Task 1

Aircraft
identifier: 4 byte integer
manufacturer identifier: 1 byte
serial number: 4 byte integer
cargoVolume: 4 byte integer (liters)
fuelCapacity: 4 byte integer (liters)
fuelConsumption: 4 byte integer (liters per 100km)
weightTotal: 4 byte integer (kilograms)
emissionsCO2: 4 byte integer (grams per kilometer)
retired: 1 byte (boolean)

Together 30 bytes, i.e. 32 bytes aligned.

Task 2 - indices

A) Primary key index

index entry has size of key+pointer $= 8B$

blocking factor = $\frac{4KiB}{8B} = 512$

number of blocks in primary file $= \frac{4KiB}{32B} = 128$

that is equal to $\lceil \frac{20000000}{128} \rceil = 156,250$ blocks

the number of blocks in which each level is stored

1. level $\lceil \frac{156.250}{512} \rceil = 306$
2. level $\lceil \frac{306}{512} \rceil = 1$

height of the index

2

the total size of the index in MB

total blocks $= 307$

size $\approx 1.2 MiB$

B) Secondary key direct index

Appropriately select a secondary key: Serial number

blocking factor stays the same, serial number has the same size as PK.

the number of blocks in which each level is stored

- 0. level $\lceil \frac{20,000,000}{512} \rceil = 39,063$
- 1. level $\lceil \frac{39,063}{512} \rceil = 77$
- 2. level $\lceil \frac{77}{512} \rceil = 1$

height of the index

3

the total size of the index in MB

total blocks = $39,141$
size $\approx 152.8 MiB$

C) Secondary key indirect index

We have the same blocking factor, because PK size is same as that of a page pointer.

The indirect index would have in this case have exactly the same properties as the direct secondary key index.

Comparison

The primary index is much smaller and has one less level of jumping.

Task 3 - lookup time

$T = 2 \cdot (s + r + btt) + 2 \cdot r + btt$
$btt = \frac{2 \cdot r}{TC \cdot B}$

Where

$r = 4.17ms$
$TC = 0.3MB$
$btt = \frac{2 \cdot 0.00417}{0.3} \cdot 0.004 \approx 0.1104ms$
$s = average\ seek\ time = 8.5ms$

Then

$T = 2 \cdot (8.5ms + 4.17ms + 0.1104ms) + 2 \cdot 4.17ms + 0.1104ms$
$T \approx 34.01ms$

Task 4 - bitmap

We can use bitmap to store manufacturer identifier to optimize for queries over vehicles of a given manufacturer.

1. Use two columns to store the number.

We have to load both columns to query the identifier.

The size of the bitmap:

Entries per block: $16,384$
Blocks: 1221
Size: $4.8MiB$

2. Use four separate columns, one for each value

To find all vehicles of one manufacturer, we have to load only one column. However, we have twice as many columns, so the sizes doubles to $9.6MiB$.