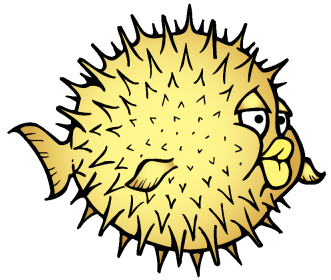
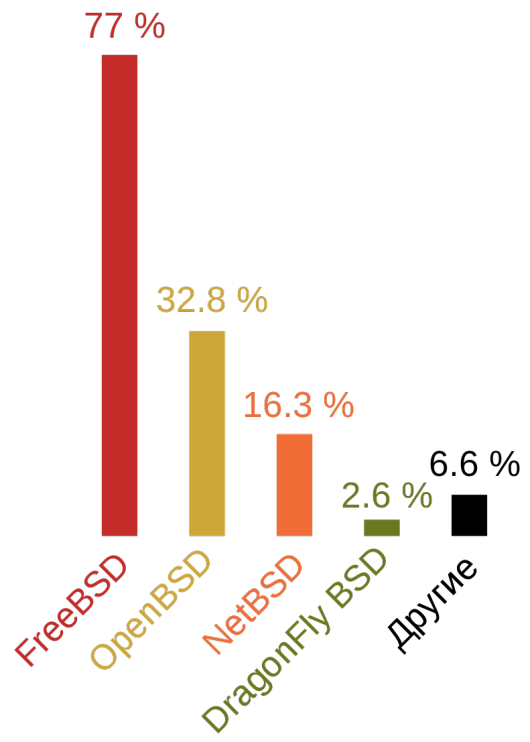
OpenBSD



Историческая справка



Популярность BSD-систем
(2005 год, BSD Certification Group Report)

- OpenBSD — самостоятельный проект, ответвление NetBSD, возникшее в конце 1995 года в результате раскола в команде разработчиков. Тео де Раадт, один из четырёх основателей NetBSD, был вынужден покинуть проект после конфронтации по поводу дальнейшего развития ОС.
- Взяв за основу дерево исходных кодов NetBSD и переделав его в соответствии со своим видением, он создал свой собственный проект — OpenBSD, в который, вслед за ним, перешли и некоторые другие разработчики NetBSD.

Тео де Раадт (Theo de Raadt)

- Тео де Раадт на первой российской технической конференции по ОС BSD – ruBSD, 2013 год

Цели проекта

- Переносимость (поддерживается 12 аппаратных платформ)
- Стандартизация, корректная работа
- Активная безопасность
- Интегрированные криптографические средства

Сферы применения

- Маршрутизаторы и точки доступа
- Персональные компьютеры: более 8000 пакетов в репозитории:
  - Xenocara – релизация X Window System
  - Рабочие окружения: GNOME, KDE, XFCE
  - Прикладное ПО: веб-браузеры, офисные пакеты и т. п.
- Всевозможные серверы: почтовые, веб, FTP, DNS, NFS

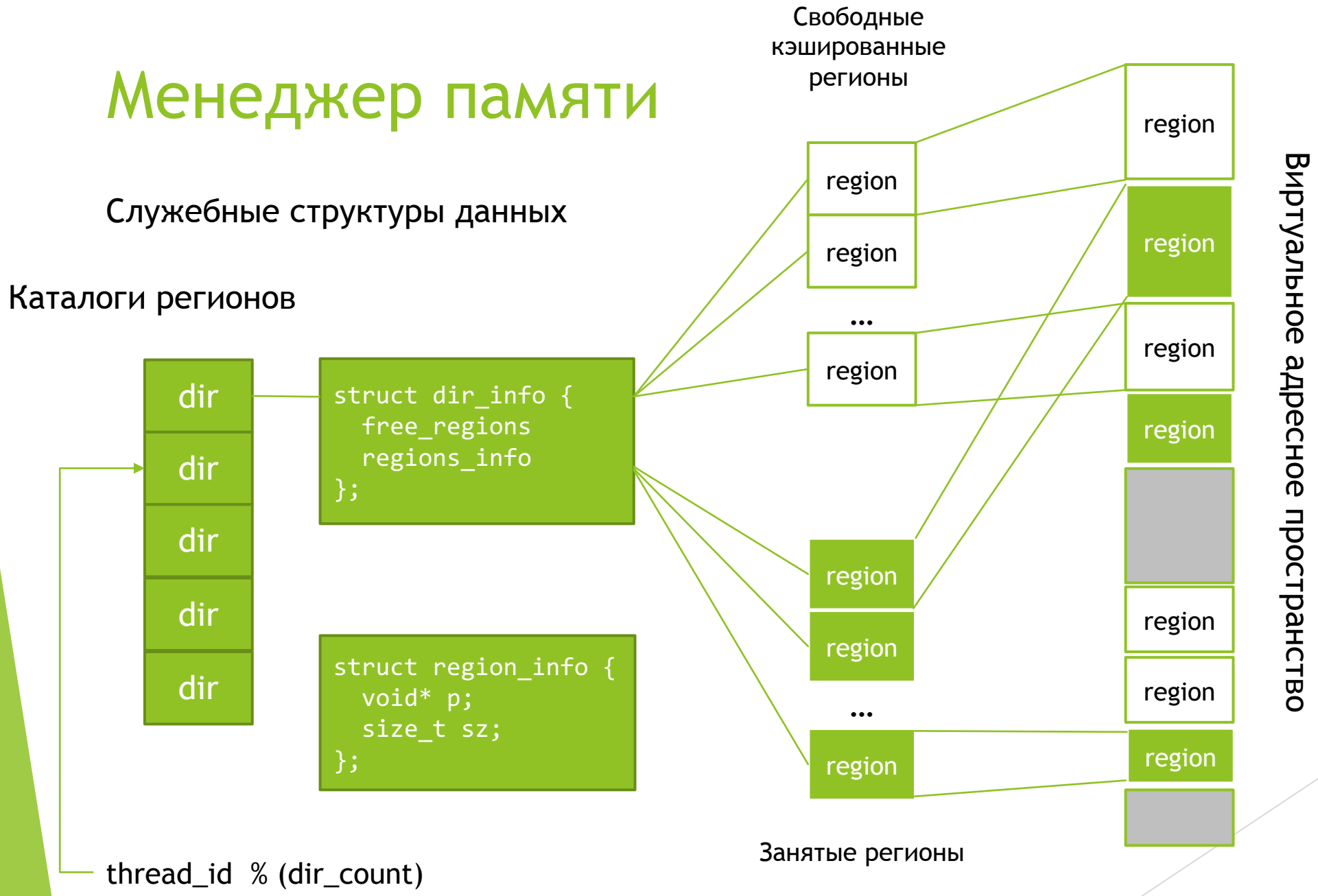
Другие проекты OpenBSD

- LibreSSL - библиотека с реализацией протоколов SSL/TLS
- OpenSSH - набор программ, предоставляющих шифрование сеансов связи по компьютерным сетям с использованием протокола SSH
- Пакетный фильтр PF (используется в macOS, NetBSD, FreeBSD)
- Демоны маршрутизации OpenBGPD и OpenOSPFD
- Утилита синхронизации файлов OpenRSYNC
- Демон синхронизации локального системного времени OpenNTPD

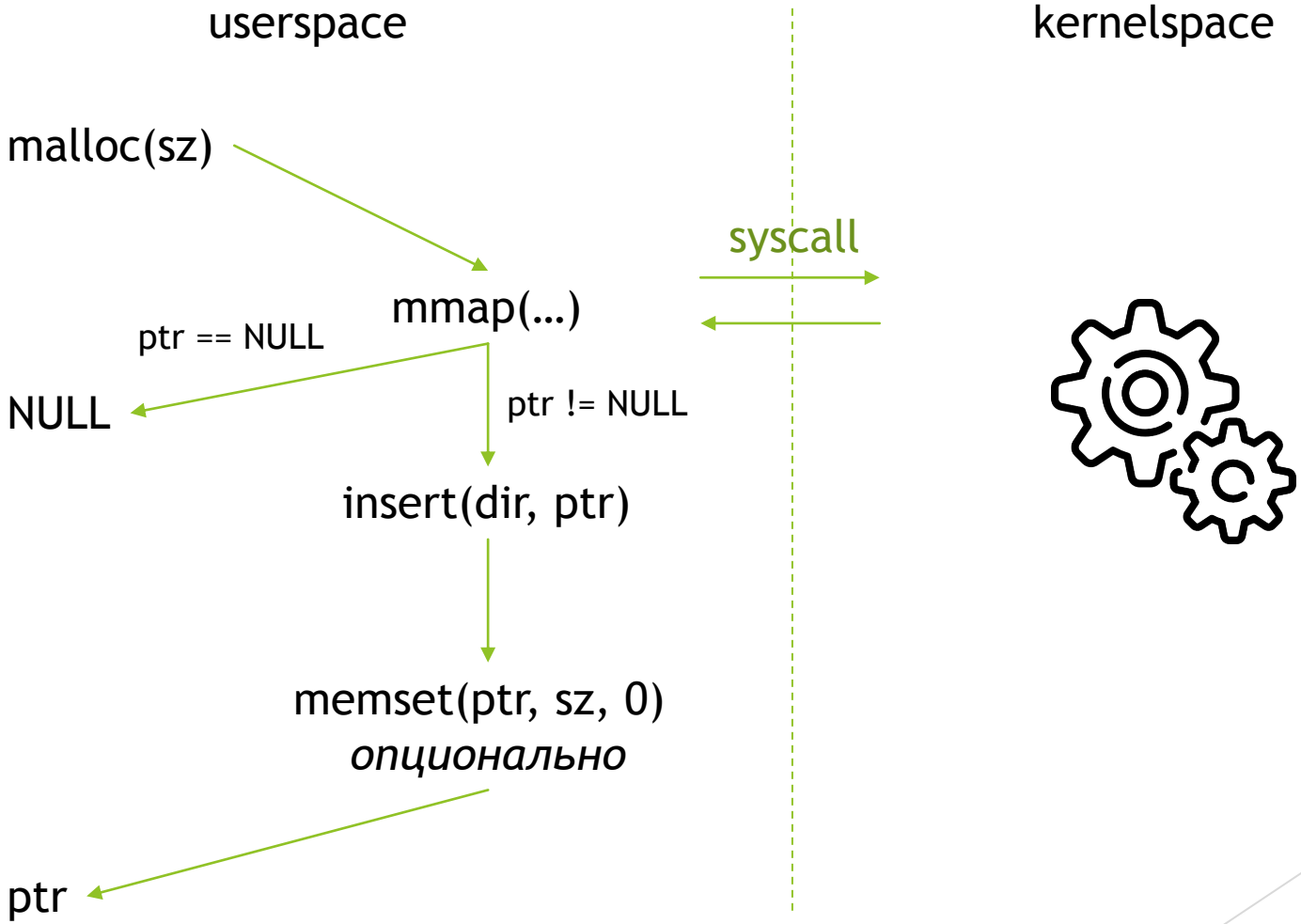
Менеджер памяти

Служебные структуры данных

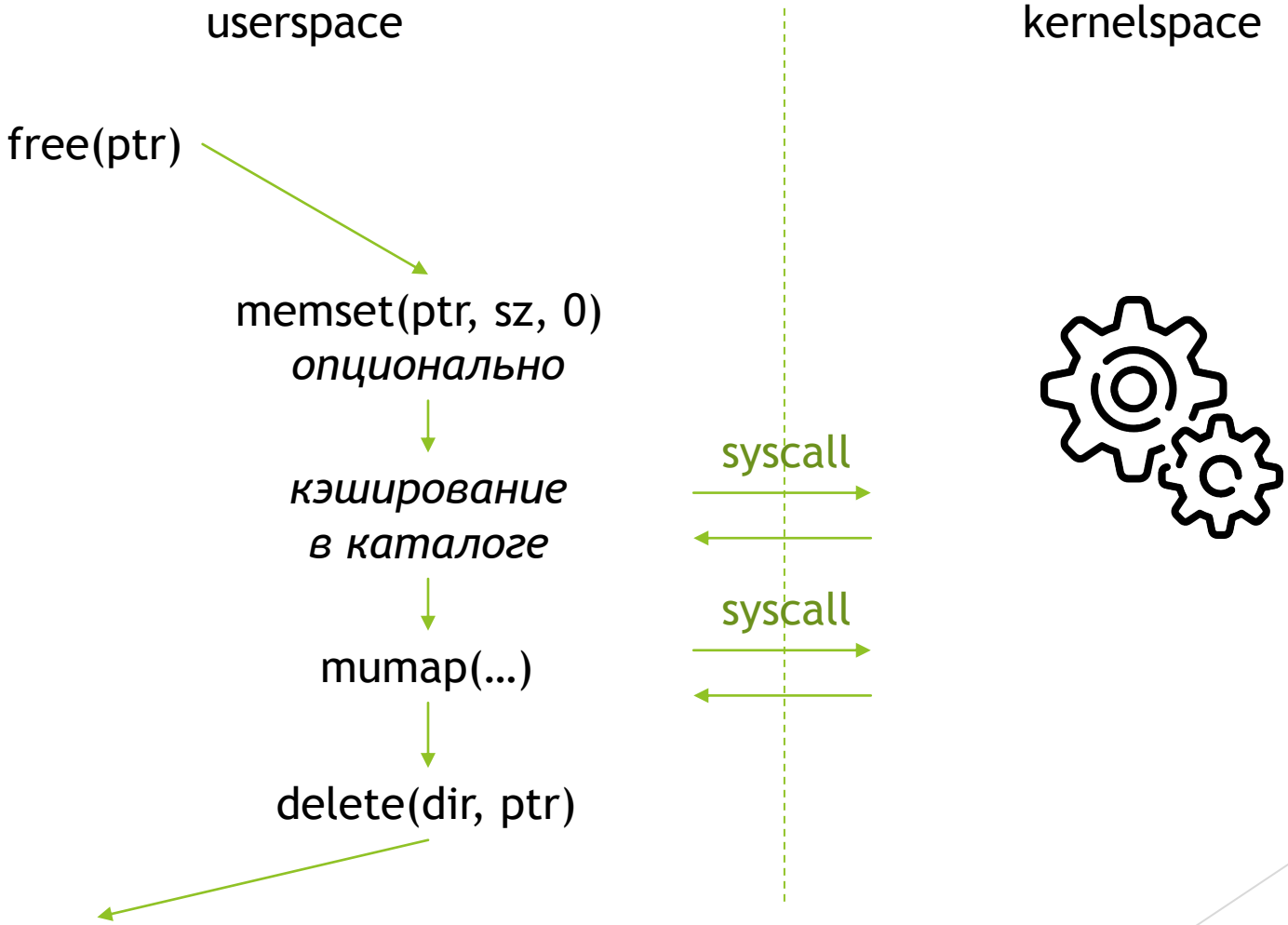
Каталоги регионов



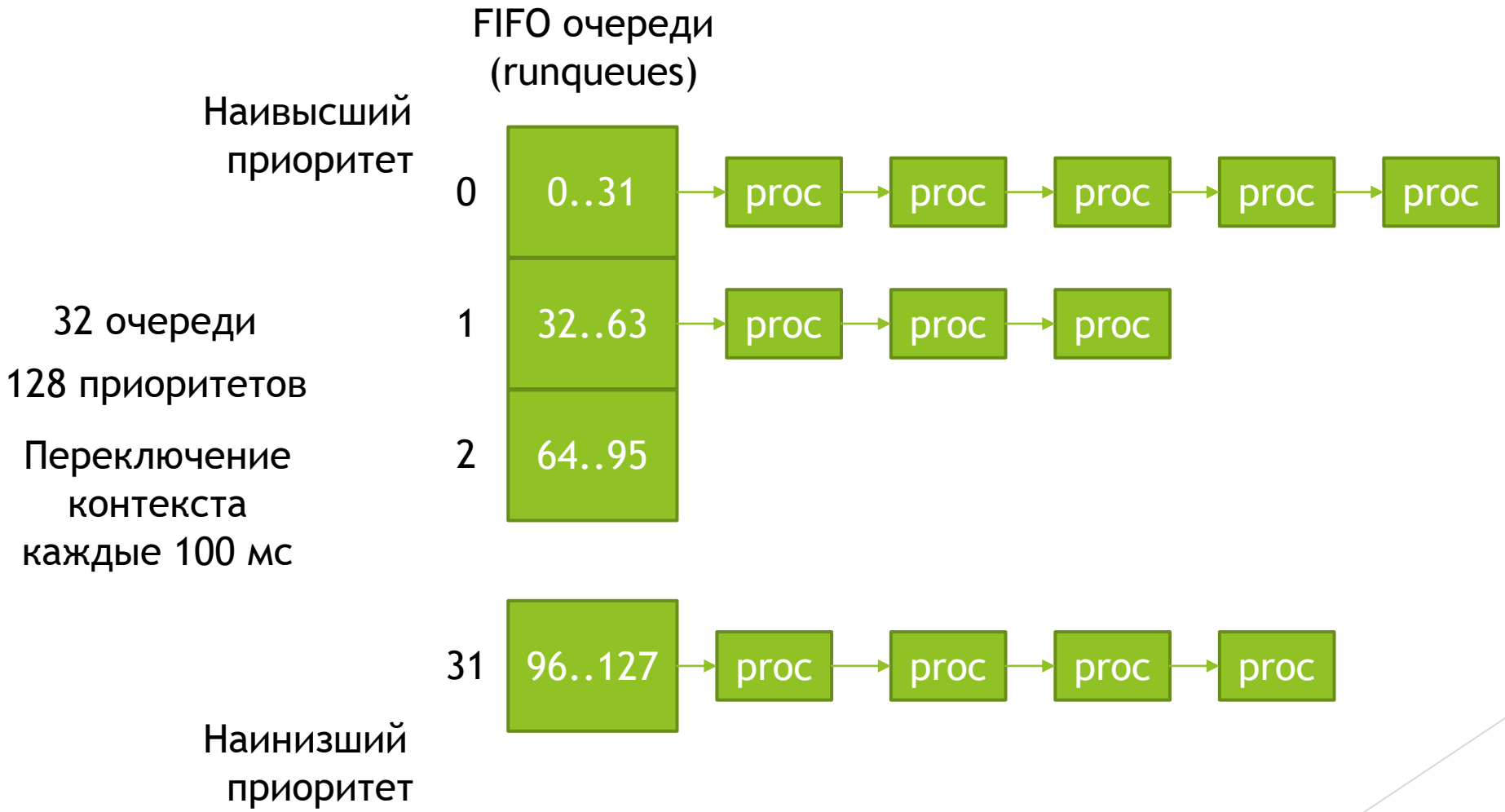
Выделение памяти



Освобождение памяти



Планировщик процессов



Параметры потока

- `priority` - приоритет потока
- `nice` - параметр, определяющий, насколько данный поток лучше других с таким же приоритетом
- `estcpu` - параметр, определяющий уровень использования потоком CPU
- `slptime` - счетчик, определяющий, сколько секунд поток провел в состоянии ожидания

Основная идея

- Чем больше поток потребляет ресурсов CPU, тем ниже становится его приоритет
- Чем дольше поток пребывает в состоянии ожидания, тем меньше значение `estcpu` и тем выше становится его приоритет
- Через определенный промежуток времени система «забывает» 90% информации о потреблении ресурсов

Пересчет estcpu и priority

- Для потоков в очереди и для исполняющегося:

$$estcpu_n = \frac{2 \times load_{avg}}{2 \times load_{avg} + 1} estcpu_{n-1}$$

$$priority = 50 + estcpu + 2 \times nice$$

Пересчет каждую секунду

- Для исполняющегося потока:

$$estcpu = \min\{estcpu + 1;\ 2 \times 20\ - 4\}$$

$$priority = 50 + estcpu + 2 \times nice$$

Пересчет с частотой ~12-16 Гц

Пересчет `estcpu`

При $load_{avg} = 1$:

- 1 сек: $e_1 = 0.66(e_0 + t_0), e_0 = 0$
- 2 сек: $e_2 = 0.66(e_1 + t_1) = 0.66t_1 + 0.44t_0$
- 3 сек: $e_3 = 0.66(e_2 + t_2) = 0.66t_2 + 0.44t_1 + 0.30t_0$
- 4 сек: $e_4 = 0.66(e_3 + t_3) = 0.66t_3 + \cdots + 0.20t_0$
- 5 сек: $e_5 = 0.66(e_4 + t_4) = 0.66t_4 + \cdots + 0.13t_0$

~10% от исходного значения

Выход потока из состояния ожидания

1 – Пересчет `estcpu`:

$$estcpu_1 = \left(\frac{2 \times load_{avg}}{2 \times load_{avg} + 1}\right)^{slptime} \times estcpu_0$$

2 - Пересчет `priority`:

$$priority = 50 + estcpu_1 + 2 \times nice$$

3 – Добавление в очередь

4 - rescheduling

Специальные средства защиты ОС

- W^X – Write XOR Execute
- Программное отключение SMT (HyperThreading)
- Системный вызов unveil() для изоляции ФС
- Механизм защиты RETGUARD
- FileFlags
- Securelevels

Write XOR Execute

- Суть метода: страницы памяти загруженной программы либо доступны для чтения, либо для исполнения
- Защита от типовых атак через переполнение буферов
- Включен по умолчанию
- Для приложений, использующих «грязный» метод JIT-компиляции, доступна опция отключения метода (например, для веб-браузеров)

Программное отключение SMT

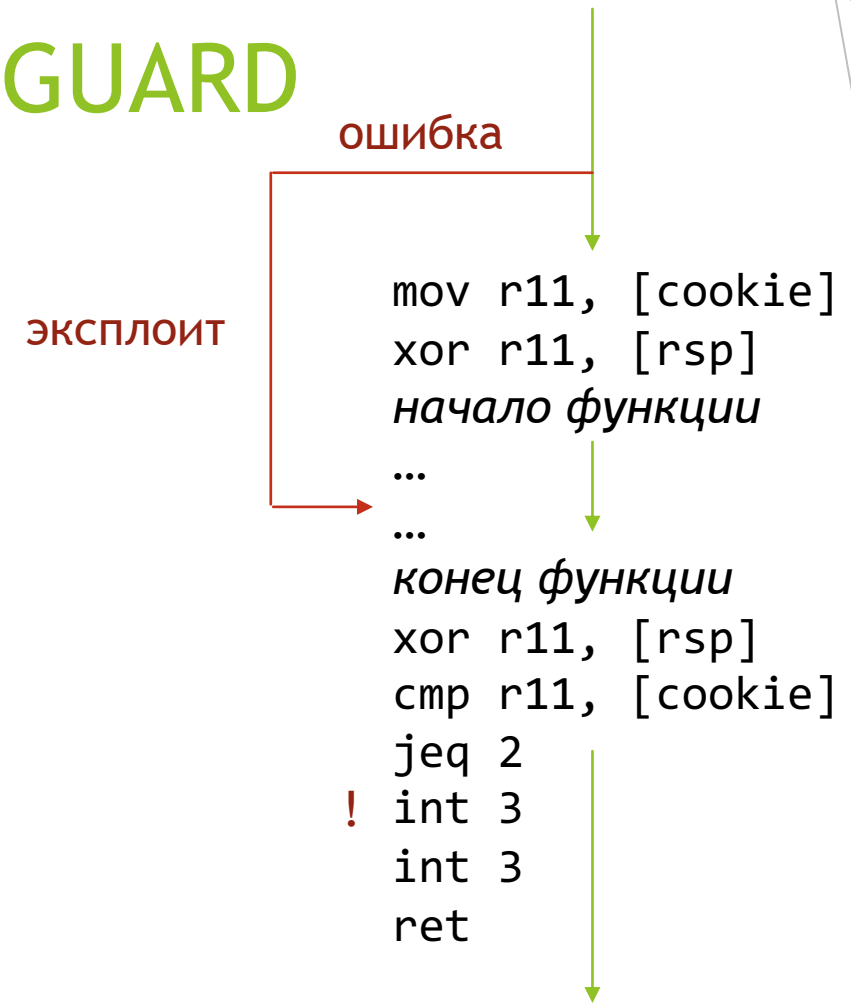
- После заявления Тео де Раадта (июнь 2018) о подозрении на наличие аппаратной уязвимости в реализации технологии SMt, в OpenBSD был разработан патч, позволяющий де-факто отключить SMT на уровне ОС
- Через интерфейс «hw.smt» можно разрешать или запрещать запуск двух потоков на одном ядре одновременно
- В скором времени было официально заявлено о наличии в процессорах Intel, AMD и ARM64 уязвимостей Spectre, Meltdown и прочие

Системный вызов unveil()

- Изоляция доступа к ФС в режиме «белого списка»
- Первым вызовом приложение полностью блокирует доступ к ФС, а последующими открывает доступ к некоторым путям
- Поддержка флагов доступа: можно отдельно открыть доступ на чтение, запись и исполнение, например:
  - /tmp – на запись
  - /bin/sh – на исполнение
  - /var/pool – на чтение

Механизм защиты RETGUARD

- Усложнение выполнения эксплоитов, построенных с использованием заимствования кусков кода и приёмов возвратно-ориентированного программирования – перезапись адреса возврата



File Flags – дополнительный уровень защиты данных

- sappnd. Файл доступен для чтения, его нельзя ни изменять, ни удалять – только дописывать. Флаг может быть установлен только пользователем root.
- uappnd. Аналогичен sappnd, за исключением того, что этот флаг может установить также владелец файла.
- schg. Файл нельзя изменить никоим образов. Устанавливается пользователем root.
- uchg. Аналогичен schg, за исключением того, что этот флаг может установить также владелец файла.
- nodump. Файл с таким флагом должен игнорироваться при создании резервной копии.

Securelevels

Securelevels - механизм ограничения доступных для ОС действий.

- Уровень -1. Дополнительные защитные механизмы отключены
- Уровень 0. Используется только при первой загрузке системы, автоматически переходит на уровень 1
- Уровень 1. Уровень по умолчанию
  - Файлы /dev/mem и /dev/kmem доступны только для чтения
  - Символьные файлы смонтированных устройств доступны только для чтения
  - Флаги файлов sappnd и schg нельзя удалить
  - Нельзя загружать/выгружать модули ядра
- Уровень 2. Максимальный уровень
  - Символьные файлы всех устройств доступны только для чтения
  - Многие настройки сети нельзя изменить
  - Системные часы нельзя откатить назад