CS 186 Fall 2024

Introduction to Database Systems Alvin Cheung

DIS 6

1 Assorted Joins

- Companies: (company_id, industry, ipo_date)
- NYSE: (company_id, date, trade, quantity)

We have 20 pages of memory, and we want to join two tables Companies and NYSE on C.company_id = N.company_id. Attribute company_id is the primary key for Companies. For every tuple in Companies, assume there are 4 matching tuples in NYSE.

NYSE contains [N] = 100 pages, NYSE holds $p_N = 100$ tuples per page.

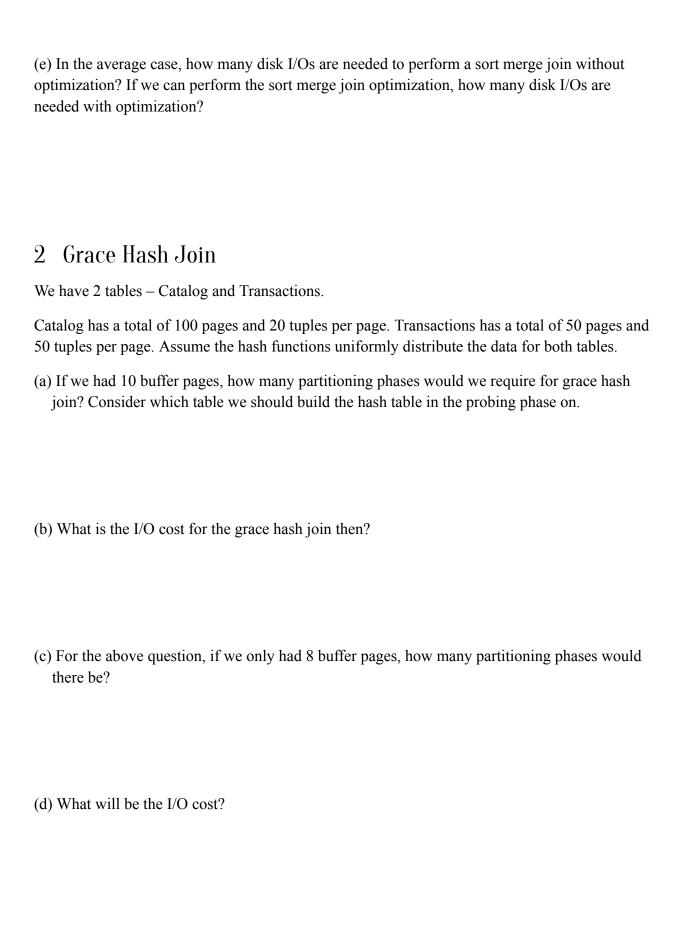
Companies contains [C] = 50 pages, C holds p_C = 50 tuples per page.

There are alternative 3 unclustered B+ tree indexes of height 1 on C.company_id and N.company_id. Throughout the problem, do not assume any caching of index nodes.

- (a) How many disk I/Os are needed to perform a simple nested loop join?
- (b) How many disk I/Os are needed to perform a block nested loop join?

- (c) How many disk I/Os are needed to perform an index nested loop join?
- (d) For this part only, assume the index on NYSE.company_id is clustered. What is the cost of an index nested loop join using companies as the outer relation?

CS 186 Fall 2024 DIS 6



CS 186 Fall 2024 DIS 6

3 Relational Algebra

	~ '	. 1	.1	1	
•	Onci	Idar	· tho	cch	ama
•	Consi	luci	unc	SOID	uma.

CS 186 Fall 2024 DIS 6