

SEQ
PAR
PIPE Scan S1 - Build J1 ENDDPIPE
PIPE Scan S3 - Build J2 ENDDPIPE
PIPE Scan S5 - Build J3 ENDDPIPE
PIPE Scan S7 - Build J4 ENDDPIPE
ENDDPAR;
PAR
PIPE Scan S2 - Probe J1- Build J5 ENDDPIPE
PIPE Scan S6 - Probe J3 - Build J6 ENDDPIPE
ENDDPAR;
PIPE Scan S4 - Probe J2 - Probe J5 -Build J7 ENDDPIPE
PIPE Scan S8 - Probe J4 - Probe J6 - Probe J7 ENDDPIPE
ENSEQ

Bushy Tree

4. Synthesis of Parallelization Methods and Choice an Approach

Table 1: Synthetesis of Optimization/Parallelization Methods

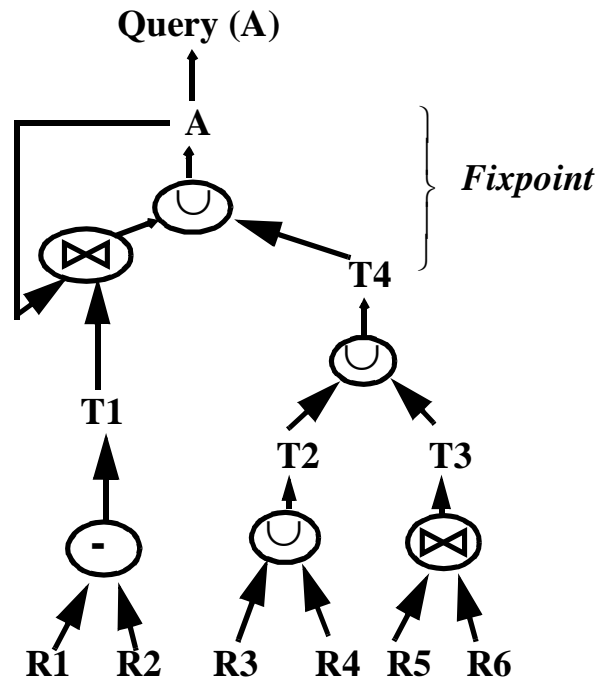
<i>Context</i>	<i>Static</i>	<i>Dynamic & Hybrid</i>
<i>Approach</i>		
<i>One-Phase</i>	<i>Sch 90, Che 92 Zia 93, Lan 93</i>	<i>Bru 97, Kos 97 Rah 95</i>
<i>Two-Phase</i>	<i>Gar 96/97 Has 94/95 Ham95</i>	<i>Hon 92, NM 98 Kab 98, Bon 98/99 Ives 98</i>

Table 2: Parameters to Choose an Approach ?

<i>Approach</i>	<i>Two- Phase</i>	<i>One-Phase</i>
<i>Parameters</i>		
<i>Search Space</i>	<i>Reduced</i>	<i>Large</i>
<i>Search Strat.</i>	<i>Extension</i>	<i>“Re-design”</i>
<i>Cost Models</i>	<i>F1 + F2</i>	<i>F</i>
<i>Implementation</i>	<i>+ “Easy”</i>	<i>- “Easy”</i>

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

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

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

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

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

|R| : Number of Tuples in R=10⁶

||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

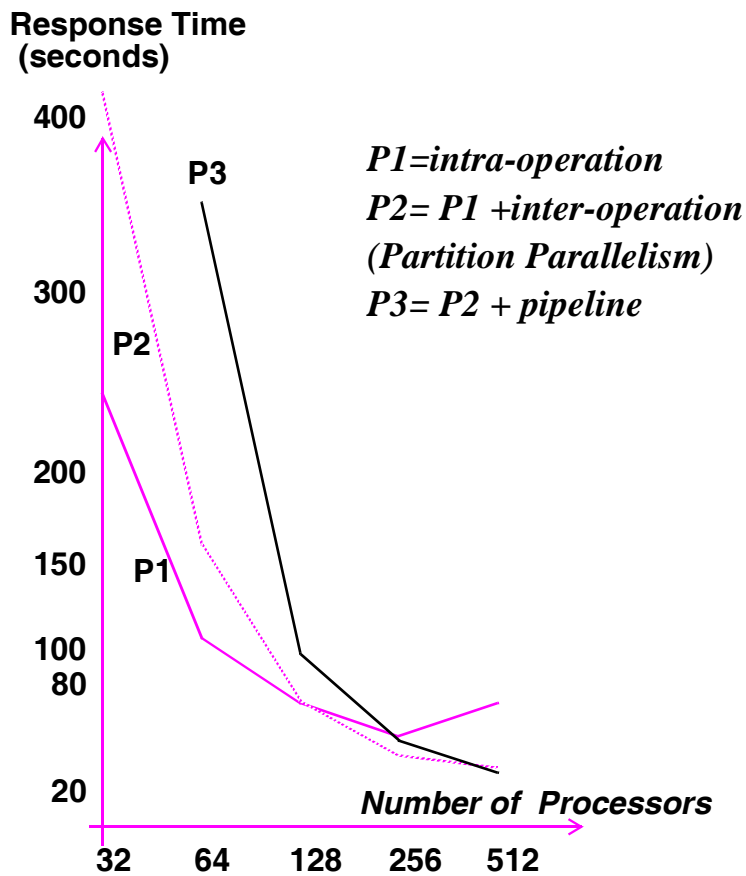
d : Number of proc. of source operation

p : Number of proc. of destination operation

Trf : Time to transfer a tuple

msg : Time to send a message

CPU = 4 MIPS



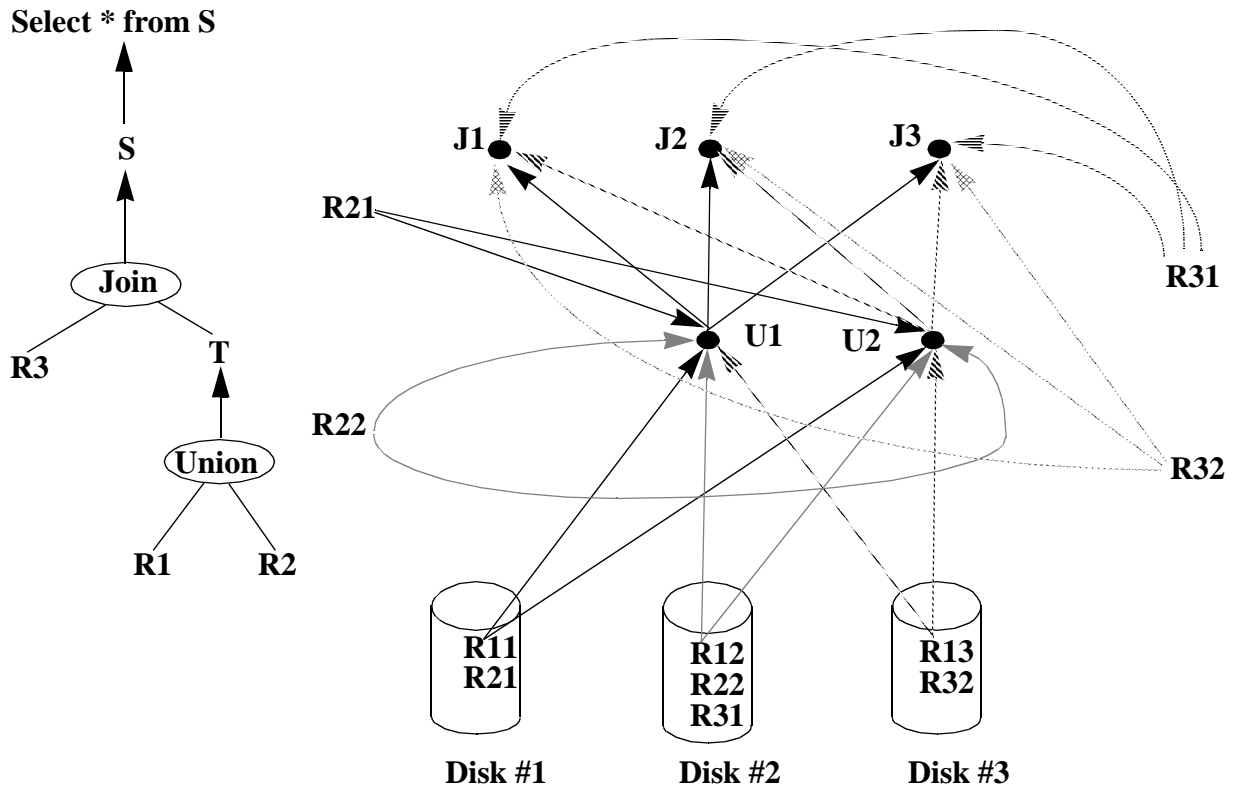
◆ *Efficiency of Parallelism :*

- *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]