CSE 4125: Distributed Database Systems
Chapter – 6
(Part – C)

Optimization of Access Strategies.

Topics to be discussed -

- Query Optimization
- Data Transmission
- Comparison between different Query Optimization Strategies

Query Optimization

- Permutation of the ordering of operations within a query can provide many equivalent **strategies** to execute it.

- Finding an "**optimal**" ordering of operations for a given query is important.
  - –Done by query optimization layer(or optimizer for short).

Data Transmission

- Data transmission requirement can be evaluated by –

➢ **Transmission cost**

–i.e. cost to initiate a transmission, routing cost etc.

➢ **Transmission delay**

–i.e. elapse time between activation and completion of an app.

Data Transmission (cont.)

- Data transmission requirement can be evaluated by –

- **Transmission cost**

  $$TC\ (x) = C_0 + x * C_1$$

- **Transmission delay**

  $$TD\ (x) = D_0 + x * D_1$$

x = Transmitted data

C's and D's are system dependent constants.

C_0 = initialization fixed cost
C_1 = network wide unit cost
D_0 = connection initialization fixed time
D_1 = network wide unit transfer rate

Data Transmission (cont.)

- Data transmission requirement can be evaluated by *(more detailed characterization)*–

➢ **Transmission cost**

$$TC\ (x) = C_0^{ij} + x * C_1^{ij}$$

Total amount of data

➢ **Transmission delay**

$$TD\ (x) = D_0^{ij} + x * D_1^{ij}$$

Total / Longest amount of data (Total = Not parallel, Longest = Parallel execution)

i and *j* denote source and destination respectively.

Comparison between different strategies

- ✓ We will see different versions of a query.
- ✓ Measure their cost and delay to see which one is better.

Example

- You will be given a query with fragmentation schema and their database profile.
- Draw operator tree for the query.
- Apply strategy 1 as per the question
- Apply strategy 2 as per the question
- Compare between strategies

Scenario

For the following query compare strategy 1 and strategy 2 based on the TC and TD at site 2.

$$\mathbf{PJ}_{\text{SNUM}}\ ((\mathbf{PJ}_{\text{SNUM, DEPTNUM}}\ \mathbf{SUPPLY})\ \mathbf{JN}_{\text{DEPTNUM=DEPTNUM}}\ (\mathbf{PJ}_{\text{DEPTNUM}}\ \mathbf{SL}_{\text{AREA="North"}}\ \mathbf{DEPT}))$$

Assume that SUPPLY has two horizontal fragments and DEPT has three horizontal fragments. The database profiles of $SUPPLY_1$ and $SUPPLY_2$ are similar. Also, profiles of $DEPT_{1,}$ $DEPT_2$ and $DEPT_3$ are identical. The database profiles are shown on the next slide.

Determine the transmission delay for the above strategy when network-wide transfer rate is 10000 bits/second. Note that, the system provides the benefit of parallel processing and the initial delay is 0.

Scenario

- **Profiles of $SUPPLY_1$ and $SUPPLY_2$:**

card ($SUPPLY_1$) = 30000
card ($SUPPLY_2$) = 20000

site($SUPPLY_1$) = 1
site($SUPPLY_2$) = 4

	snum	pnum	deptnum	quan
size	6	7	2	10
val	1800	1000	20	500

- **Profiles of $DEPT_1$, $DEPT_2$ and $DEPT_3$:**

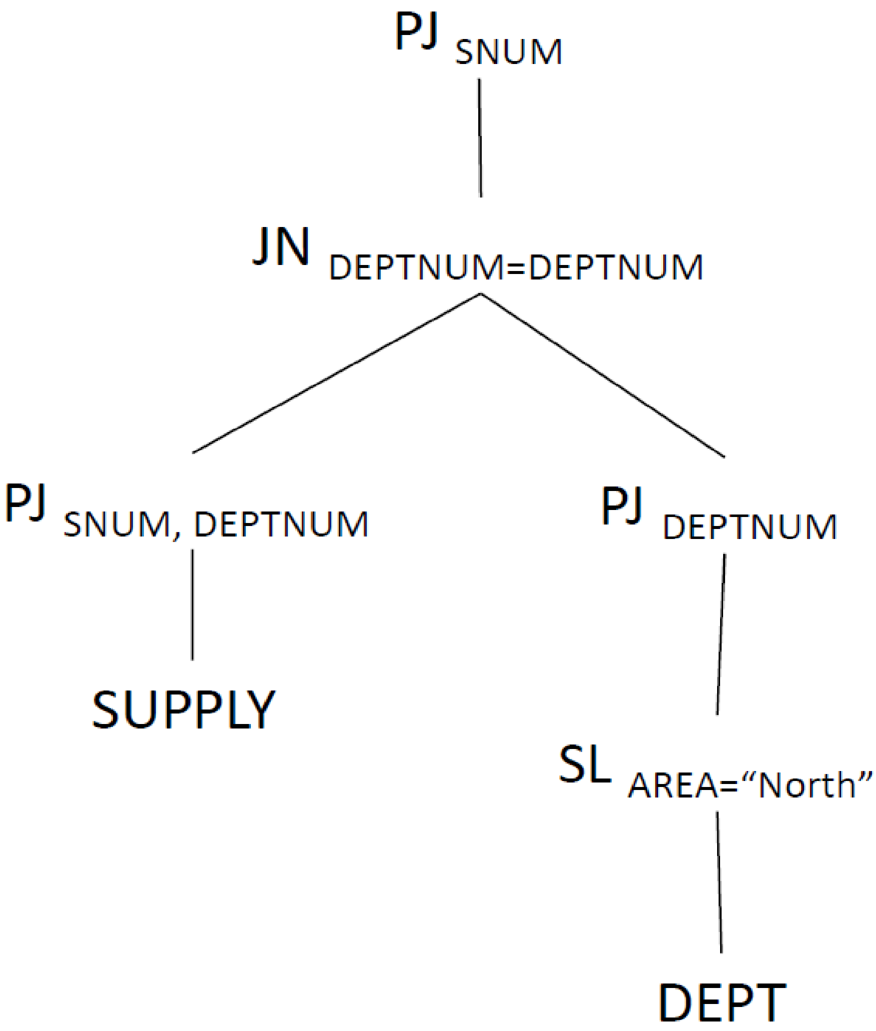
card ($DEPT_1$) = 10
card ($DEPT_2$) = card ($DEPT_3$) = 20

site($DEPT_1$) = 2
site($DEPT_2$) = 3
site($DEPT_3$) = 5

	deptnum	name	area	mgrnum
size	2	15	1	7
val	10	10	2	10

Given Scenario

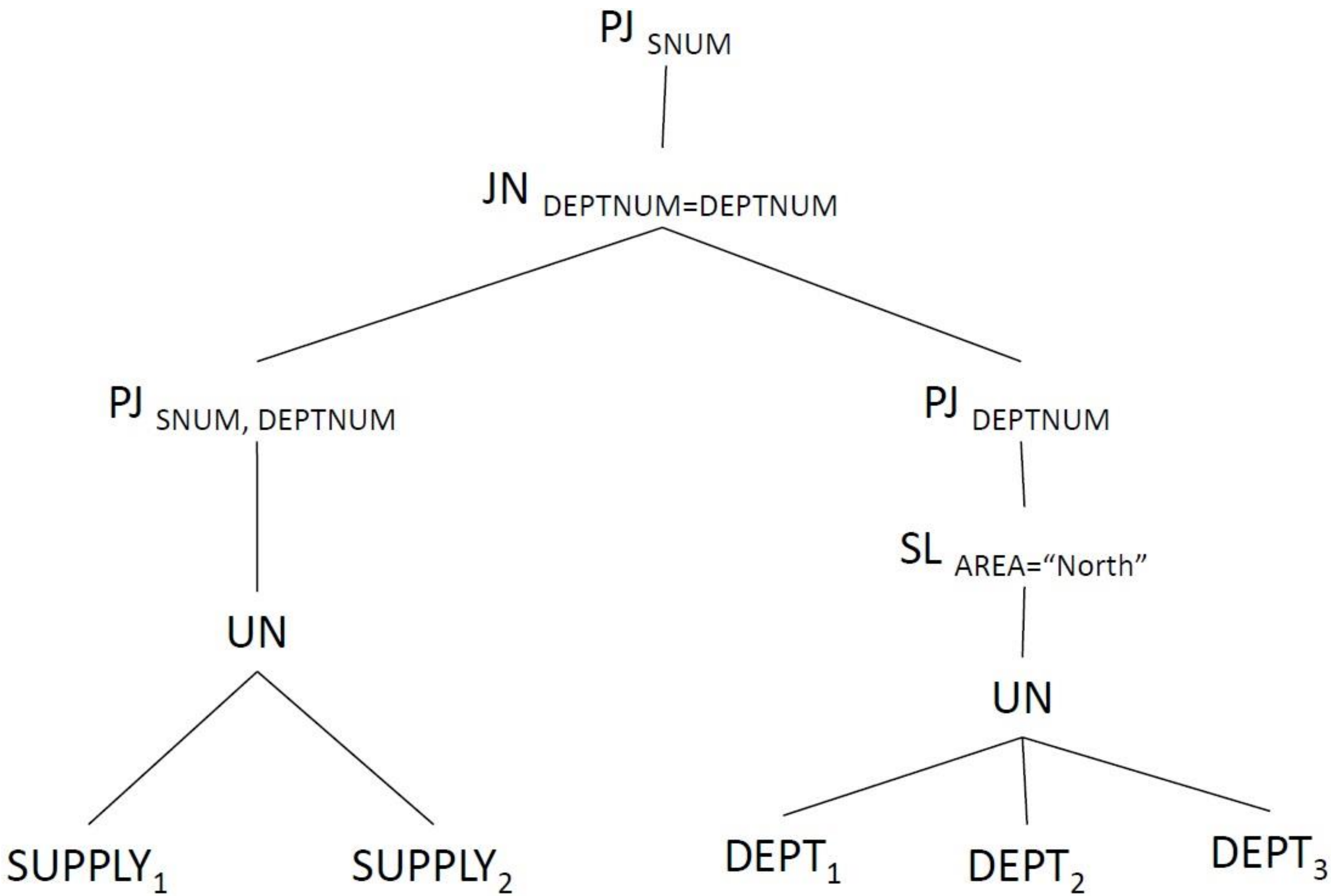
Input Query (Operator Tree):



Strategy – 1

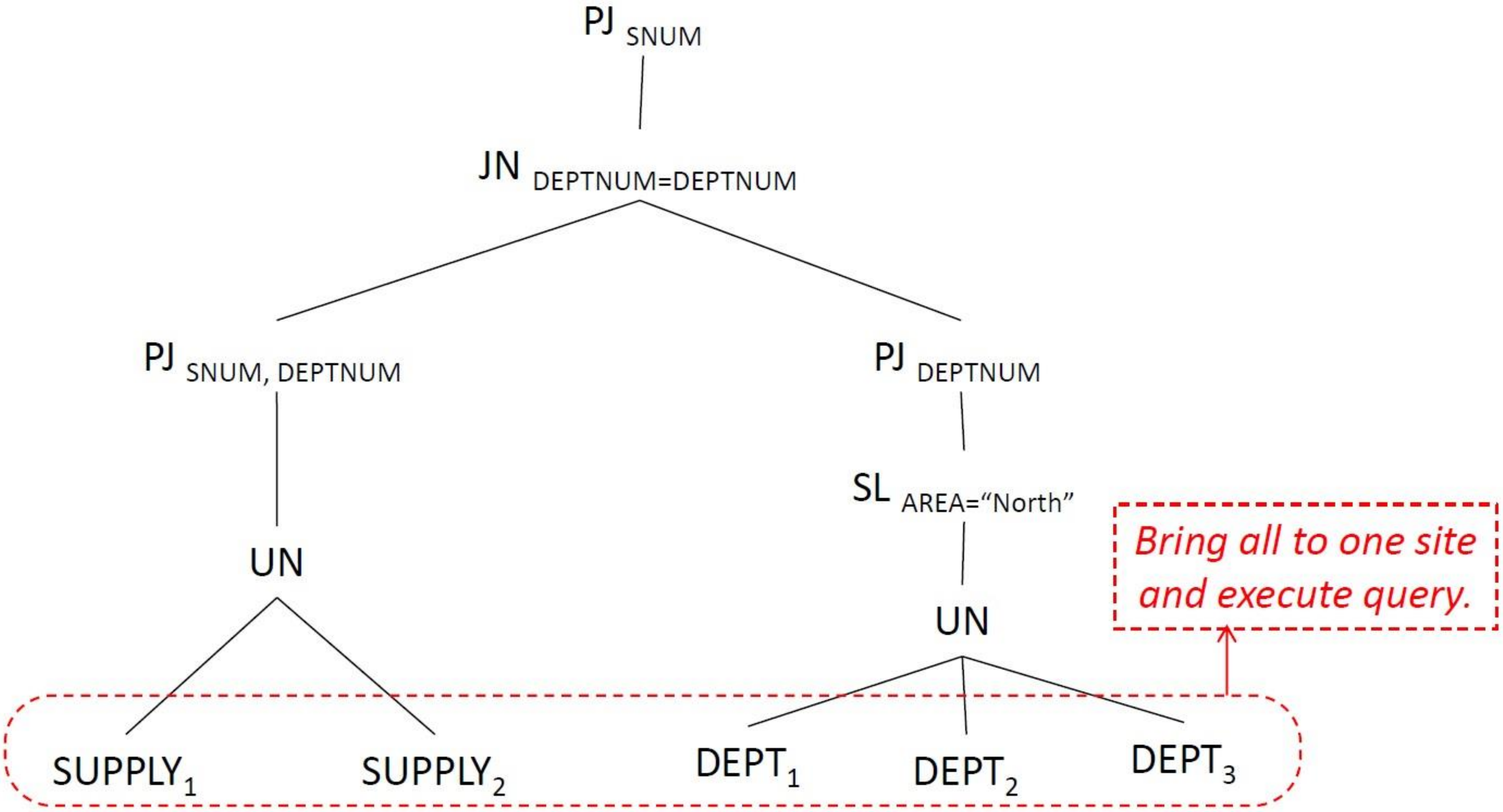
No simplification

Strategy – 1



PJ SNUM
JN DEPTNUM=DEPTNUM
PJ SNUM, DEPTNUM
PJ DEPTNUM
SL AREA="North"
UN
UN
SUPPLY1
SUPPLY2
DEPT1
DEPT2
DEPT3

Strategy – 1 (cont.)



Strategy – 1 (cont.)

- Let us execute the query at site 2.
  - –We need to collect all the fragments there (in parallel).

Strategy – 1 (cont.)

- Let us execute the query at site 2.
  - –We need to collect all the fragments there (in parallel).

- Transmitted amount:

  $X_{SUPPLY1} = ?$

card (SUPPLY$_1$) = 30000

	snum	pnum	deptnum	quan
size	6	7	2	10
val	1800	1000	20	500

Strategy – 1 (cont.)

- Let us execute the query at site 2.
  - –We need to collect all the fragments there (assume in parallel).

- Transmitted amount:

$$X_{SUPPLY1} = card(SUPPLY_1) \times size(SUPPLY_1) \times 8 \text{ bits}$$
$$= 30000 \times (6+7+2+10) \times 8 \text{ bits}$$
$$= 30000 \times 25 \times 8 \text{ bits}$$
$$= 6000000 \text{ bits}$$

Strategy – 1 (cont.)

- Transmitted amount for other fragments:

$X_{SUPPLY2}$ = ?

X_{DEPT1} = ?

X_{DEPT2} = ?

X_{DEPT3} = ?

card ($SUPPLY_2$) = 20000

	snum	pnum	deptnum	quan
size	6	7	2	10
val	1800	1000	20	500

card ($DEPT_1$) = 10
card ($DEPT_2$) = card ($DEPT_3$) = 20

	deptnum	name	area	mgrnum
size	2	15	1	7
val	10	10	2	10

Strategy – 1 (cont.)

- Transmitted amount for other fragments:

$X_{SUPPLY2}$ = 20000 × 25 × 8 bits = 4000000 bits

X_{DEPT1} = 0 bits

X_{DEPT2} = 20 × 25 × 8 bits = 4000 bits

X_{DEPT3} = 20 × 25 × 8 bits = 4000 bits

Strategy – 1 (cont.)

Assume $C_0 = 0$ and $D_0 = 0$

$TC(x) = C_0 + (\text{sum of all the amount}) \times C_1$

$= (X_{SUPPLY1} + \ldots + X_{DEPT3}) \times C_1 = 10008000 \times C_1$

$TD(x) = D_0 + (\text{largest amount}) \times D_1$

$= 6000000 \times D_1$

If $D_1 = 10000$ bit/second,

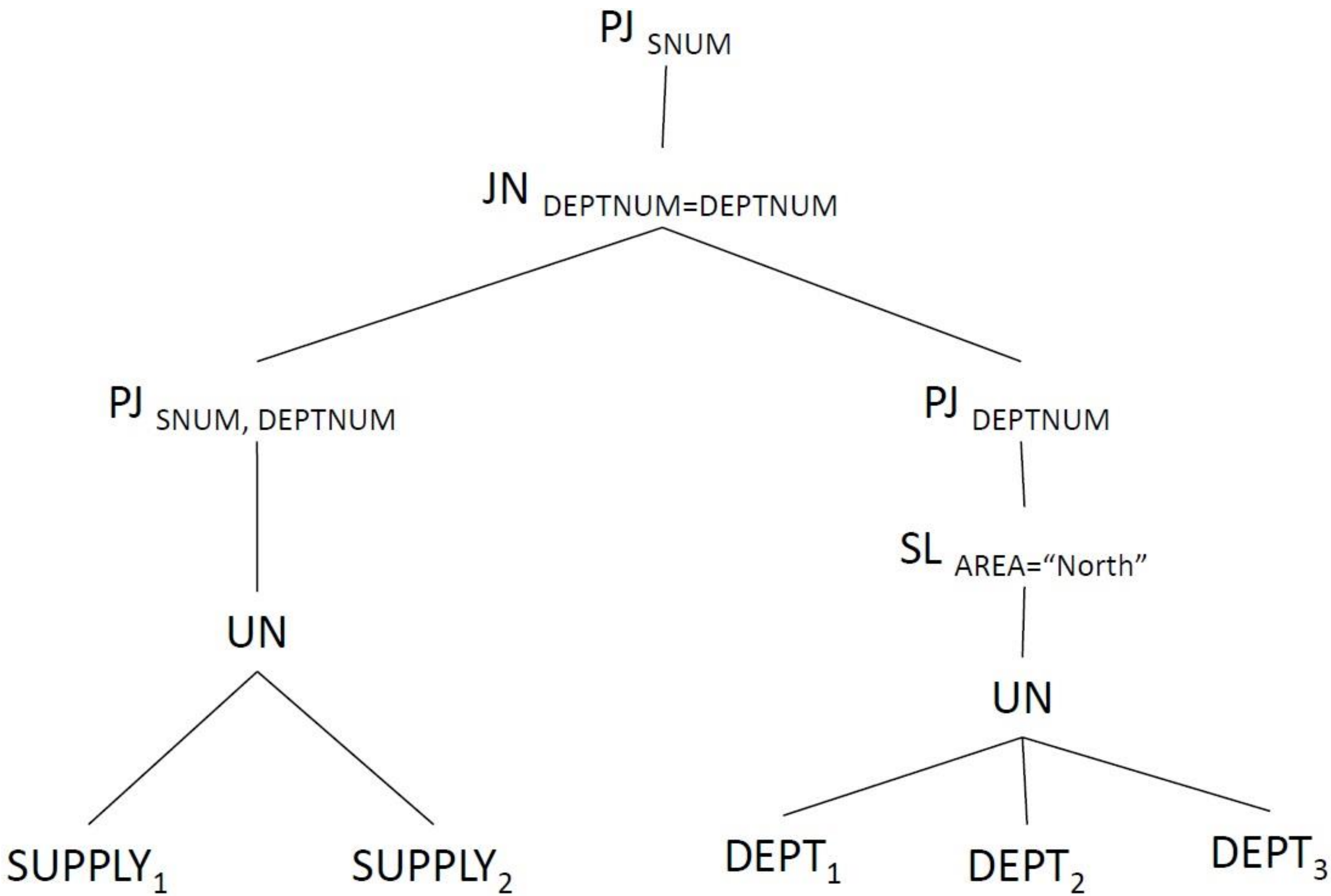
Transmission Delay = $6000000 \times (1/10000)$ s = 600 s = 10 mins

Strategy – 2

Simplification applied

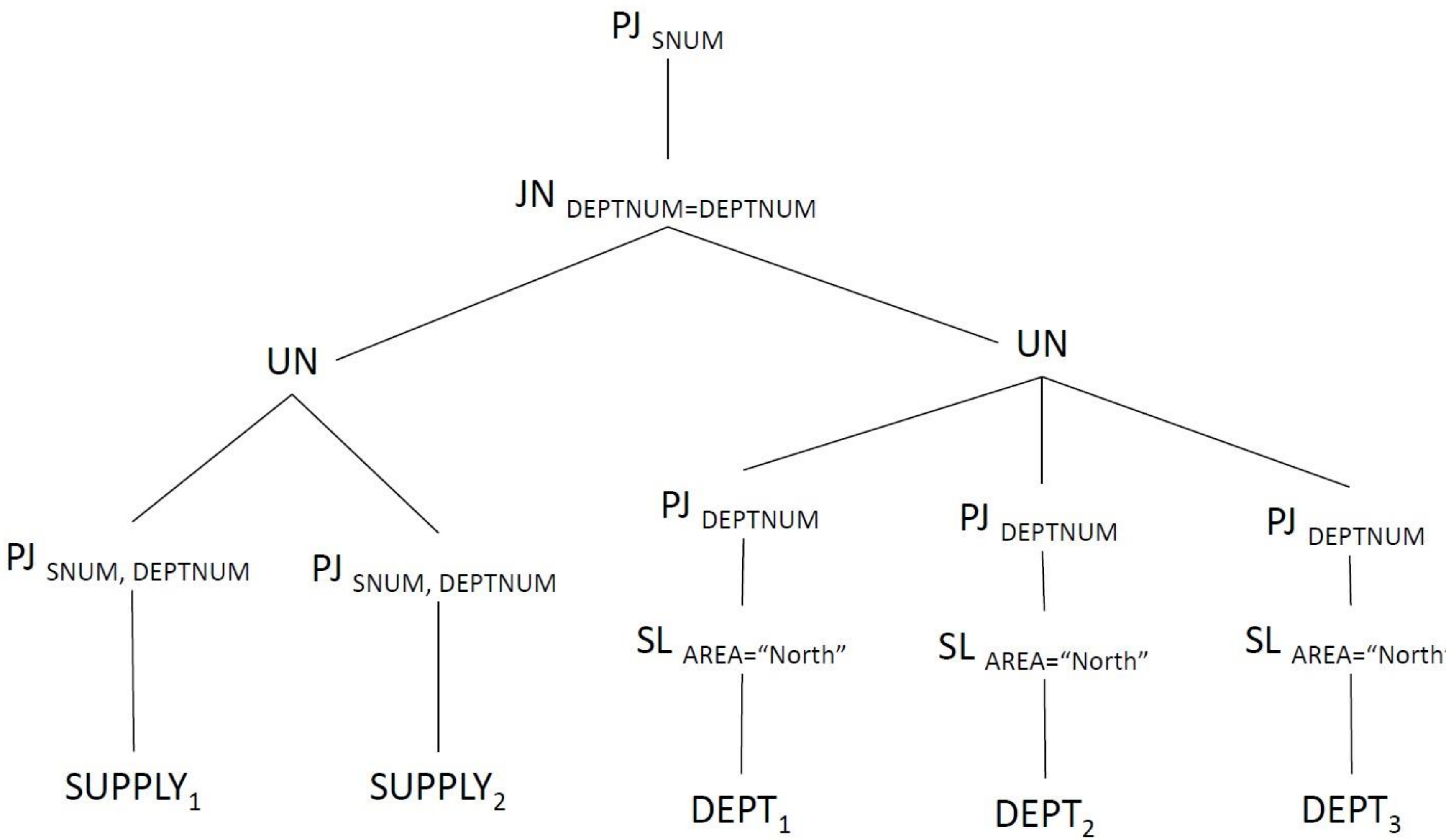
(Rules & Criterion -1 and 2).

Strategy – 2



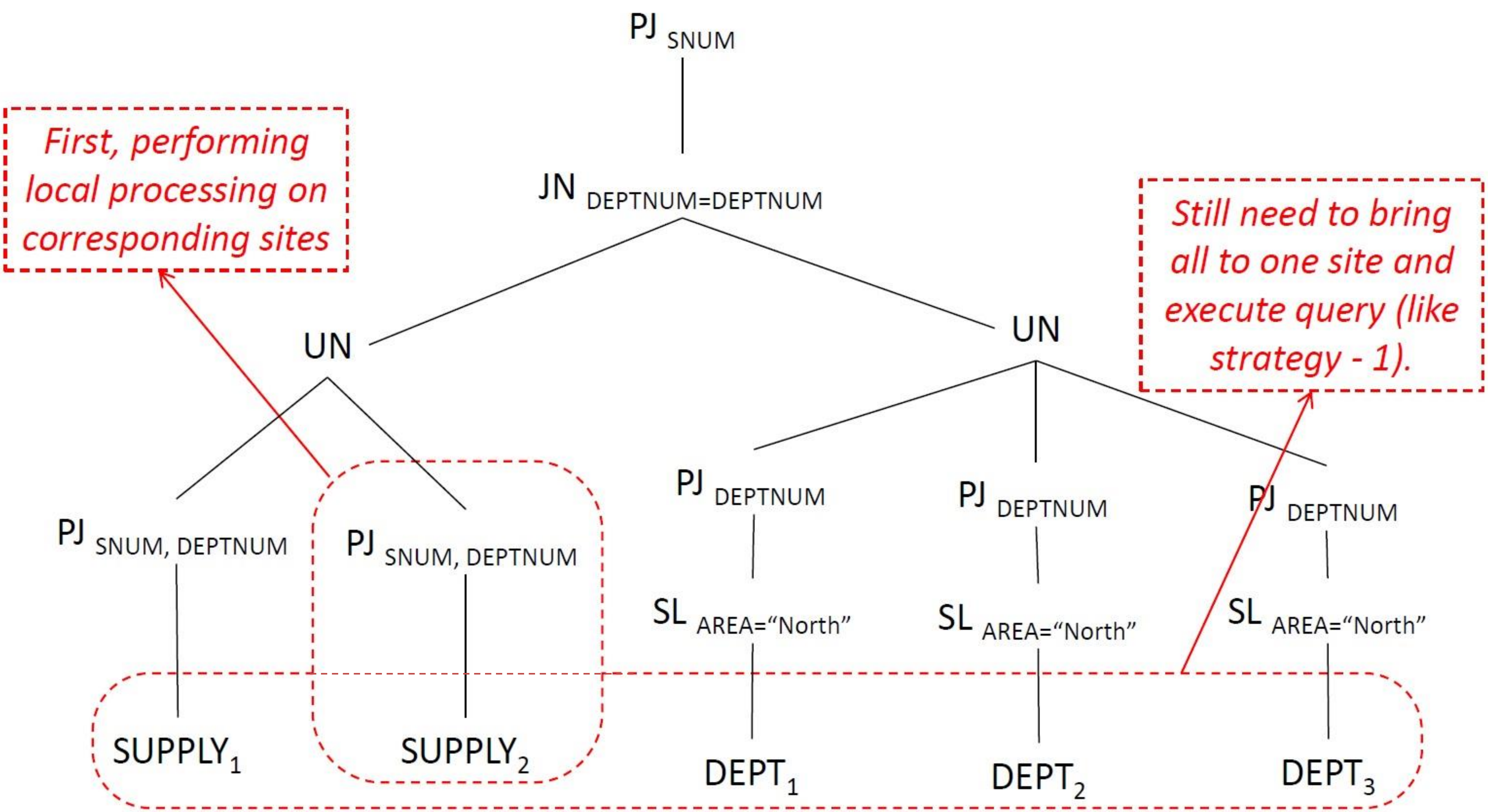
PJ SNUM
JN DEPTNUM=DEPTNUM
PJ SNUM, DEPTNUM
PJ DEPTNUM
SL AREA="North"
UN
UN
SUPPLY1
SUPPLY2
DEPT1
DEPT2
DEPT3

Strategy - 2



PJ SNUM
JN DEPTNUM=DEPTNUM
UN
UN
PJ SNUM, DEPTNUM
PJ SNUM, DEPTNUM
PJ DEPTNUM
PJ DEPTNUM
PJ DEPTNUM
SL AREA="North"
SL AREA="North"
SL AREA="North"
SUPPLY1
SUPPLY2
DEPT1
DEPT2
DEPT3

Strategy – 2 (cont.)



Strategy – 2 (cont.)

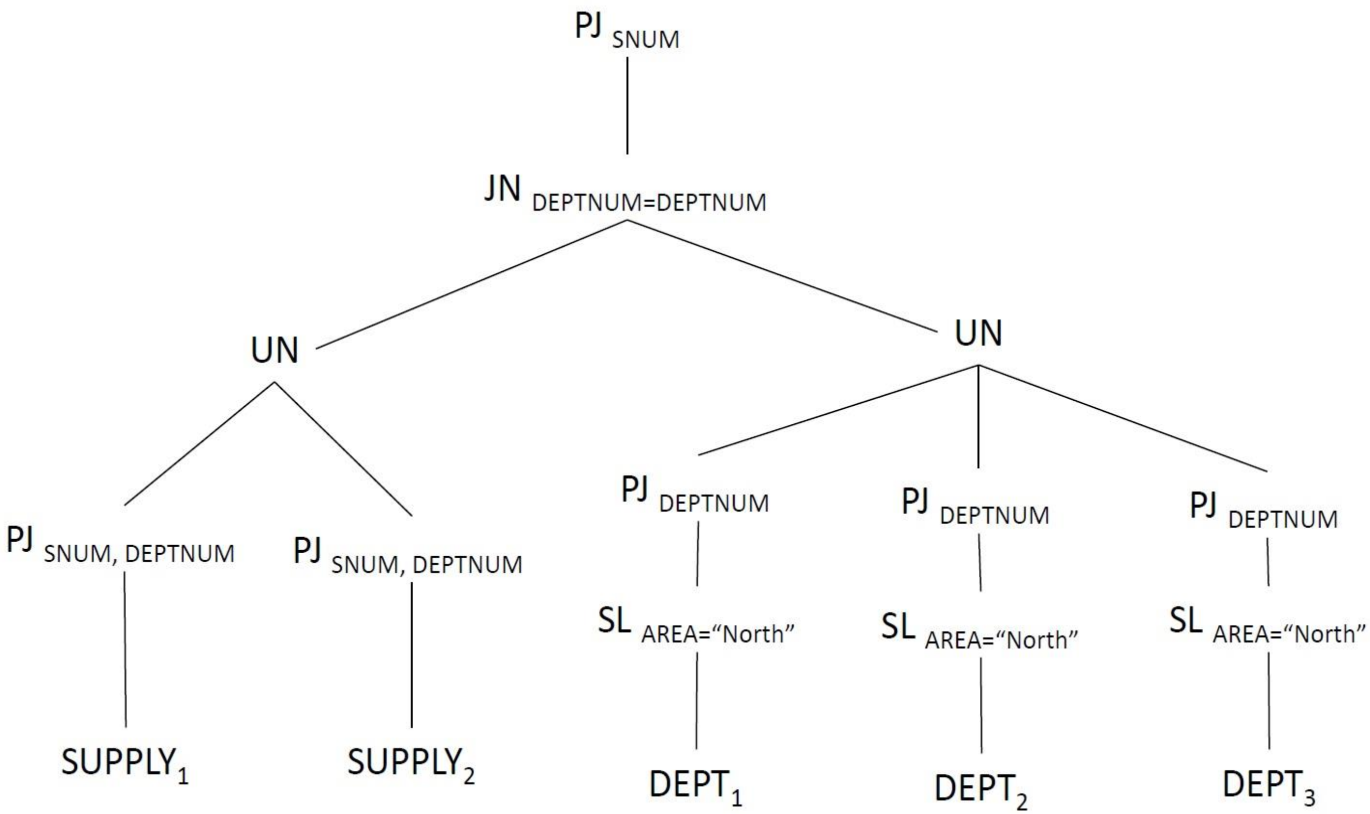
- Performing local processing on fragments.
  - –Fragment reducers.

- Then sending reduced fragments to the executing site (i.e. site - 2) in parallel.

Strategy – 2 (cont.)

- Transmitted amount:

  $X_{SUPPLY1} = ?$



Strategy – 2 (cont.)

- Transmitted amount:

  $X_{\text{SUPPLY1}} = \text{card}(\text{SUPPLY}_1) \times \{\text{size}(\textit{snum}) + \text{size}(\textit{deptnum})\} \times 8 \text{ bits}$

  $= 30000 \times (6+2) \times 8 \text{ bits}$

  $= 1920000 \text{ bits}$

Strategy – 2 (cont.)

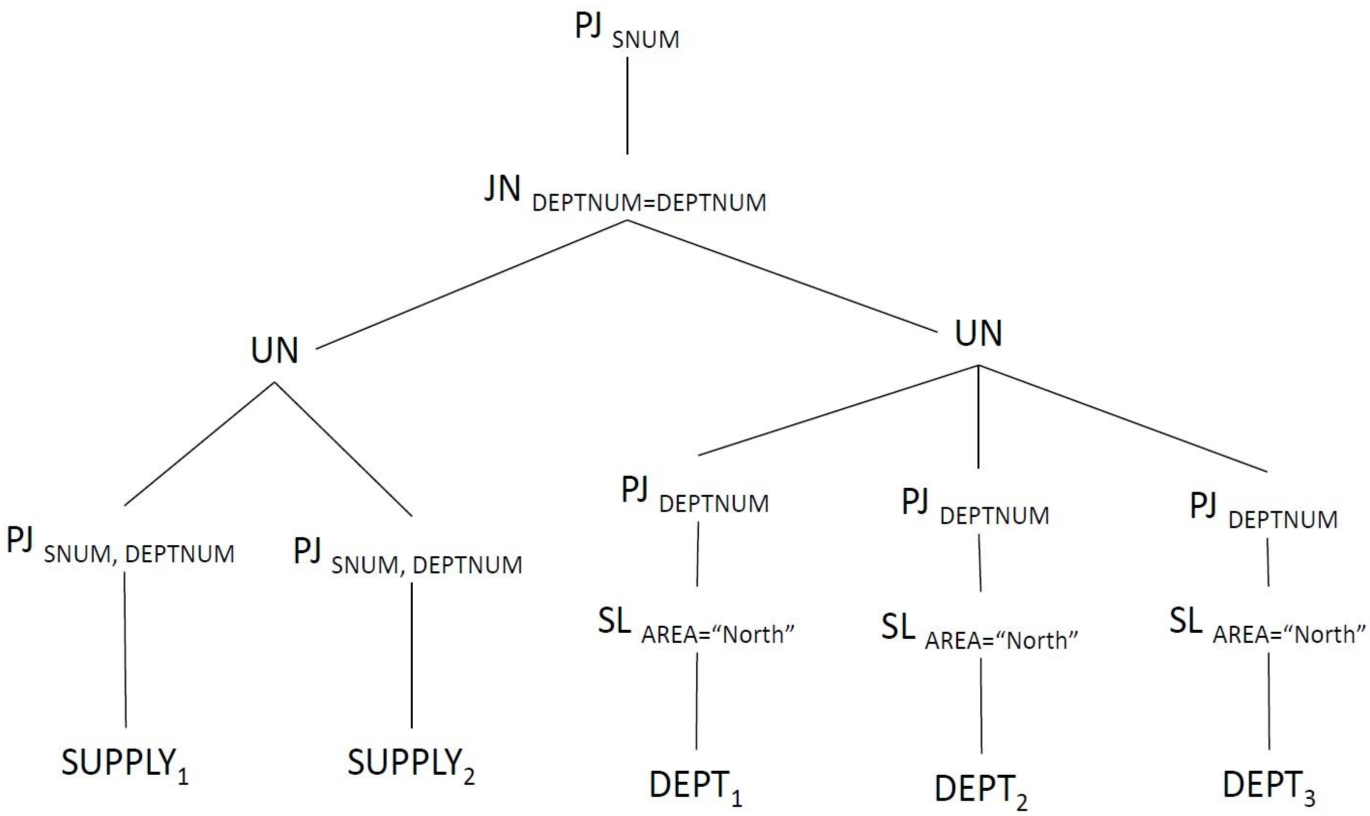
- Transmitted amount:

  $X_{SUPPLY2}$ = ?

  X_{DEPT1} = ?

  X_{DEPT2} = ?

  X_{DEPT3} = ?



Strategy – 2 (cont.)

- Transmitted amount for other fragments :

  $X_{SUPPLY2} = 20000 \times 8 \times 8 \text{ bits} = 1280000 \text{ bits}$

  $X_{DEPT1} = 0 \text{ bits}$

  $\boxed{X_{DEPT2} \approx 0 \text{ bits}}$ Try to investigate why

  $X_{DEPT3} = X_{DEPT2} \approx 0 \text{ bits}$

card ($DEPT_1$) = 10
card ($DEPT_2$) = card ($DEPT_3$) = 20

	deptnum	name	area	mgrnum
size	2	15	1	7
val	10	10	2	10

Strategy – 2 (cont.)

Assume $C_0 = 0$ and $D_0 = 0$

$TC(x) = C_0 + (\text{sum of all the amount}) \times C_1$

$\quad\quad = 3200000 \times C_1$

$TD(x) = D_0 + (\text{largest amount}) \times D_1$

$\quad\quad = 1920000 \times D_1$

If $D_1 = 10000$ bit/second,

Transmission Delay = $1920000 \times (1/10000)$ s = 192 s $\approx$ 3 mins

Comparisons

Strategy	Description	Time
1 (very bad)	▪ No simplification, no optimization. ▪ All fragments are brought to one site to execute the query.	10 m
2 (bad)	▪ Simplification applied (Criterion -1 and 2). ▪ No optimization. ▪ Processing on fragments are done on the site locally. ▪ Then, all fragments are brought to one site to execute the query.	3 m

Questions

What will happen for *strategy – 1* if the data collection from all the fragments are done sequentially, instead of in parallel?

Exercise

4. Consider the following global relational schemata.

EMP (ID, NAME, SAL, AGE, MGRNUM, DEPTNUM)
DEPT (ID, AREA, DEPTNUM, MGRNUM)

Corresponding fragmentation schemata:

$$EMP_1 = SL_{SAL \leq 25K}\, EMP$$
$$EMP_2 = SL_{SAL > 25K}\, EMP$$
$$DEPT_1 = SL_{AREA = \text{"North"}}\, DEPT$$
$$DEPT_2 = SL_{AREA = \text{"South"}}\, DEPT$$

Also consider the following global query.

$$\textbf{Q} = PJ_{NAME,\ AREA}(((SL_{SAL > 25K}\, EMP\ JN_{ID=ID}\, SL_{AREA = \text{"North"}}\, DEPT)\ DF\ (SL_{SAL \leq 25K} EMP\ JN_{ID=ID}\, SL_{AREA = \text{"North"}}\, DEPT))\ NJN\ (SL_{AREA = \text{"North"}}\, (EMP\ JN_{ID=ID}\, DEPT)))$$

Profiles of EMP1 and EMP2:

Card(EMP1) = 1800
Card(EMP2) = 1500

Site(EMP1) = 3
Site(EMP2) = 5

	ID	NAME	SAL	AGE	MGRN.	DEPTN.
Size	3	4	3	3	2	2
Val	1000	900	50	30	200	200

Profiles of DEPT1 and DEPT2:

Card(DEPT1) = 1300
Card(DEPT2) = 2000

Site(DEPT1) = 2
Site(DEPT2) = 4

	ID	AREA	MGRN.	DEPTN.
Size	2	4	2	2
Val	2000	30	200	200

Compare strategy 1 and 2 followed by assembling all the data at **site 3** and execution of the given query ***Q***.

Assume that the database profiles of EMP_1 , EMP_2 are similar. Also, profiles of $DEPT_1$ and $DEPT_2$ are identical. Determine the transmission delay for both the strategy when network-wide transfer rate is 20 bytes/second. Note that, the system does not provide any benefit of parallel processing and the initial delay is 0.