Операционные системы
Ч. II

Лекция 1
Гипервизоры и виртуализация ресурсов

Боровцов Е.Г.

Преимущества виртуализации информационной инфраструктуры

● **Консолидация серверов** За счет консолидации нагрузок на одной аппаратной платформе можно поддерживать схему «одно приложение —один сервер», в то же время предотвращая увеличение числа физических серверов.

● **Максимизация полезного времени** За счет разделения нагрузок исключается воздействие одного приложения на работу другого и предотвращаются сбои системы.

● **Надежное аварийное восстановление** Стратегия виртуализации позволяет поддерживать план мгновенного аварийного восстановления

● **Сокращение регрессивного тестирования приложений на совместимость** За счет виртуализации приложений и их использования на настольных ПК по запросу практически исключаются конфликты приложений между собой. Это значительно сокращает регрессивное тестирование перед развертыванием и устраняет большинство проблем, связанных с совместимостью

Преимущества виртуализации информационной инфраструктуры

- **Поддержка устаревших и специализированных приложений** виртуализация настольных систем позволят новым ОС поддерживать приложения, написанные для устаревших операционных платформ без исправления их программных кодов

- **Эффективное обслуживание серверов** Гибкость распределения нагрузок между физическими серверами с минимальным ущербом для их работы позволяет планировать техническое обслуживание серверов без прерывания обслуживания

- **Оптимизированный ввод в эксплуатацию** Ввод в эксплуатацию рабочих ресурсов можно ускорить и отделить от процесса приобретения оборудования

- **Снижение сложности** При управлении виртуальной инфраструктурой с помощью таких же средств, которые используются для физических активов, можно упростить сложность и оптимизировать изменения, которые вносятся во всю инфраструктуру.

Объекты виртуализации

- Виртуализация серверов
- Виртуализация хранилищ
- Виртуализация сети
- Виртуализация десктопов (технология BYOD)
- Виртуализация приложений

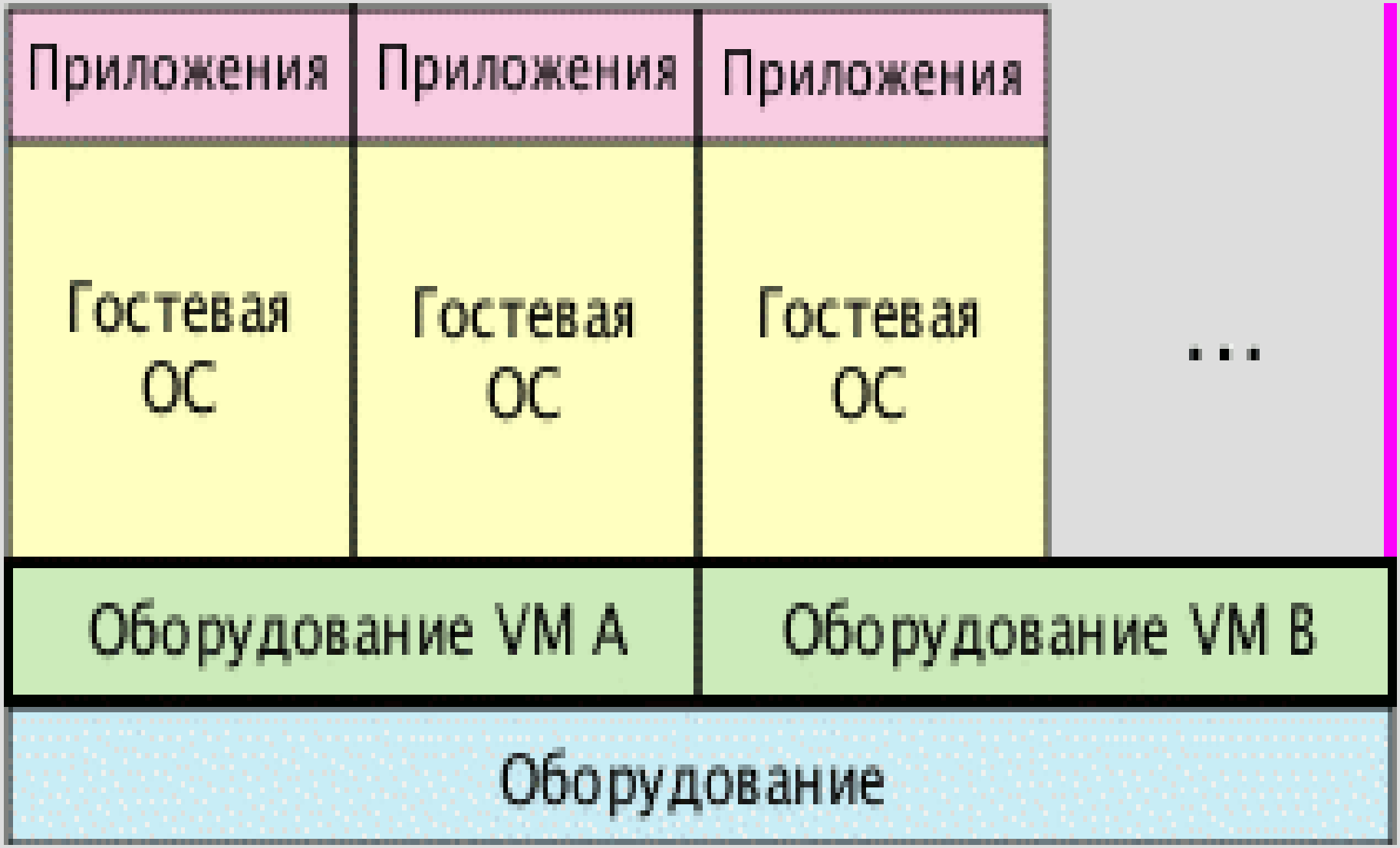
Проблемы виртуализации

- Производительность:
- исполнение инструкций
- доступ к памяти
- операции ввода-вывода
- время отклика системы

- безопасность, стабильность

- точное соответствие поведения

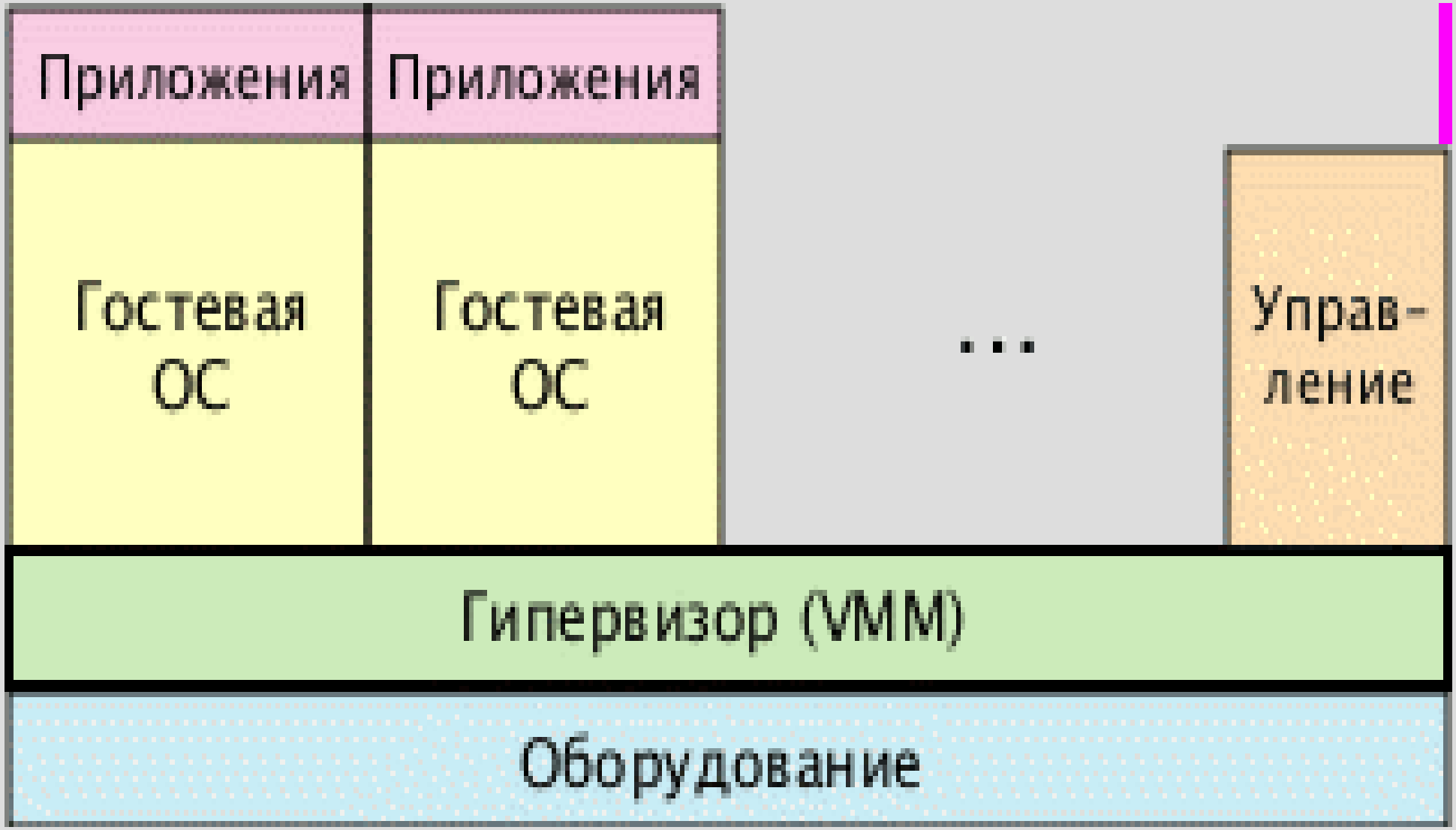
Типы виртуализации - эмуляция



Эмуляция - примеры

Эмуляция архитетуры 80x86 и системы DOS:
DosBox;
Bosch;
PCEmu
Эмуляция одной платформы на другой:
QEMU (x86, ARM,SPARK, MIPS,
PowerPC,m68k,x86-64) на (x86, x86-64,PowerPC, DEC Alpha, SPARC, ARM,S390)
Эмуляция оборудования и операционной среды:
Эмулятор Android;
Эмулятор Windows Phone;

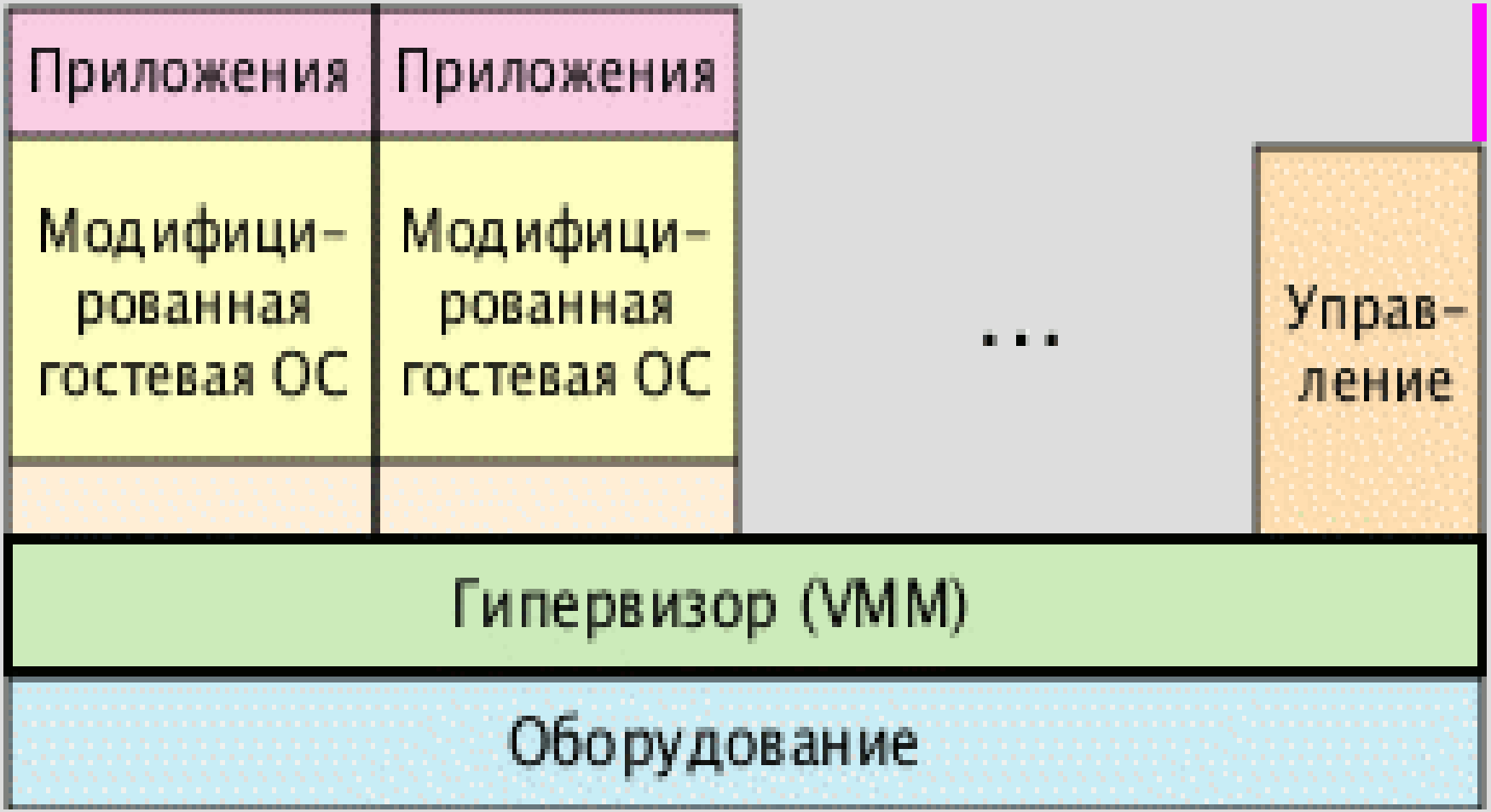
Типы виртуализации — полная виртуализация



полная виртуализация - примеры

VMware:
VMware ESXi, VMware Workstation, WMware Player, Vmware Fusion;
Microsoft:
HyperV, VirtualPC;
Open Virtualization Alliance:
KVM;
FreeBSD Team: BHyVe;
MAC OS X: xHyVe, Parallels

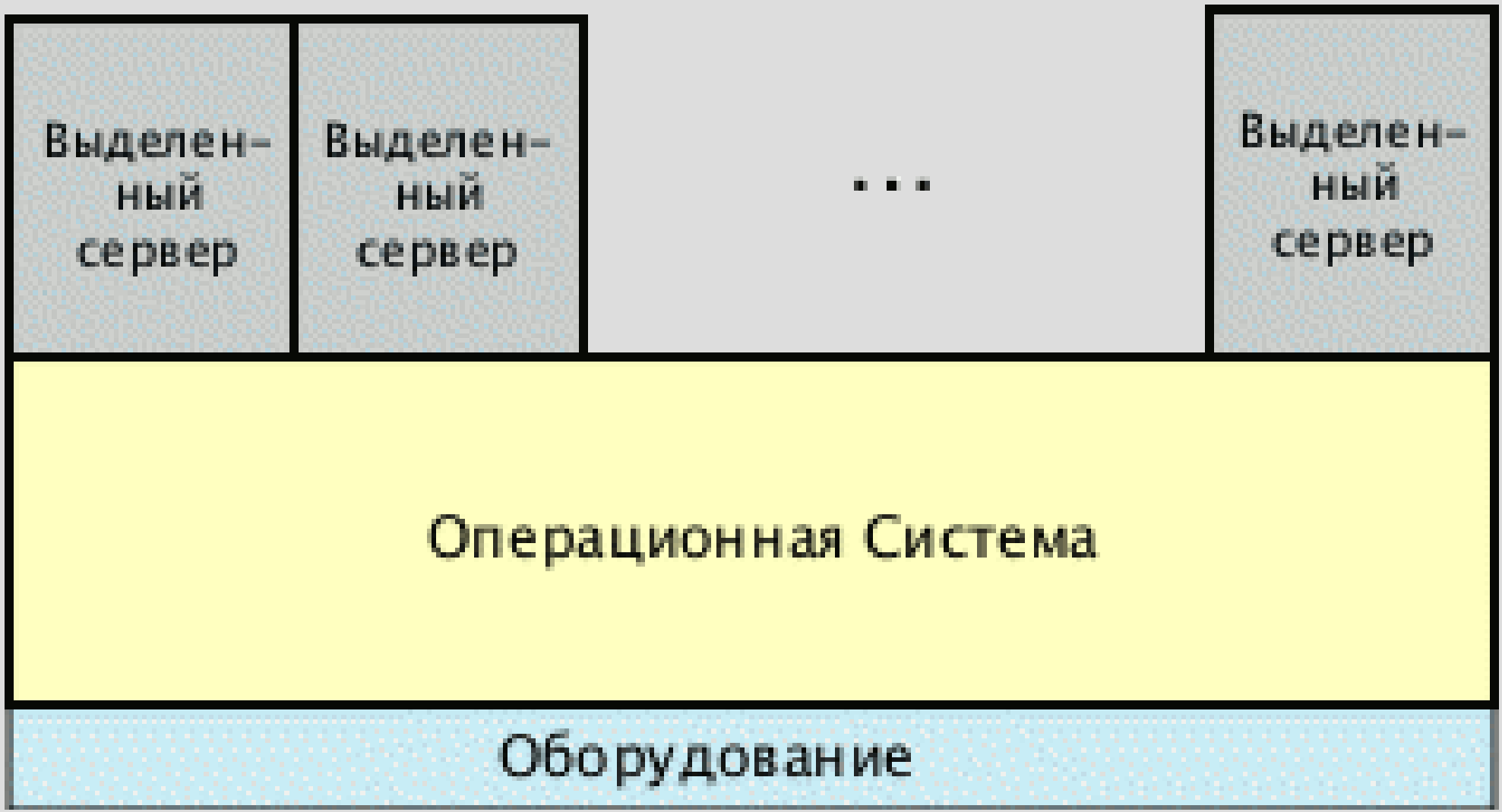
Типы виртуализации - паравиртуализация



Паравиртуализация - примеры

XEN — гипервизор общего назначения;
XtratuM — гипервизор для систем реального времени;
Trango — гипервизор реального времени для платформ ARM и MIPS. В 2008 г. приобретен Vmware. Переименован в **Vmware Horizon Mobile**.

Типы виртуализации — виртуализация уровня ОС



виртуализация уровня ОС - примеры

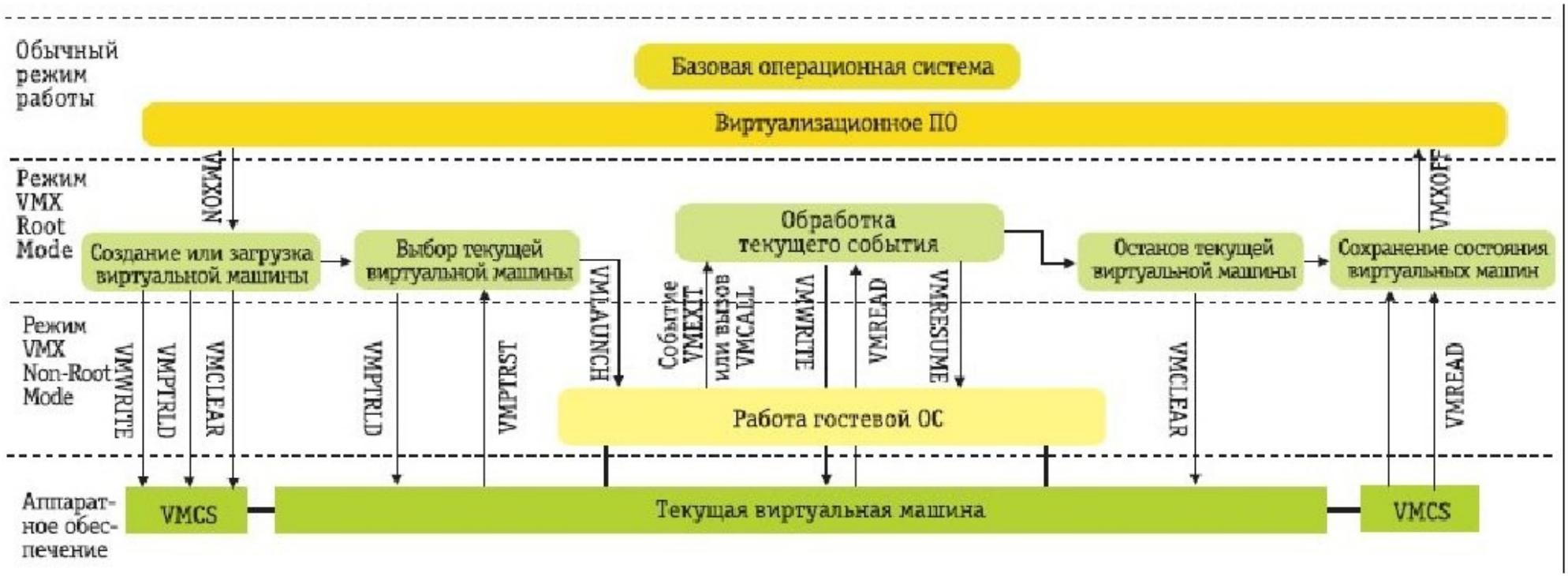
OpenVZ- открытый проект;
Virtuozzo — коммерческая надстройка над OpenVZ;
LXC — открытый проект;
FreeBSD Jails — открытый проект;
Solaris Zones — коммерческий проект;
Windows Server Containers — Windows Containers, Hyper-V Containers — проект от Микрософт (Windows Server 2016);

Аппаратная виртуализация — Intel VT-x

Набор команд Intel VT-x

VMXON и VMXOFF включают и выключают режим VMX.
VMWRITE позволяет записывать данные в структуру VMCS, описывающую виртуальную машину;
VMREAD - аналогично читать данные из VMCS.
VMPTRLD позволяет выбрать текущую виртуальную машину (указатель на VMCS).
VMPTRST, аналогично, сохранить указатель на текущую виртуальную машину.
VMLAUNCH позволяет запустить выбранную виртуальную машину (описывающуюся ранее установленным указателем на корректную текущую VMCS).
Исполнение кода работающей виртуальной машины прерывает либо наступление указанного в VMCS события (внешнего прерывания, попытки выполнить ту или иную инструкцию), либо выполнение инструкции **VMCALL** (если она разрешена в настройках VMCS).
VMRESUME позволяет продолжить прерванное событием выполнение кода на виртуальной машине.
VMCLEAR используется для инициализации пустой структуры VMCS и для перевода выбранной виртуальной машины в «остановленное» состояние (с сохранением данных VMCS).

Использование команд VT-x



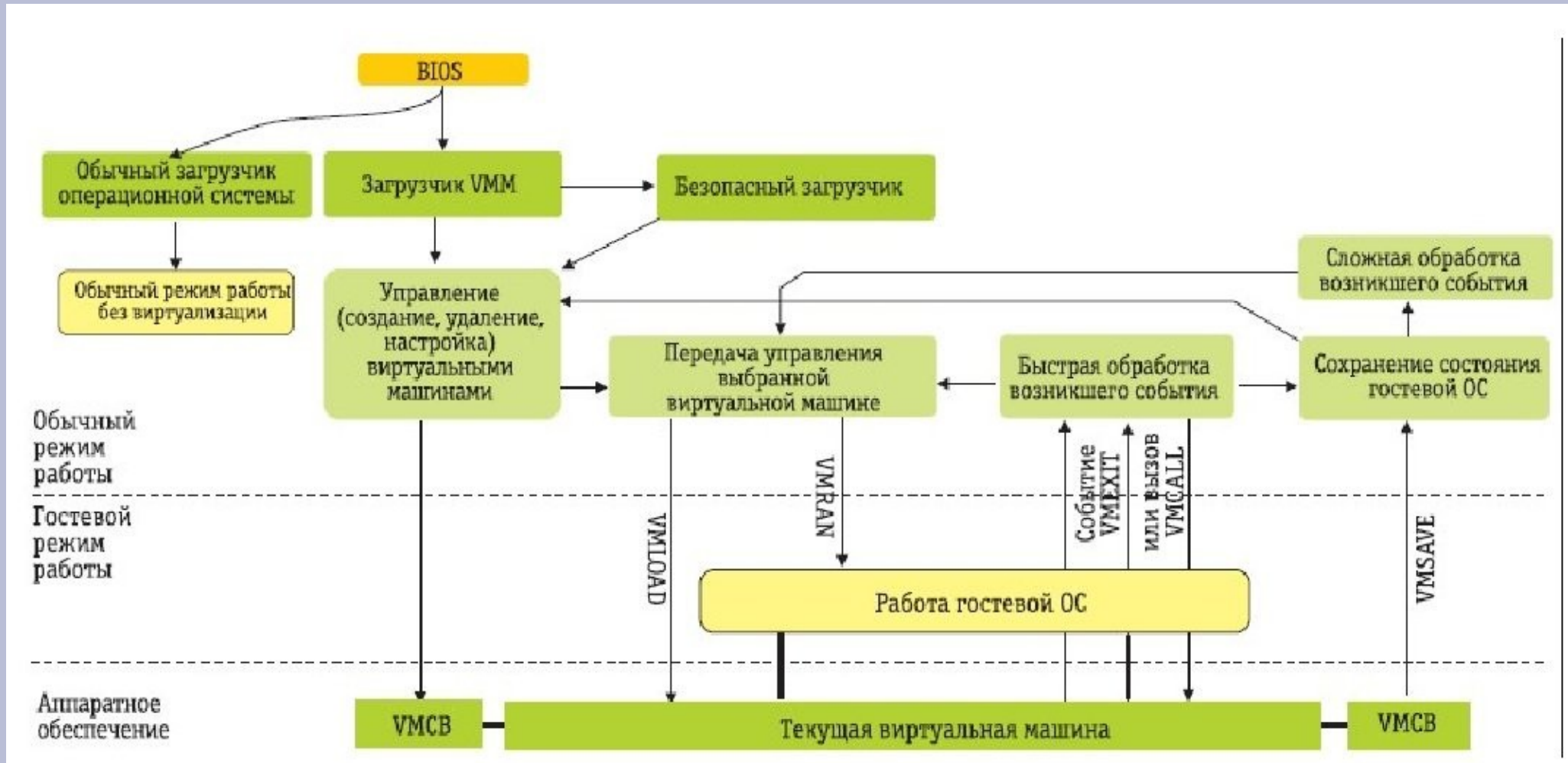
Виртуализация AMD-V

Набор команд AMD-V

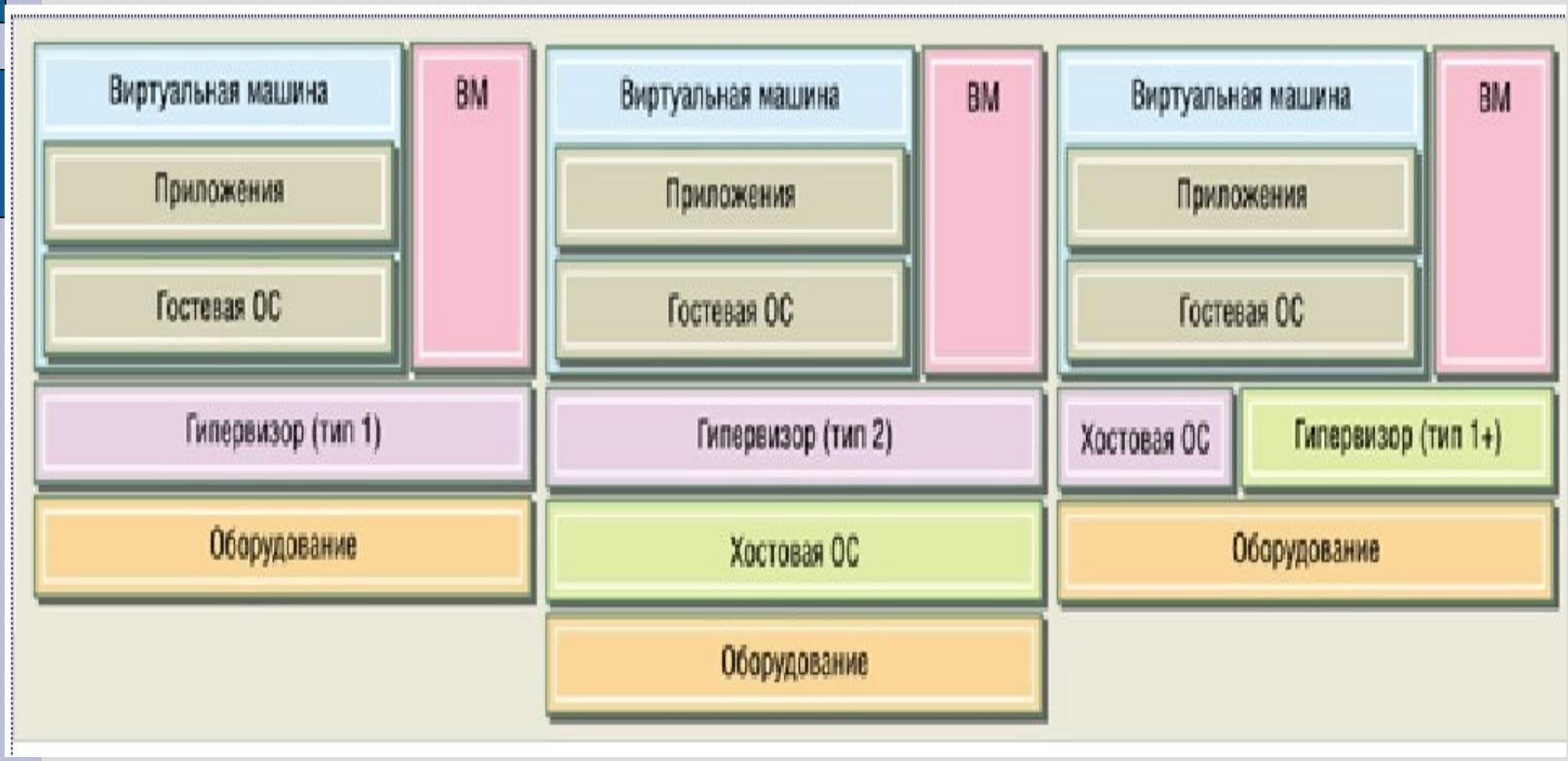
Команда **VMRUN** переключает выполнение на выбранную виртуальную машину.
Из виртуальной машины управление возвращается либо по перехвату одного из указанного в настройках виртуальной машины событий, либо при вызове специальной инструкции **VMMCALL** (если последняя разрешена настройками).
Информация о виртуальной машине хранится в VMCB (Virtual Machine Control Block) известного формата. VMM работает с данной структурой напрямую, изменяя при необходимости соответствующие поля
Самые необходимые операции по переключению контекста при переходах VMM к гостевой ОС и обратно выполняются автоматически. Однако, чтобы не совершать лишних действий, сохраняется и загружается действительно только самое необходимое, и при необходимости каких-либо сложных действий или переключения на другую гостевую ОС, «дополнительные» операции сохранения состояния процессора в VMCB и обратной загрузки выполняются инструкциями **VMLOAD** и **VMSAVE**.
Инструкция **SKINIT** позволяет начать загрузку процессора в «безопасном режиме», на аппаратном уровне гарантировав соответствие загрузчика (до 64 Кбайт кода) указанной в аппаратуре (в модуле TPM) цифровой подписи

Дополнительные инструкции для повышения производительности:
STGI, CLGI - управляют схемой перехвата прерываний в Pacifica (включают-выключают «глобальный перехват прерываний»).
INVLPGA - сбрасывает TLB, но не целиком, а только те записи TLB, которые относятся к конкретной гостевой ОС (или к VMM)

Использование команд AMD-V



Типы гипервизоров



Организация обмена с периферией