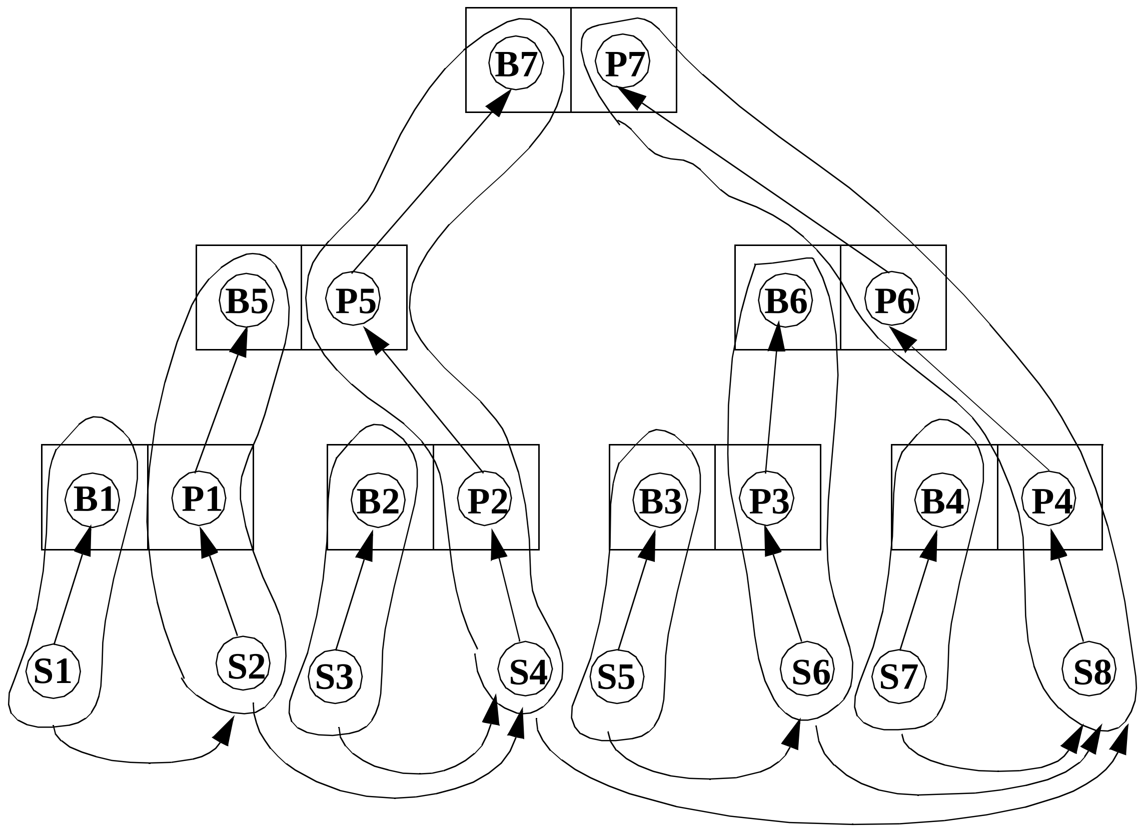


SEQ

PAR

PIPE Scan S1 - Build J1 ENDPIPE

PIPE Scan S3 - Build J2 ENDPIPE

PIPE Scan S5 - Build J3 ENDPIPE

PIPE Scan S7 - Build J4 ENDPIPE

ENPAR;

PAR

PIPE Scan S2 - Probe J1- Build J5 ENDPIPE

PIPE Scan S6 - Probe J3 - Build J6 ENDPIPE

ENDPAR;

PIPE Scan S4 - Probe J2 - Probe J5 -Build J7 ENDPIPE

PIPE Scan S8 - Probe J4 - Probe J6 - Probe J7 ENDPIPE

ENSEQ

Bushy Tree

4. Synthesis of Parallelization Methods and Choice an Approach

Table 1: Synthetesis of Optimization/Parallelization Methods

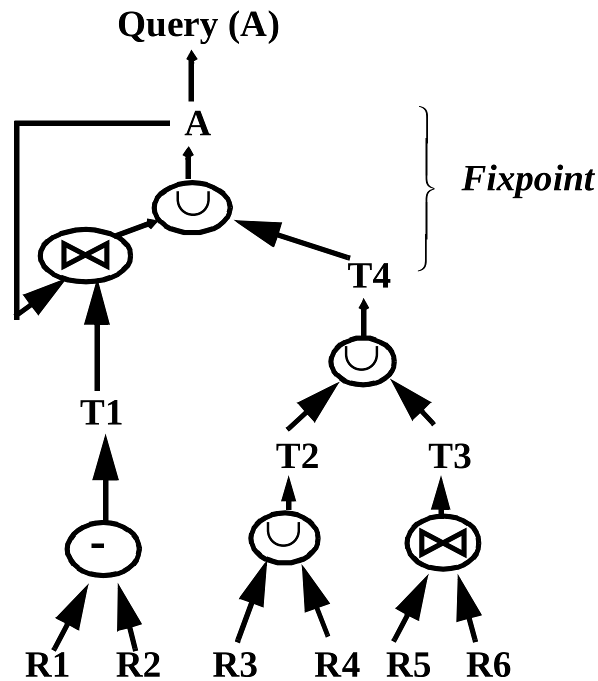
Context / *Approach*	*Static*	*Dynamic & Hybrid*
One-Phase	*Sch 90, Che 92* *Zia 93, Lan 93*	*Bru 97, Kos 97* *Rah 95*
Two-Phase	*Gar 96/97* *Has 94/95* *Ham95*	*Hon 92, NM 98* *Kab 98, Bon 98/99* *Ives 98*

Table 2: *Parameters to Choose an Approach ?*

Approach / *Parameters*	*Two- Phase*	*One-Phase*
Search Space	*Reduced*	*Large*
Search Strat.	*Extension*	*"Re-design"*
Cost Models	*F1 + F2*	*F*
Implementation	*+ "Easy"*	*- "Easy"*

III. Efficiency of Parallelism

- ***Shared-Nothing Architecture***
- ***Rel. Size [Bit 83], & Parameters [Sch 90], [Val 88]***



- **Simple Hash-Join Algorithm (Build + Probe)**
  Build (R) holds in memory

LRT $(T \leftarrow R \bowtie S) = T_{ef} + T_{d} + T_{com}$ where

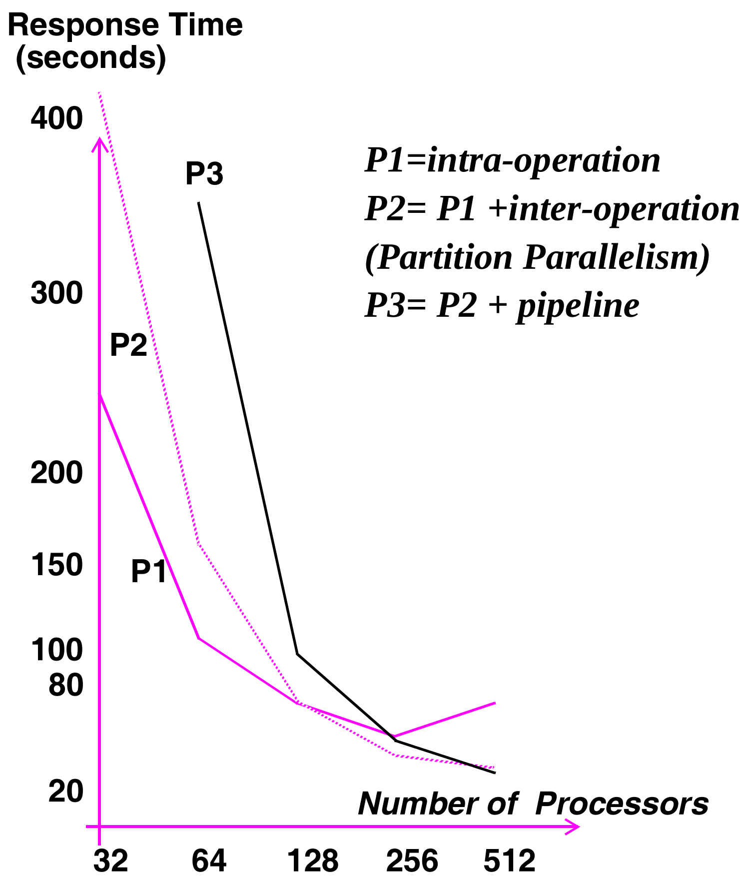
$T_{ef} = (|R|/d).th + ((||R||/d) + (||S||/d)).CR + (||R||/d/q).(||S||/d).CJO + |T|.I + ||T||/d.CW$

Build Time + Read Time + ComparisonTime +
Time for Moving a Tuple + WriteTime

$T_{d} = (|T|/d).th$ **and** $T_{com} = ((|T|/d).trf + p.msg).\lceil d/p \rceil$

|R| : **Number of Tuples in R=10^6**
||R|| : **Numbre of Pages in R**
th : **Time to hash a tuple (200 B)**
CR : **Time to read 1 page (18KB)= 8 ms**
CW : **Time to write 1 page= 16 ms**
CJO : **Time for joining 2 unsorted pages**
d : **Number of proc. of source operation**
p : **Number of proc. of destination operation**

Trf : Time to transfer a tuple
msg : Time to send a message
CPU = 4 MIPS



◆ *Efficiency of Parallelism :*

- *Intra-Operation with Lower NB of Processors*

- *Pipeline with Large NB of Processors*

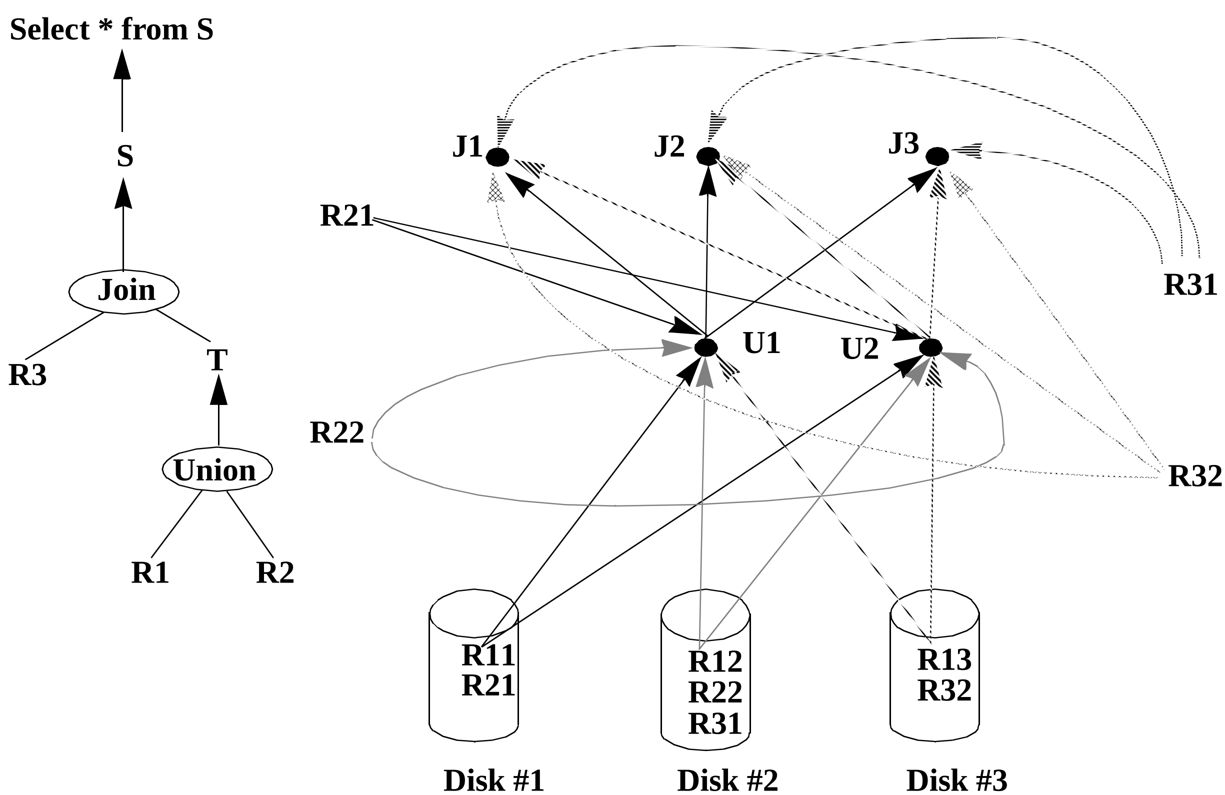
➥ *The Plague of Parallelism : Cost of Data Communication*

IV. Optimization of Data Communication

1. Logical Optimization : JSP -->PSJ (Reducing the Vol. of Data)

2. Physical Optimization : the order in which the joins are executed

3. Parallelization Phase :
Cost of Tuple Redistributing



A Simple SQL Query and Associated Data Flow Graph

Methods : Tree Coloring [HAS 95] Propagation Method[Ham 93]

- *Partitioning Attributes &*
  /Same Partitioning Function
- *Number of Processors*

Propagation Method : Partition Attribute & Number of Processors [Ham 93 in Dexa 93]