

CRIU: дедупликация анонимной разделяемой памяти

Студент: Федор Бочаров
Руководитель: Евгений Баталов

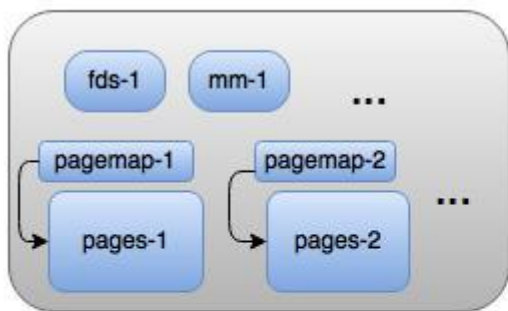
Что такое CRIU

- Набор системных утилит для сохранения и восстановления состояния дерева процессов в Linux
- CRIU позволяет сохранять и восстанавливать:
 - Отношения процессов
 - Память процессов
 - Потoki процессов
 - Идентификаторы процессов
 - Сокеты
 - Объекты IPC
 - Пространства имен Linux, cgroup'ы
 - И многое другое
- Основная часть написана на Си

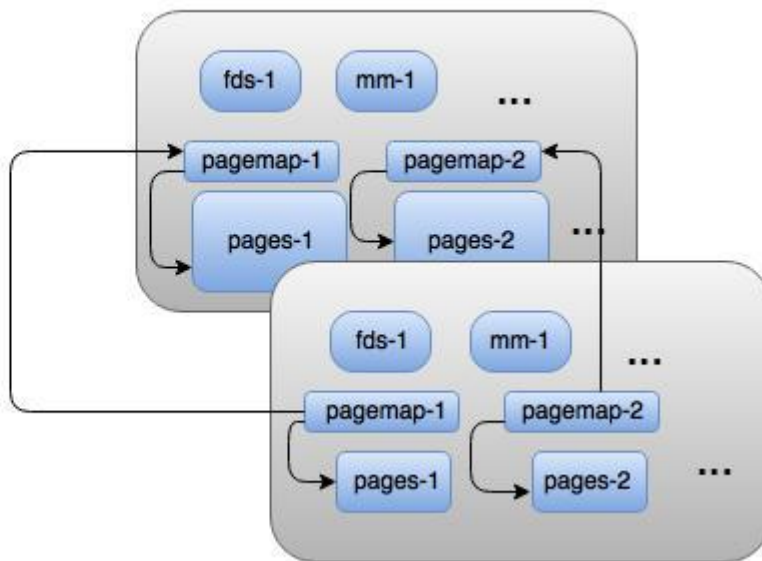


Дампы дерева процессов

Full dump



Incremental dump



Цели и задачи практики

Цель: убрать из дампов деревьев процессов дубликаты содержимого анонимной разделяемой памяти

Задачи:

- Реализовать инкрементальное сохранение содержимого анонимной разделяемой памяти в дампах дерева процессов
- Реализовать удаление копий анонимной разделяемой памяти в существующих дампах дерева процессов
- Реализовать очистку дампа дерева процессов от содержимого анонимной разделяемой памяти в процессе восстановления дерева

<https://github.com/xemul/criu/issues/37>

Что было до меня

- В CRIU реализован аналогичный функционал для анонимной неразделяемой памяти
- Реализованы утилиты для сохранения/восстановления содержимого анонимной неразделяемой памяти (PMC, page-reader, page-xfer)
- Сохранение анонимной разделяемой памяти, находящейся в RSS процесса на каждой итерации инкрементального дампа

Инкрементальное сохранение содержимого анонимной разделяемой памяти

Отслеживаем изменения в анонимной разделяемой памяти между запусками CRIU. Сохраняем в дамп только изменившиеся данные.

Что и как было сделано:

- Реализован алгоритм отслеживания изменений в анонимной разделяемой памяти. Используются биты `PME_PRESENT`, `PME_SWAP`, `PME_SOFT_DIRTY` из `/proc/PID/pagemap`.
- Переиспользованы: `PMC`, `page-reader`, `page-xfer`
- Их код был доработан для использования с двумя видами памяти, а не с одним

Удаление копий анонимной разделяемой памяти в существующих дампах

У нас есть несколько полных дампов дерева процессов. Хотим сделать из них цепочку инкрементальных дампов с удалением всех дубликатов данных.

Что и как было сделано:

- Реализация для анонимной разделяемой памяти была доработана для использования с двумя видами памяти, а не с одним
- Переиспользован page-reader

Очистка дампа во время восстановления дерева процессов

Восстановленная страница памяти процесса удаляется из дампа с помощью вызова к ФС `fallocate(FALLOC_FL_PUNCH_HOLE)`.

Что и как было сделано:

- Реализация для анонимной разделяемой памяти была доработана для использования с двумя видами памяти, а не с одним
- Переиспользован `page-reader`

Результаты

- 10 патчей к CRIU, 2 приняты, остальные находятся на 2ой итерации ревью
- Патчи приняты руководителем
- К новой функциональности в CRIU обычно пишут автотесты. На них времени не хватило. Ручные тесты проходят успешно

Что я получил от практики

- Поработал с production-кодом на Си
- ... и попробовал писать такой же сам
- Получил представление о том, что такое разработка в open-source проекте
- Узнал некоторые возможности git

Спасибо за внимание!