

Лабораторная 5 (6) — отображенные в память файлы, барьеры POSIX

Задача — обработать файл размером, превышающим доступный размер памяти.

Обработка выполняется несколькими потоками.

Файл отображается на память (`mmap()`) блоками.

Размер блока кратен размеру страницы памяти (`page_sz = sysconf(_SC_PAGESIZE)`) и количеству потоков N , соответственно, блок может быть разделен на равные куски, каждый из которых будет обрабатываться своим потоком независимо.

```
chunk_sz = k * page_sz;  
block_sz = chunk_sz * N;
```

Ранг (номер, индекс) главного потока 0, ранги остальных — от 1 до $N-1$. Зная свой ранг, поток может вычислить начало своего куска.

Отображением блоков из файла на память занимается главный поток (с рангом 0).

Он же сбрасывает результат обработки в файл, когда все потоки закончат обработку своего куска. Здесь можно использовать условную переменную на счетчике, который будет устанавливаться главным потоком в N и декрементироваться остальными по мере того, как они заканчивают свой кусок.

Остаток необработанного файла на последнем раунде будет меньше размера блока — эта ситуация должна правильно обрабатываться.

Потоки ожидают завершения отображения нового блока на барьере — как только главный поток подготовит им работу, он приходит последним к барьеру и всех освобождает.

Когда поток в очередном раунде завершает обработку своего куска, он приходит к барьеру ожидания нового.

На последнем раунде потоки самостоятельно определяют, есть ли у них работа и ее размер.

Суть обработки — любая асинхронная работа, например перестановка пар длинных беззнаковых целых по возрастанию (частичная сортировка). Оставшиеся байты последнего раунда остаются как есть.

Большой файл можно получить, читая `dev/random`. Его размер должен превышать в несколько раз объем свободной оперативной памяти и быть, по возможности, простым числом.

Как получить большой файл:

~20Гбайт

```
$ dd if=/dev/random of=random.bin bs=16384 count=1273936  
$ dd if=/dev/random bs=1 count=1783 >> random.bin  
$ wc -c random.bin  
20872169207
```

~200Мбайт

```
$ dd if=/dev/random of=random.bin bs=16384 count=12739  
$ dd if=/dev/random bs=1 count=5845 >> random.bin  
$ wc -c random.bin  
208721621
```

