Fabian Klopfer (956507)|Simon Suckut (710134)

Exercise 1: Free Blocks

a. Name an advantage and a drawback for each block list and block bitmap.
Linked List: + no external fragmentation, can grow in file size - no pre-allocation, mostly not contingous
Bitmap: + smaller, faster look-up, pre-allocatable - not growable (reallocation for additional blocks possible)

b. 8192MB DB file(s), 16KB blocksize, block bitmap. How many KB are needed to reference all blocks?
$(2^{13} << 10)KB >> 4\frac{KB}{block} = (2^{13} << 6)blocks = 2^{19}blocks$ and
$1\frac{bit}{block} \cdot 2^{19}blocks = 2^{19}bit = (2^{19} >> 8)byte = (2^{11} >> 10)KB = 2^1 KB = 2KB$

c. see modified source file

Exercise 2: Buffer Management: OS vs. DB

a. Virtural memory is designed to store files of arbitary size and large objects like videos, images, computer graphics models for gaming and so on. It is also designed to map devices to logical addresses in order to control the hardware

b. The buffer manager is optimized for data of a specific kind, e.g. RDBMS are optimized for fixed size data, elastic serach is optimized for documents, ...

c. The buffer manager and Virtual memory have some things im common. They both split the data in Pages and they both swap pages to the disk if it is not needed at the moment.

d. a combination, e.g. if one starts at the mindset of a relational db and wants to search documents, it seems reasonable to index and store the files differently. There are also interesting methods for semantic clustering e.g. hebbian learning in order to find frequent itemsets. But the virtual memory can not replace the buffer manager as the database has not enouth controll over it.

Exercise 3: Measuring Performance

a. $\sum_{i=0}^{100000000}(i)^2$

b. The runtimes vary, but mostly decrease due to vm optimization/caching

c. The benchmark framework JMH creates multiple forks of the JVM and per fork multiple warm-up iterations in order to reach a stable state for the benchmark. Further it measures the performance of the whole system in mean runtime and standard deviation (and other things)