

Конструирование ядра операционных систем (IX-X)

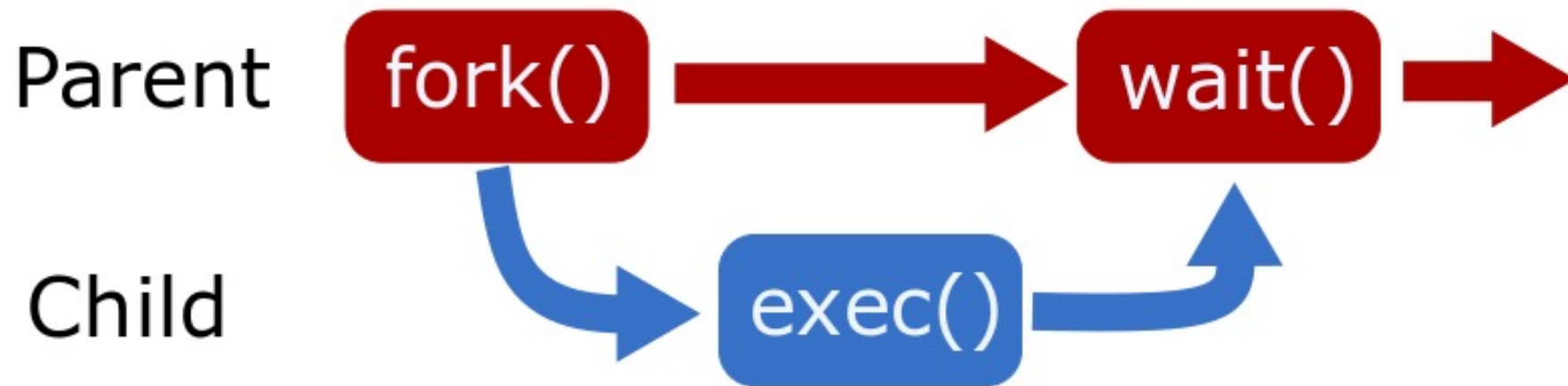
Многозадачность и файловая система

План

- Порождение новых процессов через `fork/exec`
- Оптимизация по скорости и памяти в создании процессов.
- Механизмы взаимодействия между процессами
- Файловая система

Создание процессов (fork/exec)

- Процесс клонируется (fork) и замещается новым (exec).
- В ребёнке fork() возвращает 0, в родителе — pid процесса.
- Функция wait() или waitpid() ожидает завершения ребёнка.



Ускорение fork и уменьшение RAM

- Механизм Copy-on-Write (COW)
 - Страницы помечаются как общие и доступны как ребёнку, так и родителю в режиме для чтения. В момент записи происходит дублирование страницы с передачей владения.
- Механизм vfork
 - При создании дочернего процесса адресное пространство не копируется и не может быть изменено ребёнком (UB). Ребёнок должен выполнить exes.
- Механизм posix_spawn
 - Создаётся новый процесс из указанного источника (файла) без клонирования родительского процесса (вызова fork).

Другие способы оптимизации ОЗУ

- Swapping
 - Реже используемые страницы памяти помещаются на диск и замещаются более часто используемыми страницами.
- Удаление копирования
 - Использование механизмов подсчёта ссылок для общих объектов.
- Сжатие страниц
 - Реже используемые страницы памяти сжимаются (e.g. LZSS) и занимают меньше пространства в ОЗУ.
- Кэш библиотек
 - Системные библиотеки отображаются один раз для всех процессов в режиме COW (e.g. dyld_shared_cache).

Inter-process communication (IPC)

- Для нормальной работы процессам необходимо передавать данные друг другу через некоторый канал.
- IPC должны быть быстрыми (минимальное количество копирований и переключений контекста) и использовать понятное количество ресурсов.

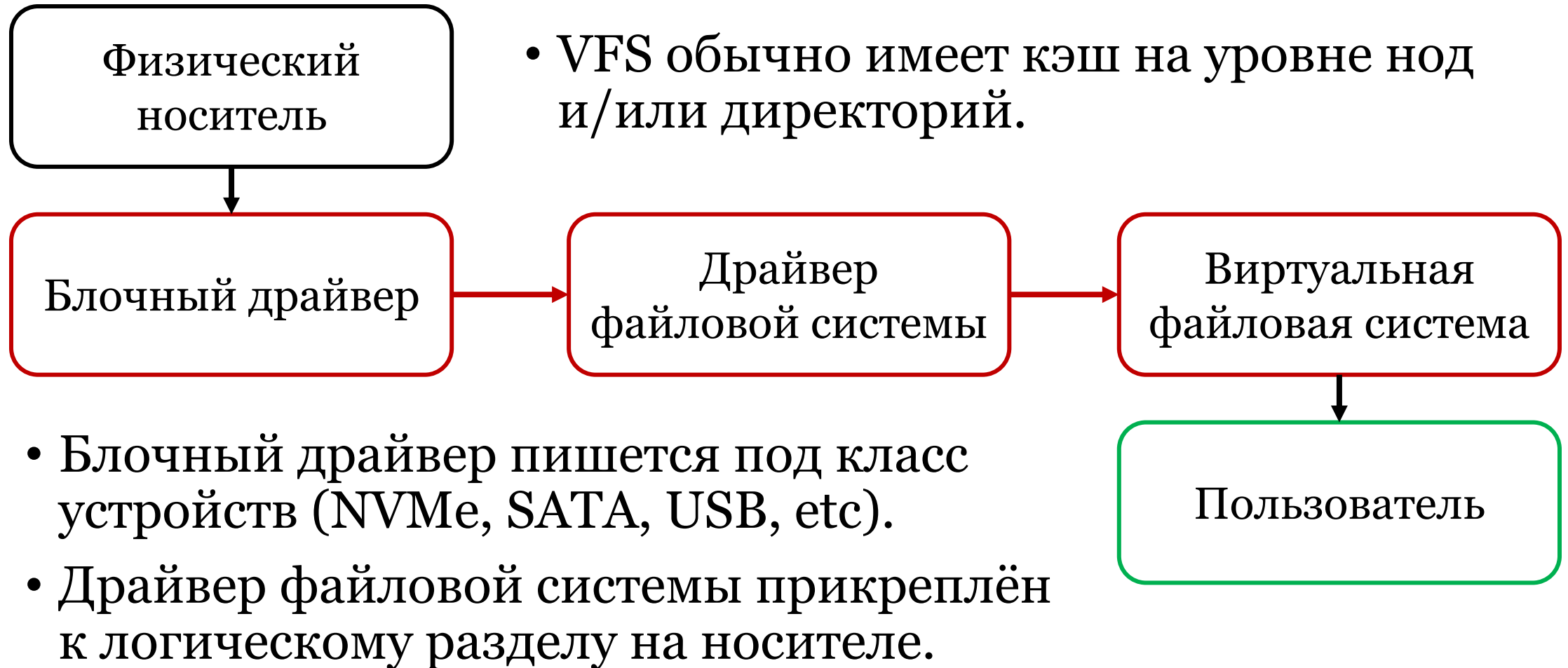
IPC в JOS

- Реализован парой функций: `ipc_send` и `ipc_recv`.
 - `ipc_recv` усыпляет текущий процесс в ожидании послылки.
 - `ipc_send` отправляет послылку целевому процессу.
 - Передача данных идёт через прямое отображения страницы отправителя в адресное пространство получателя. За счёт этого максимизируется производительность.
-
- Нет таймаута ожидания послылки.
 - Размер ограничен одной страницей.

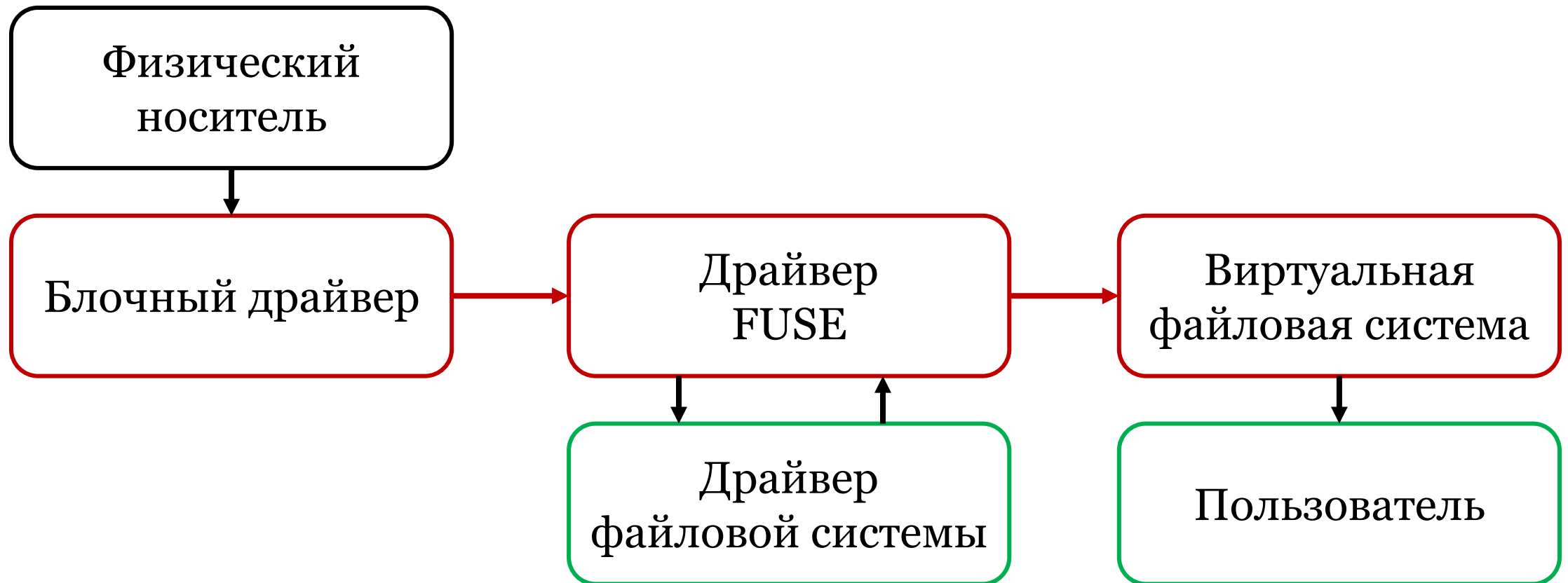
Другие виды IPC

- IPC через файловую систему
 - Файлы
 - Именованные/неименованные каналы
 - Файловые сокеты
- IPC через оперативную память
 - Очереди сообщений (mq)
 - Сигналы
 - Общая память
- IPC через устройства
 - Сетевые сокеты

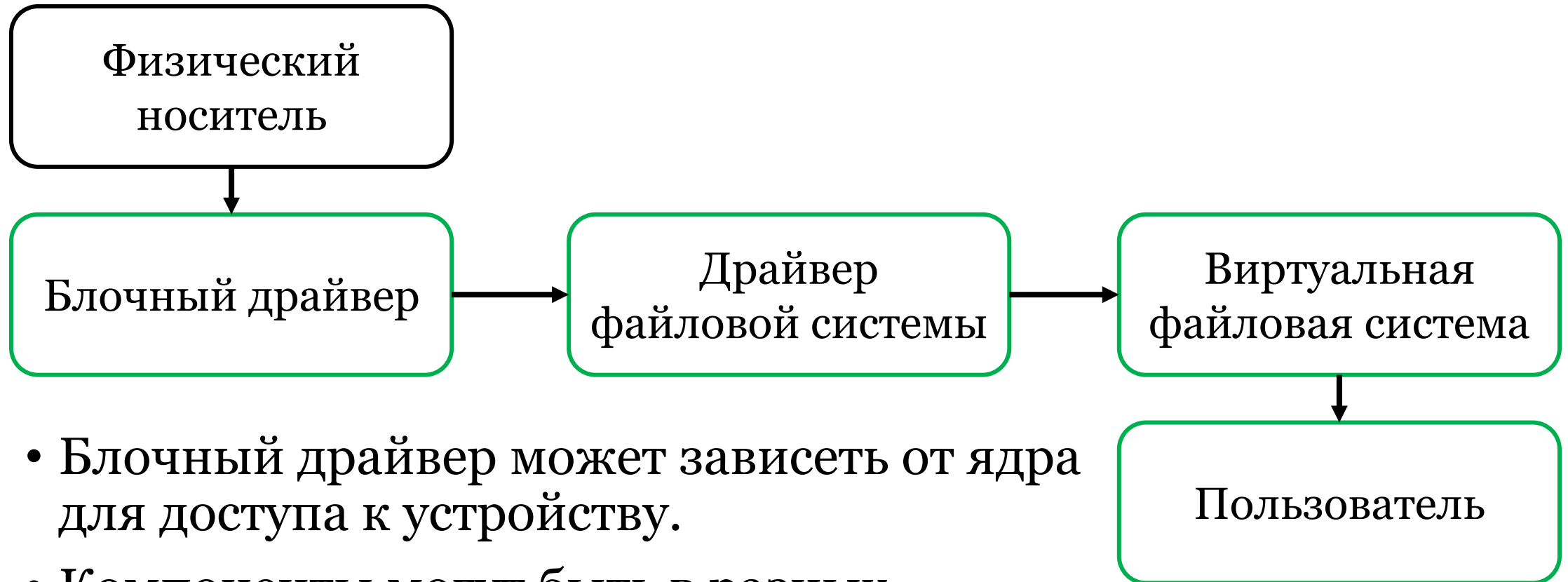
Классическая схема ФС



ФС в userspace (e.g. FUSE)



ФС в микроядерной архитектуре



- Блочный драйвер может зависеть от ядра для доступа к устройству.
- Компоненты могут быть в разных процессах для обеспечения изоляции.

Ограничения со стороны VFS

- Порядок получения доступа к файлам
- Порядок сериализации транзакций к файлам
- Количество одновременно открытых файлов
- Количество одновременных операций
- Размер одной транзакции

Спасибо за внимание!

Вопросы?