Quiz Sheet No. 5 for *Architecture and Implementation of Database Systems*
Prof. Rudolf Bayer, Ph. D.
Institut für Informatik
SS 2003

Exercises for **Chapter 6: Distributed DBS**

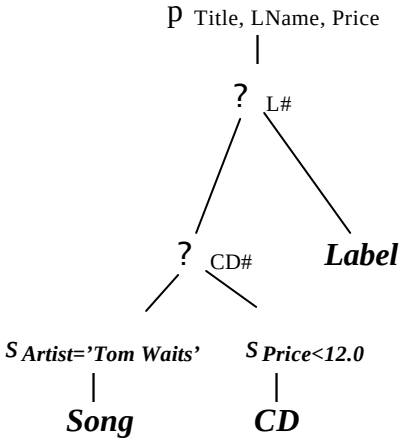
1. Consider the database schema:

```
create table CD
( CD#        integer,
  Title      string,
  Price      real,
  L#         integer )     key is CD#

create table Label
( L#         integer,
  LName      string,
  LCountry   string  )     key is L#

create table Song
( CD#        integer,
  Track#     integer,
  STitle     string,
  Artist     string  )     key is (CD#, Track#)
```

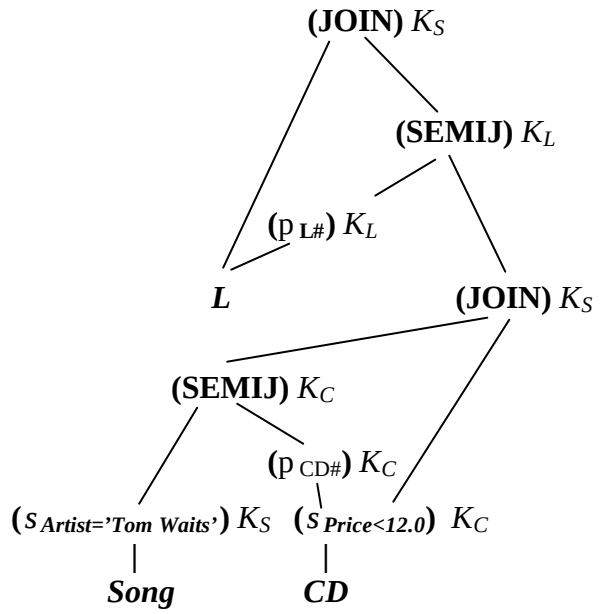
a) Given below is an operator tree for query "*Retrieve the Title, LName and Price of CDs that cost less than $12 and contain songs by the Artist Tom Waits*":



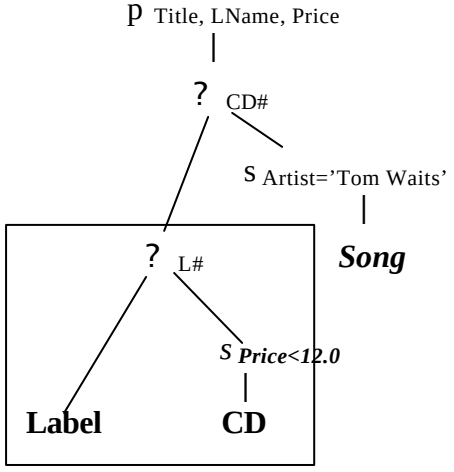
Assume a distributed DBS with nodes K_L (holding relation Label), K_C (holding relation CD) and K_S (holding relation Song). Sketch an operator graph using semi-joins for executing this query on the distributed DBS.

Answer:

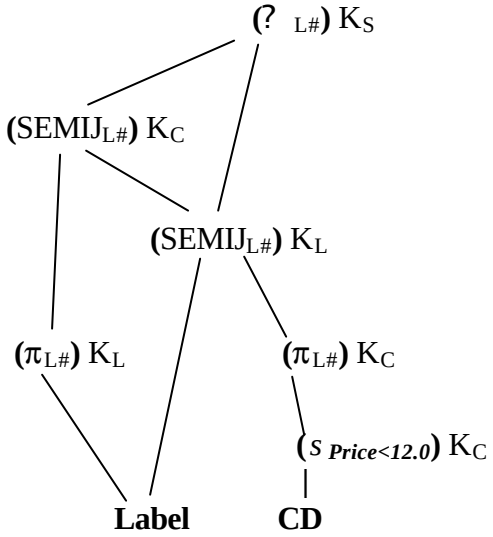
There is a large variety of possibilities for such an operator graph. Given below is only one of these solutions:



b) The following operator tree is another representation of the query in exercise 1a):



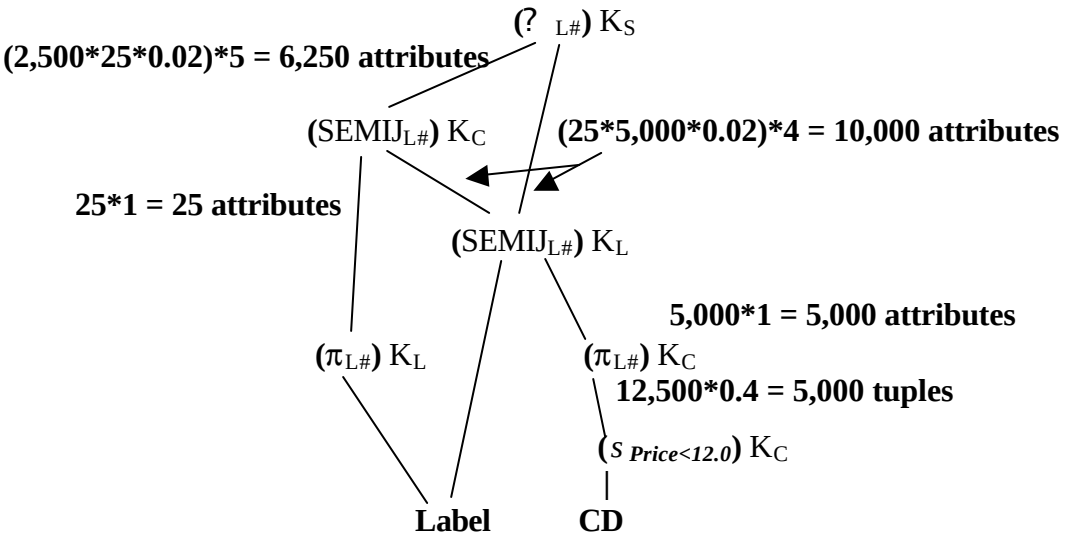
A possible operator graph for the join Label? CD, designed for the distributed DBS hosting K_L, K_C and K_S is:



Compute the transport volume (the number of attributes, regardless of their types) that must be sent from node to node and determine the total transport volume of this query plan. Assume that:

|Label| = 25
|CD| = 12,500
Selectivity ("*Price < 12.0*") = 0.4
Selectivity for all semi-joins = 0.02

Answer:
Assuming that *both* join attributes in semi joins are being sent from node to node:



total transport volume: 5,000 + 10,000 + 25 + 6,250 = 21,275 attributes