

bulkmt

Условие

Основная цель переработать Задание 7 так, чтобы вывод обработанных команд выполнялся параллельно. При старте программы должно быть создано, в дополнение к существующему основному потоку, еще три дополнительных. Дадим им условные имена:

- main - основной поток
- log - поток для вывода в консоль
- file1 - первый поток для вывода в файл
- file2 - второй поток для вывода в файл

Основная логика обработки меняется таким образом, что блок команд, сразу после своего формирования должен быть отправлен в консоль и файл одновременно. При этом, отправка блока в файл должна быть распределена между двумя потоками.

Следует обратить внимание на недостаточную точность часов для формирования уникального имени. Необходимо, сохранив timestamp в имени, добавить дополнительную секцию, которая будет гарантирована отличаться.

В условиях многопоточной обработки ведение логов может оказаться затратной процедурой и невольным синхронизатором потоков. В данной задаче необходимо ввести дополнительные счетчики операций, которые должны быть выведены на консоль в конце работы программы. На каждый дополнительный поток два счетчика:

- количество блоков
- общее количество команд в блоках

В основном потоке три счетчика:

- количество строк считанных из файла (команды + { + })
- количество команд
- количество блоков

Для последовательности

```
{
cmd1
cmd2
{
cmd3
cmd4
}
cmd5
```

```
cmd6  
}
```

метрики должны иметь значения соответственно:

- main поток - 10 строк, 6 команд, 1 блок
- log поток - 1 блок, 6 команд
- file1 поток - 1 блок, 6 команд
- file2 поток - 0 блоков, 0 команд

Опционально провести эксперимент по подбору оптимального количества потоков. Необходимо существенно увеличить CPU нагрузку в потоках для записи в файл, например сделать очень большое количество случайных перемешиваний или подсчетов hash. Подойдет любая операция, которая сильно тратит только CPU. Делать эту операцию циклически, наблюдая за скоростью чтения данных с диска, до тех пор, пока чтение с диска существенно не сократится. Вторым этапом будет увеличение количества потоков, увеличивать количество потоков следует медленно, продолжая наблюдать за скоростью чтения с дисков. Должны заметить следующую динамику:

- увеличение CPU стоимости в одном потоке приводит к снижению скорости чтения с диска
- увеличение количества потоков приводит к увеличению скорости чтения с диска
- дальнейшее увеличение количества опять приводит к снижению скорости чтения

Обратить внимание на то, что проведению эксперимента могут помешать любые очереди, которые могут быть в программе.

Самоконтроль

- пакет `bulkmt` содержащий исполняемый файл `bulkmt` опубликован на `bintray`
- отправлена на проверку ссылка на страницу репозитория

Проверка

Задание считается выполненным успешно, если после установки пакета и запуска с тестовыми данными вывод соответствует описанию Задания 7 с дополнительным выводом метрик. Данные подаются на стандартный вход построчно с паузами в 1 секунду для визуального контроля и в существенно большем объеме без пауз.

Будет отмечен способ взаимодействия главного потока с дополнительными. Соответствие выходных метрик поданным на вход данным.