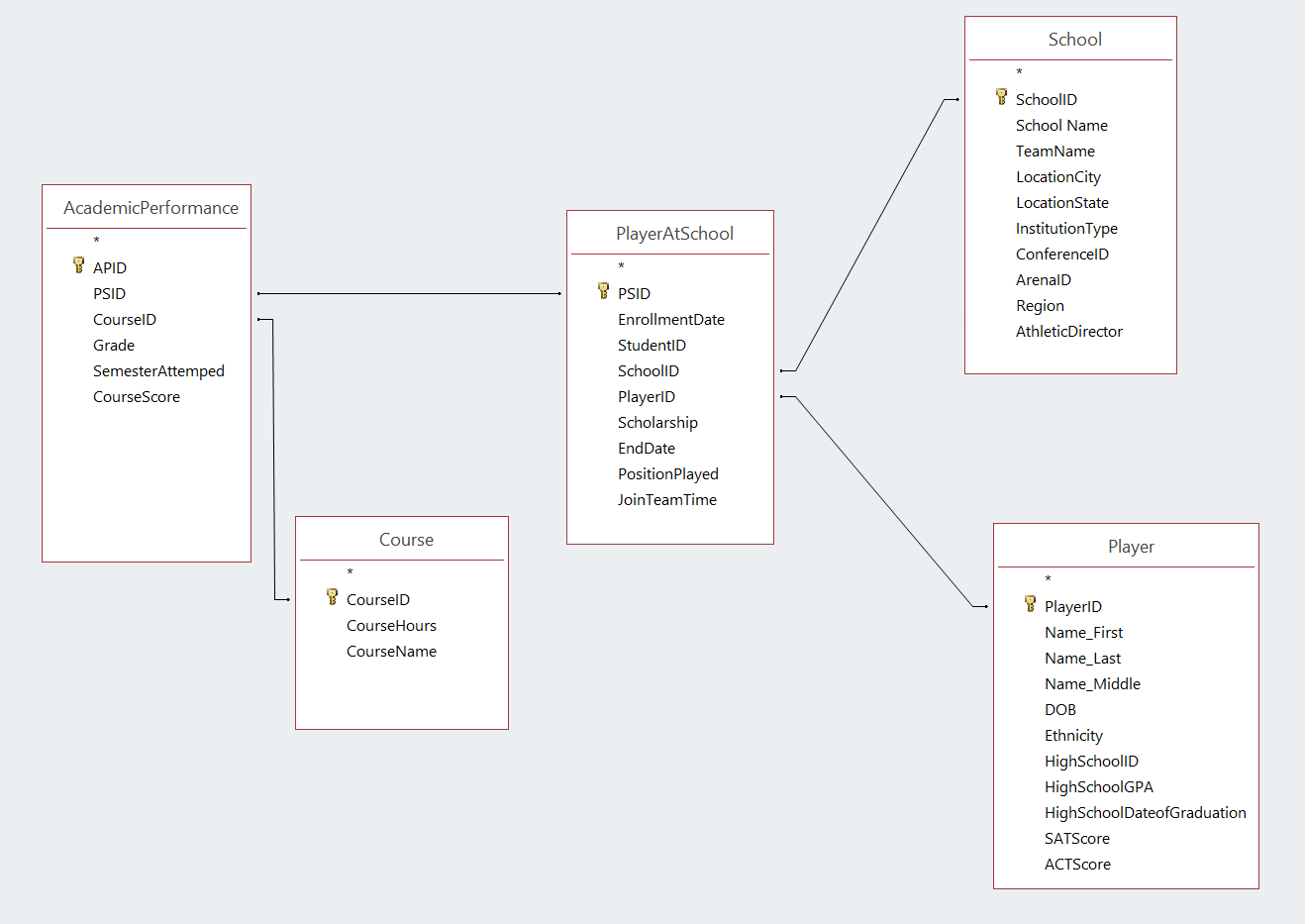
Group 6: Physical Database Design



Course (CourseID, CourseName)

AcademicPerformance (CourseID, SemesterAttempted)

Assume sequential access time is 1 ms and random access time is 10 ms. The size of a block is 8K. Assume a predominantly transaction-based environment. If the scan time is less than 0.25 seconds (i.e., 250 ms), then you don’t need to index *any* field since the query time is limited.

Assume the following number of records and size of a record in each file of the active database:

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Table | Records | Length | Bf | No. Blocks | Scan Time |
| Course: | 10,000; | 50 bytes | 8,000/50=160 | 10,000/160=62.5 | 62.5/1,000=0.06s |
| AcademicPerformance: | 30,000,000; | 120 bytes | 8,000/120=66.6 | 30,000,000/66.6=450,450.45 | 450,450.45/1000=450s |

Formulas used:

BF = (Block Size) / (Record Length)

No. Blocks = (Records) / (BF)

Scan Time = (No. Blocks) \* (Scan Time / Block = 1ms)

Answer the following (and show the calculations): 1) Consider access through the following four attributes in a single-table query with restrict condition of type attribute = value: Show calculations for each field and decide if you need to have an index on each of the following four fields.

The following are the distinct values for some of the attributes:

CourseHours: 4

SemesterAttempted: 3

|  |  |  |  |
| --- | --- | --- | --- |
| Field | No. of Records Touched (n) | Index Time | Index (I) or Scan (S) |
| Course\_CourseID | Small table | N/A | Scan |
| Course\_CourseHours | Small table | N/A | Scan |
| AcademicPerformance\_CourseID | 30,000,000/10,000=3,000 | 3,000/100=30s | Index |
| AcademicPerformance\_SemesterAttempted | 30,000,000/3=10,000,000 | 10,000,000/100=100,000s | Scan |

Formulas used:

n = (Number of Records of Table) / (Number of Records in Attribute’s Table)

s = ( n + 2 ) \* (10 ms) OR, for larger n… s = (n) / 100 = sec

If the scan time is less than 0.25 s (250ms) then you don’t need to index, you scan