Homework 3

Simon Zheng
260744353

April 5th, 2018

1

1.1 a)

1.1.1 i)

- In general: 3000 (page) accesses (have to search everywhere)
- Concrete values: 3000 (page) accesses (same reason)

1.1.2 ii)

- In general: 3000 (page) accesses (unclustered index for range search not very useful)
- Concrete values: 3000 (similar reason)

1.1.3 iii)

- In general: 1+10+1
- Concrete values:

1.1.4 iv)

- In general:
- Concrete values:

1.2 b)

Yes, especially on price, as it is very useful for range look-ups. It allows us to make less I/O since we can use the assumption the range is in order on the pages.

2

2.1 a)

$200,000 + 200,000 \times (1\ data\ page + 1\ leaf\ page)) = 600,000$

2.2 b)

$20,000 + \frac{20,000}{51-1} \times 200,000 = 80020000$

2.3 c)

$$1 + \lceil \log_{50-1}(\frac{20,000}{50}) \rceil = 3$$
$$1 + \lceil \log_{50-1}(\frac{200,000}{50}) \rceil = 4$$
$$200,000 \times 4 + 20,000 \times 3 + 200,000 + 20,000 = 1,080,000$$

2.4

Number of output tuples: $Parts \times Catalog = 4,000,000,000$

3

3.1 a)

$\pi_{sname}\left(\left(\sigma_{country\ =\ 'China'}(Suppliers) \times \sigma_{pname\ =\ 'bearing'}(Parts)\right) \bowtie Catalog\right)$