

Методы виртуализации подсистемы ввода-вывода

Рыбаков С.А.

12.03.2020

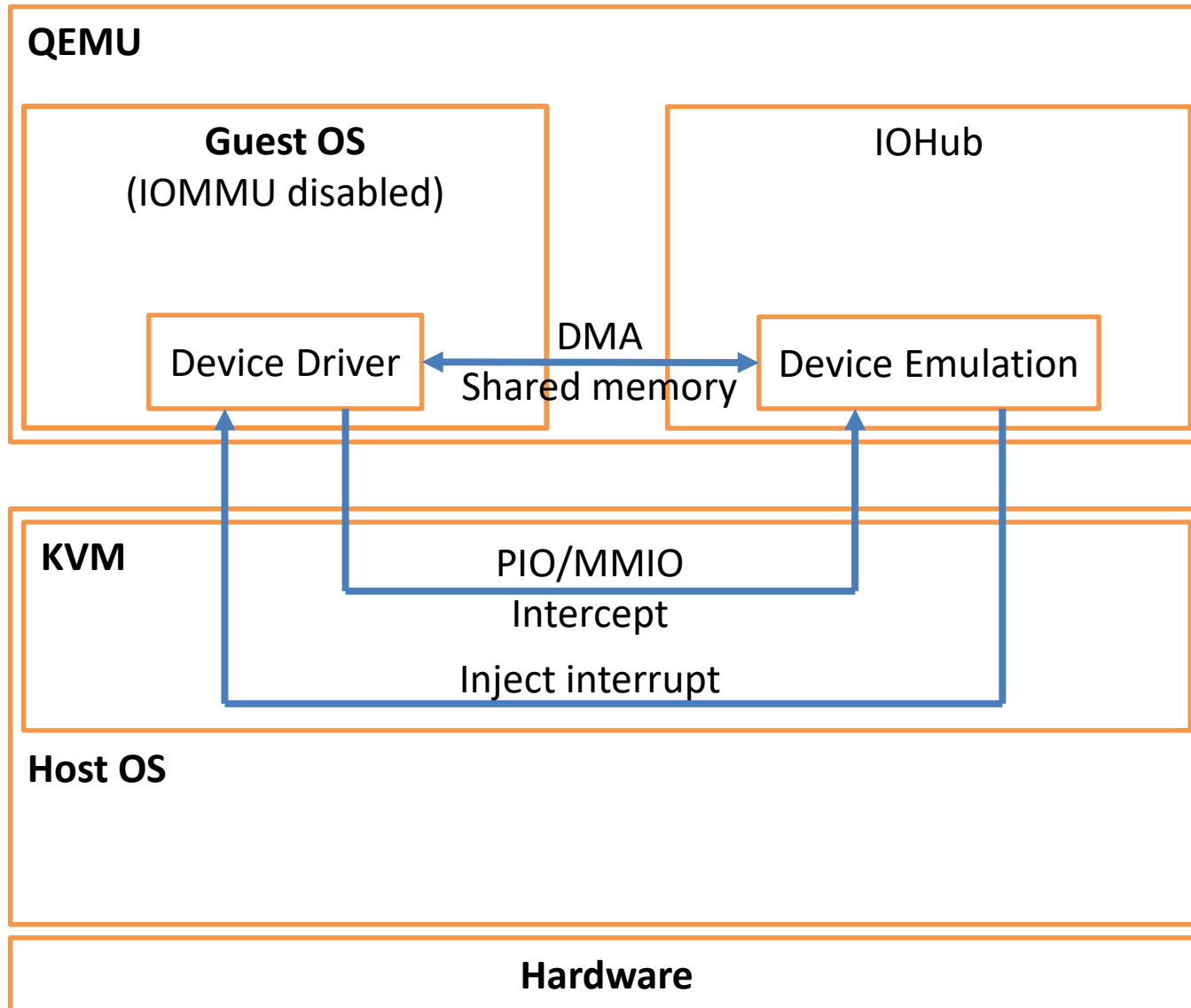
ПАРАВИРТУАЛИЗАЦИЯ И ПОЛНАЯ ВИТУАЛИЗАЦИЯ

	Паравиртуализация	Полная виртуализация
Требуется модификация гостевой ОС	✓	✗
Гость знает, что он запущен внутри виртуальной машины	✓	✗
Необходима аппаратная поддержка	✗	✓

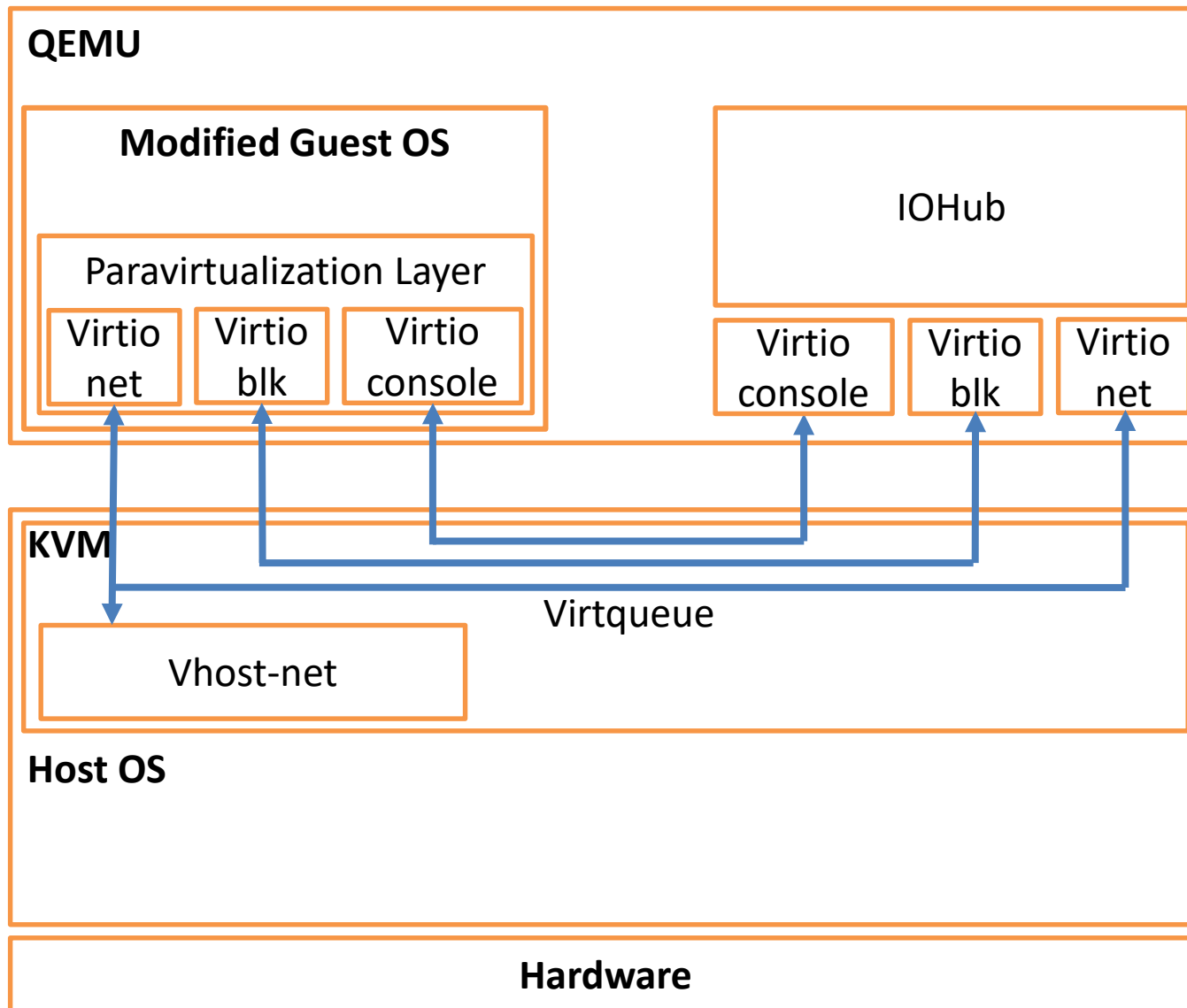
ИНТЕРПОЗИЦИЯ ВВОДА-ВЫВОДА

- Контроль гипервизора над состоянием VM, включая устройства ввода-вывода
 - Возможность приостановить работу VM, и сохранить ее состояние
 - Live Migration – перенос VM на другой сервер без приостановки
 - Замена оборудования на сервере во время работы VM
- Консолидация ввода-вывода, увеличение утилизации устройств
- Возможность объединения нескольких физических устройств в одно виртуальное
- Возможность поддержки в виртуальном устройстве функциональности, не поддерживаемой реальным оборудованием
- Стандартные оптимизации памяти применимы к гостевому физическому адресному пространству
 - Memory overcommitment
 - Demand-based paging
 - Lazy allocation
 - Swapping
 - Copy-on-write
 - Page migration
 - Transparent huge pages

Эмуляция (ПОЛНАЯ ВИРТУАЛИЗАЦИЯ ВВОДА-ВЫВОДА)



VIRTIO (ПАРАВИРТУАЛИЗАЦИЯ ВВОДА-ВЫВОДА)



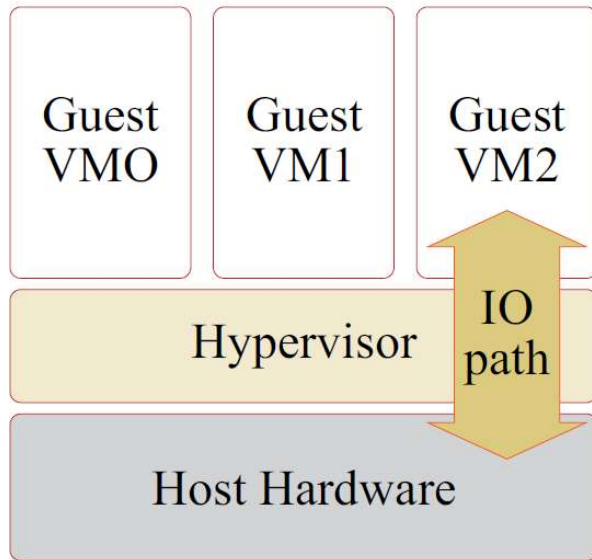
СРАВНЕНИЕ ПРОИЗВОДИТЕЛЬНОСТИ ЭМУЛЯЦИИ И VIRTIO

Netperf TCP stream

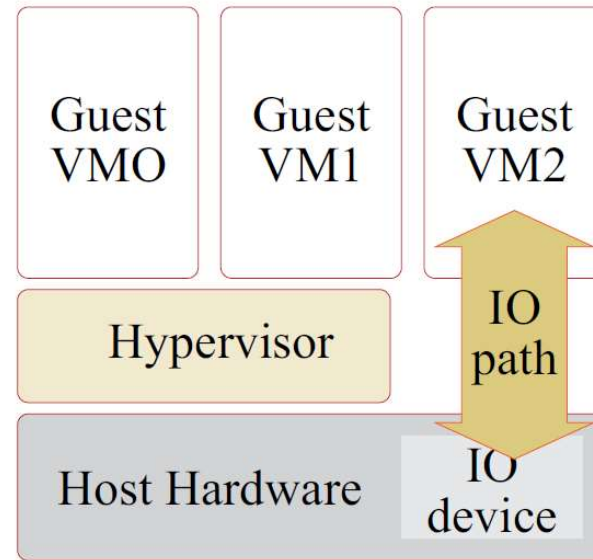
Linux 3.13, QEMU 2.2, Intel Xeon E5620

Метрика	e1000	Virtio-net	Соотношение
Пропускная способность (Мб/с)	239	5230	22x
Кол-во перехватов/с	33783	1126	1/30x
Кол-во прерываний/с	3667	257	1/14x

ПРОБРОС УСТРОЙСТВА ГОСТЮ



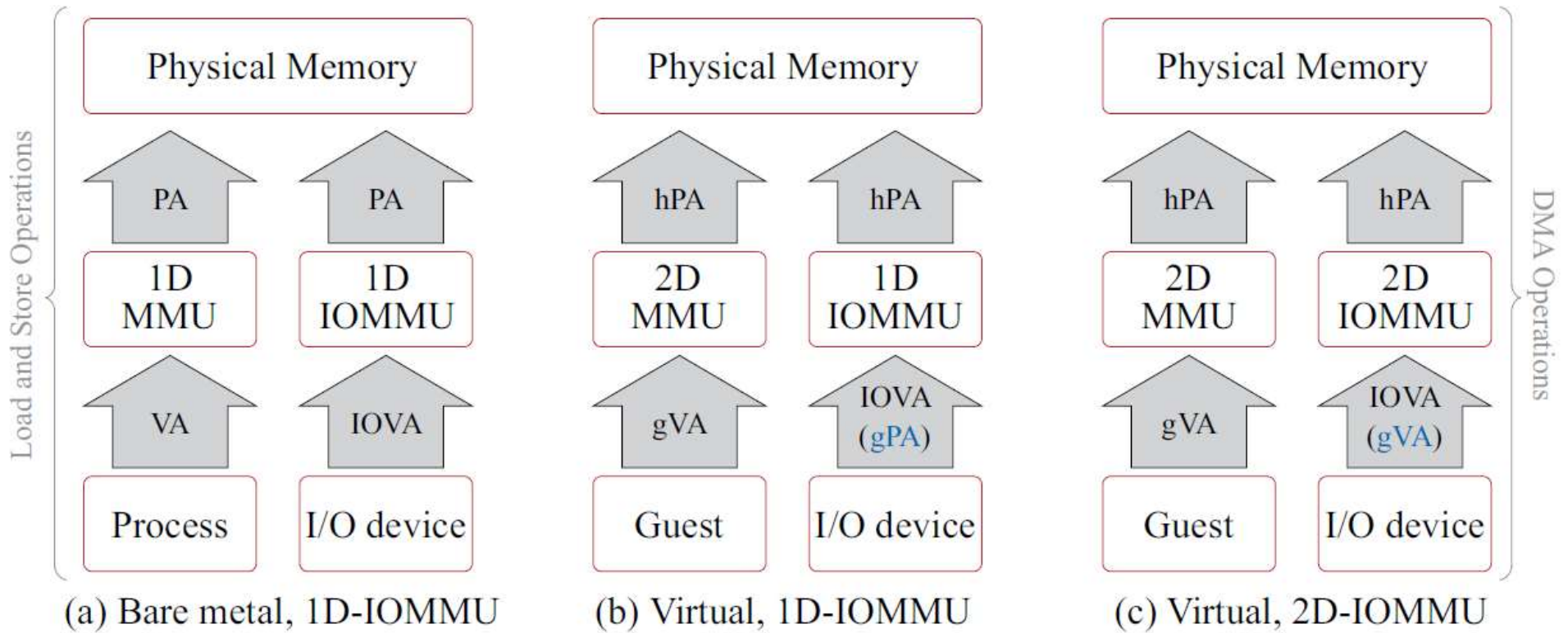
(a) Emulation/paravirtualization



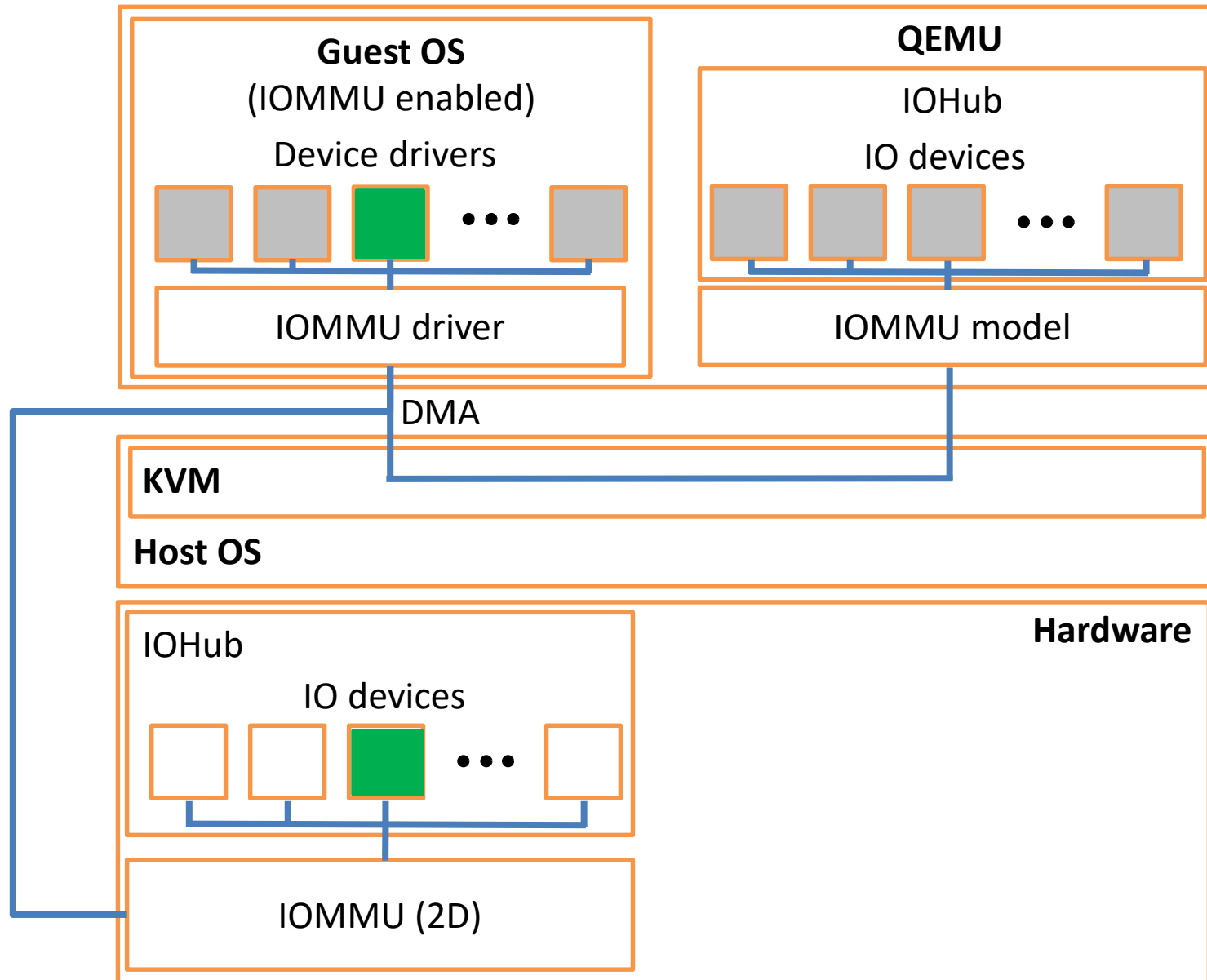
(b) Direct device assignment

Преимущества	Недостатки
- Нативная производительность (без издержек на эмуляцию)	- Отказ от интерпозиции - Плохая масштабируемость - Необходима защита от DMA и прерываний от проброшенного устройства

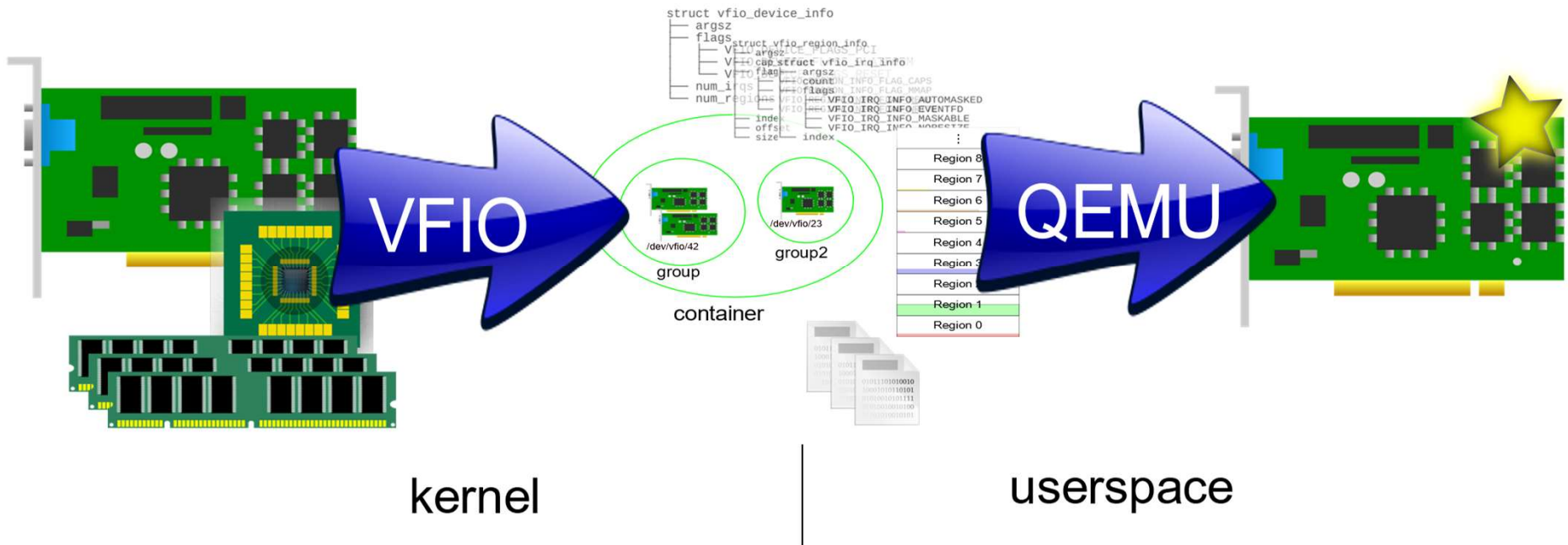
Трансляция DMA при пробросе



ПРОГРАММНАЯ МОДЕЛЬ IOMMU



VFIO – VIRTUAL FUNCTION I/O



СРАВНЕНИЕ МЕТОДОВ ВИРТУАЛИЗАЦИИ ВВОДА-ВЫВОДА

	Эмуляция	Паравиртуализация (Virtio)	Проброс
Переносимость, запуск гостевой ОС без модификации	✓	✗	✓
Интерпозиция и оптимизации памяти	✓	✓	✗
Производительность	Низкая	Высокая	Нативная