

Алгоритмы для работы с большими объемами данных, практика 1

В данном задании требуется провести измерения скорости работы вашего диска, оперативной памяти и кэшей. Ваше решение должно состоять из двух частей:

1. код программы, выполняющей необходимые измерения;
2. отчет, в котором будет описаны проведенные измерения.

Программа

Ваша программа должна принимать на вход (в качестве аргумента командной строки) тип измерения, которое требуется произвести. На выходе программа должна печатать результат проведенного измерения и его погрешность там где это применимо.

Требуется реализовать следующие проверки:

1. Измерение скорости последовательного чтения данных с диска (выводить надо в Мб/с).
2. Измерение времени сика вашего диска (время надо выводить в миллисекундах, в случае SSD можно в микросекундах).
3. Измерение времени случайного доступа к памяти (время надо выводить в наносекундах).
4. Измерение времени работы кэша/последовательного доступа к памяти (время выводить в наносекундах, либо как скорость в Мб/с). Стоит провести измерение для разных размеров блока, к которому будет производится доступ и показать, как получаемые значения зависят от размера блока.

Отчет

В отчете должно содержаться следующее:

1. Краткая характеристика вашего железа:
 - Объем и модель RAM;
 - Тип жесткого диска, и его характеристики, указанные в спецификации.
2. Описание того, как собирать вашу программу и как она измеряет указанные требуемые величины.
3. Полученные значения измерений на вашем железе.
4. Если значения сильно отличаются от ожидаемых для вашего железа (в десятки раз), то обоснование почему так получилось.

Дополнительная информация

- При работе с файлами в Linux следует пользоваться низкоуровневым API: `fopen`, `fread`, `fwrite`, `fflush`
- При измерение времен работы с диском следует проверить опции монтирования: <http://linux.die.net/man/8/mount>, также надо не забывать очищать кэш операционной системы <http://unix.stackexchange.com/questions/87908/how-do-you-empty-the-buffers-and-cache-on-a-linux-system>.
- Для повышения достоверности измерения стоит проводить несколько раз, после чего вычислять среднее значение и среднеквадратичное отклонение.
- При произведении замеров, желательно максимально убрать всю фоновую нагрузку на вашем компьютере, или если это невозможно, то попытаться оценить её влияние на получаемые результаты.