверов средствами программной виртуализации подтолкнуло Intel и AMD добавить в свою архитектуру ее аппаратную поддержку. Предложенные технологии получили названия Intel VT и AMD Pacifica.

В решении Intel гипервизор (монитор виртуальных машин) запускается как приложение базовой операционной системы, а работа с виртуальными машинами осуществляется в режиме VMX. Код гостевой ОС выполняется как непривилегированный и при перехвате потенциально опасной инструкции, которая может повлиять на работу остальных ОС, управление передается гипервизору, который программно имитирует результат ее выполнения. Для каждой виртуальной машины вводится специальная структура данных VMCS, содержащая информацию о состоянии гипервизора и этой машины, условия и причины последнего выхода. Все эти структуры хранятся в оперативной памяти и описываются достаточно сложной машиной конечных состояний. Всего в режиме VMX добавилось 10 новых инструкций, шесть из которых управляют структурой VMCS.

Со своей стороны, компания AMD ввела 8 команд, причем три из них служат для ускорения переключения между гипервизором и виртуальными машинами. Базовая ОС не используется, сам гипервизор уже является системным кодом и исполняет роль ядра некоторой основной операционной системы. Виртуальные машины также описываются структурой данных VMCB, но в отличие от предложенной Intel структуры VMCS она не содержит состояния гипервизора. Для него выделяется отдельная внутренняя память процессора, что позволяет избежать многочисленных копий поля состояния, упрощает работу со структурами VMCB и уменьшает вероятность ошибки при переходах.

Помимо расширения системы команд оба производителя ввели для виртуальных машин технологии аппаратной поддержки ввода-вывода: Intel VT-d и AMD IOMMU. Основные примененные в них решения совпадают и относятся к трансляции виртуальных адресов и обеспечению доступа к памяти по DMA без вмешательства гипервизора.

При трансляции виртуальных адресов используется двухуровневая схема. На первом этапе адрес преобразуется из виртуального в физический с помощью собственной таблицы страниц виртуальной машины. Он находится в виртуальной памяти хост-машины и на следующем этапе транслируется в ее физический адрес с помощью “теневой” таблицы страниц, недоступной виртуальным машинам. Полное преобразование записывается в TLB. Для того чтобы привязать все эти записи к конкретным машинам и не очищать буфер после каждого переключения между гостевыми ОС, к каждой строке TLB приписывается тэг-идентификатор. Механизм доступа по DMA добавляет вектор исключения устройств, который хранится в оперативной памяти. Каждый бит этого вектора определяет разрешение для внешнего устройства на доступ к соответствующей 4-КБ странице памяти.

Большинство из идей, использованных в своих технологиях аппаратной поддержки Intel и AMD, были предложены еще в 60-70 гг. компанией IBM. Однако и сегодня они позволяют существенно повысить быстродействие виртуальных машин и приблизить его к быстродействию систем без виртуализации.

Литература

1. Intel 64 and IA-32 Architectures Software Developer’s Manual // Intel. – 2010. – V. 3B., N. 2.
2. AMD64 Architecture Programmer’s Manual // AMD. – 2009. – V. 2.
3. AMD I/O Virtualization Technology (IOMMU) Specification // AMD. – 2009.