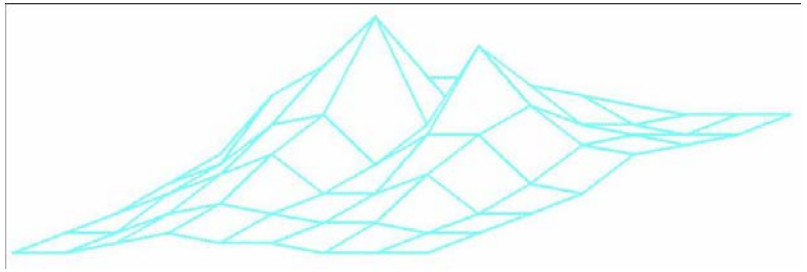


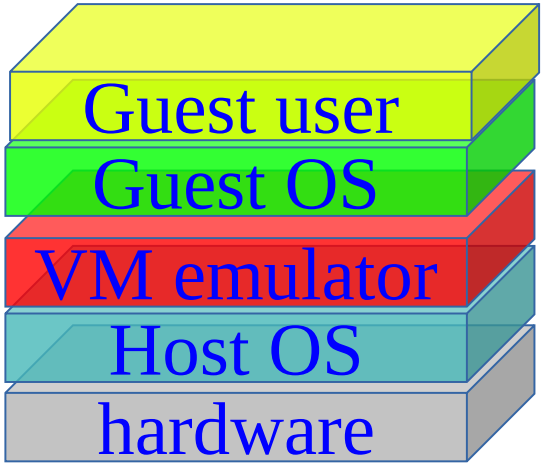
О виртуализации в «Эльбрусе»

2020

Груздов Ф.А.

Что такое виртуализация?

Виртуализация — это эмуляция вычислительной машины с точностью, достаточной для работы ОС и приложений без их модификации.



Зачем нужна виртуализация?

Виртуализация родилась в data centre-ах.
Цели:

- Снижение затрат на одного пользователя
  - Повышение степени утилизации оборудования
  - Упрощение администрирования аппаратуры и ПО
- Повышение надежности и безопасности коллективного использования оборудования
  - ВМ — это изолированная «песочница», не способная навредить другим пользователям и самому базовому оборудованию и ПО
  - ВМ дает возможность создавать контрольную точку всей машины для отката в случае сбоя оборудования или ошибок в программе
  - ВМ можно перемещать с одного сервера на другой при ненадежной работе оборудования
- Предоставление пользователям новых сервисов
  - Возможность работать под той ОС, которая требуется
  - Возможность отлаживать новые ОС и привилегированные приложения
  - Возможность создания ВМ требуемой конфигурации

Как ускорить эмуляцию?

Разрешить ВМ напрямую использовать аппаратуру.

Польза:

- Скорость исполнения

Недостатки:

- Привязка архитектуры ВМ к архитектуре сервера

Проблемы:

- Необходимость встраивания механизмов постановки гостя на процессор и снятия с процессора
- Необходимость встраивания механизмов перехвата некоторых действий гостя
- Необходимость встраивания механизмов аппаратной поддержки виртуализации ресурсов

Механизмы запуска и снятия гостя (1)

Эти механизмы похожи на запуск и приостановку обычного пользовательского приложения:

контекст ОС → контекст польз. → контекст ОС

Отличие состоит в объеме переключаемого контекста — добавляется почти полный набор привилегированных регистров (в Эльбрусе — 16 шт.), например:

- режимы и состояния
- управление стеками
- управление виртуальной памятью

Дополнительная проблема — переключение аппаратных верхушек стеков.

Термины:
Запуск гостя — guest launch — операция glaunch.
Приостановка работы гостя — interception (перехват).

Механизмы запуска и снятия гостя (2)

Переключение контекстов выполняется аппаратно-программно.

Контекст разделен на две части: «активную» и «пассивную».

«Активная» - это та, которая требуется гипервизору для немедленной нормальной работы, но без «изысков», например, без AAU. Сюда включены: регистры режимов, регистры управления стеками и сами стеки, регистры управления виртуальной памятью и тп.
«Активная» часть переключается аппаратно.
Для этого введены теневые регистры, содержащие целевой контекст.

«Пассивная» - всё остальное.
«Пассивная» часть переключается программно гипервизором.

Механизмы приостановки работы гостя (1)

Эти механизмы называются «перехват» и требуются для следующего:

- Обработка событий (прерываний), адресованных гипервизору
- Защита аппаратуры от нежелательных действий гостя
  - *например, программное обнуление машины*
- Обработка гипервизором обращений гостя к ресурсам, которые виртуализованы
  - программно
  - с аппаратной поддержкой, но еще не активированы
    - *например, гостевой адрес еще не отмэппирован в физпамять*
- Подмена для гостя значений некоторых регистров и областей памяти
  - *например, идентификационного регистра процессора*
- Обеспечение режима работы «time-share»
  - *исчерпание отведенного кванта времени*
- Освобождение аппаратуры, если гость не активен
  - *гость остановился на операции ожидания «сигнала»*
- Обработка сбоев аппаратуры

Механизмы приостановки работы гостя (2)

Гипервизору передается информация о причине и параметрах перехвата. Для этого введены структуры intc_info_cu/mu.

intc_info_cu отвечает за обращения в регистры подсистемы управления, прерывания гипервизору, исчерпание кванта времени, ...

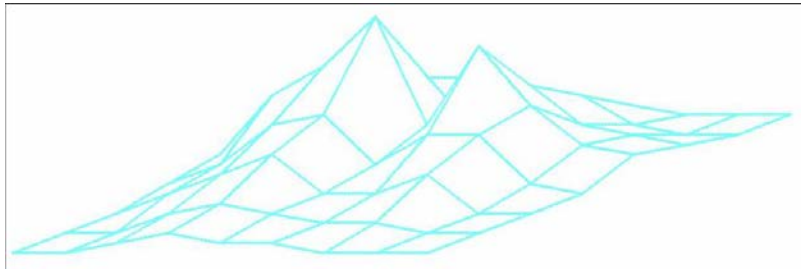
intc_info_mu отвечает за обращения в регистры подсистемы памяти, события при трансляции GPA->PA, …

После обработки перехвата эти же структуры содержат информацию для аппаратной доработки перехваченных операций.
При этом у гипервизора есть возможность подмены значений, которые, например, читаются гостем из регистра или спецобласти памяти.

Аппаратная поддержка виртуализации ресурсов

Ресурсы, виртуализация которых поддержана аппаратно:

- Вычислительные
- Оперативная память
- Периферия
- Межпроцессорные и внешние прерывания



Вопросы?

Нужно ли продолжать?