Consider the join \( \text{SELECT * FROM R, S WHERE R.a = S.b} \). The following information is known:

- Relation \( R \) contains 10,000 tuples and has 10 tuples per page.
- Relation \( S \) has 2,000 tuples and has 10 tuples per page.
- Attribute \( b \) of relation \( S \) is the primary key for \( S \).
- Both relations are stored as simple heap files.
- Neither relation has any indexes built on it.
- 53 buffer pages are available.

The cost metric is number of page I/Os, and the cost of writing out the result of join is ignored for all algorithms.

a) What is the cost of joining \( R \) and \( S \) using a block nested loops join?
b) What is the minimum number of buffer pages required for the cost in a) to remain unchanged?

c) What is the cost of joining R and S using a sort-merge join?
d) What is the minimum number of buffer pages required to sort R in two passes, assuming that in the first pass (phase 1) runs of length B are generated, where B is the size of the buffer?

**Turn-in (due before class on 23 January 2009)**

1. The completed assignment coversheet. Your comments will help us improve the course.
2. A hard copy of the answers to exercises.