

CRIU: сохранение и восстановление
состояния OFD блокировок

Бегунков П. Г.
Научный руководитель: Баталов Е. А.

Санкт-Петербургский
национальный исследовательский
Академический университет
Российской академии наук

Санкт-Петербург, 2016

- 1 О CRIU
- 2 Постановка задачи
- 3 Блокировки файлов в Linux
- 4 Итоги

CRIU (Checkpoint/Restore In Userspace) - набор утилит ОС Linux, предназначенный для сохранения состояния запущенного приложения/контейнера в образ на диске и дальнейшего его восстановления из образа на произвольном компьютере.



CRIU позволяет сохранять и восстанавливать:

- Отношения процессов
- Память процессов
- Потоки процессов
- Сокеты
- Пространства имен Linux, cgroup'ы
- **Блокировки файлов**
- etc

Цель

Реализация функционала сохранения/восстановления OFD блокировок в CRIU

<https://github.com/xemul/criu/issues/42>

Этапы:

- Исследование способов получения состояний файловых блокировок из пользовательского пространства и/или ядра Linux.
- Создание прототипа восстановления OFD блокировок.
- Интегрирование функционала с CRIU.

Файловые блокировки (file locking) - механизм ограничения/регулирования конкурентного доступа к файлам.

Используется для межпроцессной/межпоточковой синхронизации при работе с файлами.

Типы

- Read - могут удерживаться несколькими процессами одновременно.
- Write - допускается одновременное удержание только одним процессом.

POSIX locks

- Позволяют блокировать часть файла.
- Ассоциированы с процессом. непригодны для межпоточковой синхронизации.

BSD locks (flock)

- Ассоциированы с open file description.
- Блокируют файл целиком.

OFD locks объединяют лучшие качества flock и POSIX lock. Ассоциированы с open file description и позволяет блокировать часть файла.

- Не входит в стандарт POSIX
- Поддерживается Linux с 3.15 (Июнь 2014)
- Поддерживается glibc с 2.20 (Сентябрь 2014)

Сохранение

- Парсинг `/proc/<pid>/fd/<num>`
- В случае устаревшей версии ядра устанавливаем наличие блокировок с помощью `fcntl`.
- Сохранение полученного состояния с помощью `protobuf`.

Восстановление

- Считывание сохраненного состояния.
- Создание блокировок с помощью системных вызовов.

- Добавлена обработка OFD Locks на этапе сохранения состояния приложения.
- Добавлено восстановление OFD lock'ов из сохраненного состояния.
- Обработан случай старой версии glibc и ядра.
- Добавлены автоматические тесты.

Результат работы

- Разработан необходимый функционал в рамках CRIU.
<https://github.com/ISilence/criu/commits/criu-dev>
- Патч находится в состоянии ожидания ревью.

Полученный опыт

- Знакомство с кодом ядра Linux.
- Понимание принципов работы procfs и драйверов.
- Опыт разработки в крупном open-source проекте.