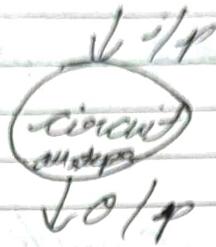


⇒ Pipeline :-



(Problem)

Data + Data manipulation

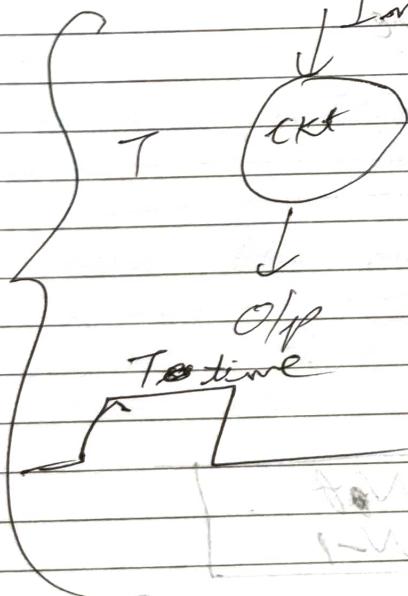
DS + Algorithms → steps of manipulation

Code language

Hardware

pipeline

Instructions



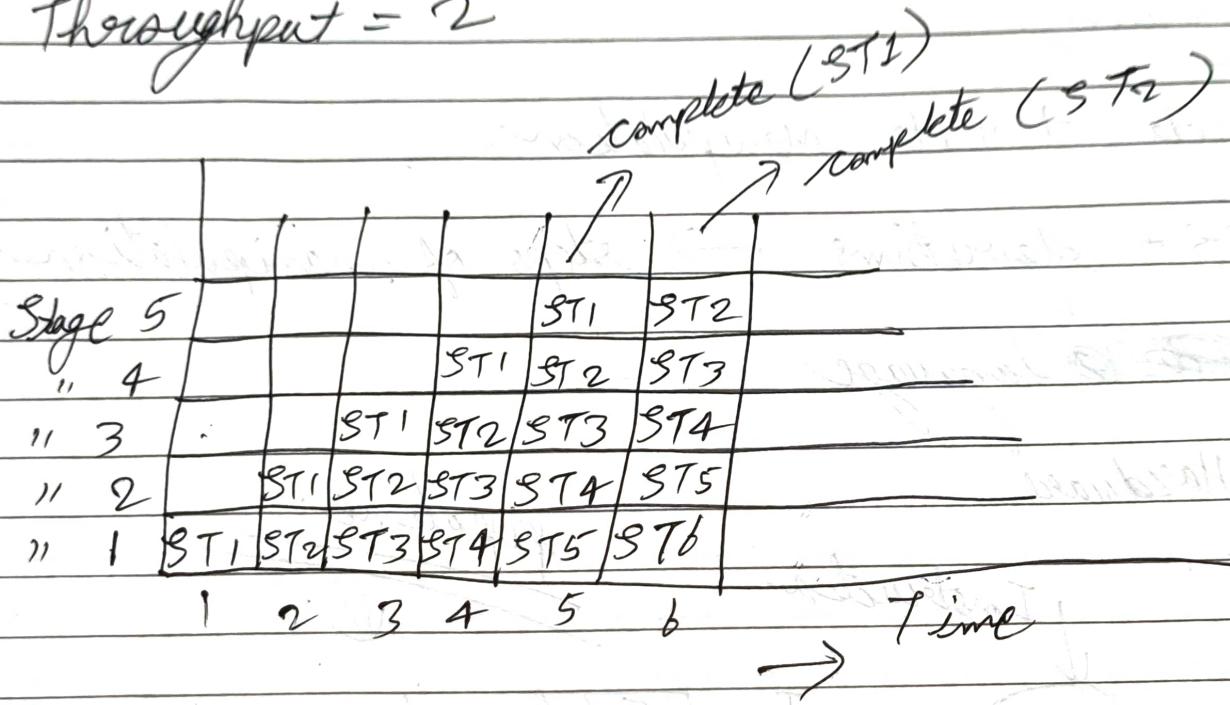
⇒ Pipeline

No. of segments / stages = M

" " tasks = N

Total time taken to complete tasks = ?

Throughput = ?



$$\text{Total time} = \underline{\underline{M+N-1}} \quad \text{Ans}$$

$$\text{Throughput} = \frac{\text{No. of tasks}}{\text{Total time}}$$

$$\boxed{\text{Throughput} = \frac{N}{M+N-1} = \frac{\text{No. of tasks}}{M+N-1}}$$

(if :- freq. of pipeline).

$$\text{Max. Throughput} = \lim_{N \rightarrow \infty} \frac{Nt}{M+N-1}$$

$$\text{Max. Throughput} = f \rightarrow \text{frequency}$$

$$\star \text{Speed up} = \frac{\text{Time without pipeline}}{\text{Time with pipeline}}$$

$$\text{Speed up} = \frac{MN \text{ clock cycles}}{M+N-1}$$

$$\text{Max. speedup} = \lim_{N \rightarrow \infty} \frac{MN}{M+N-1}$$

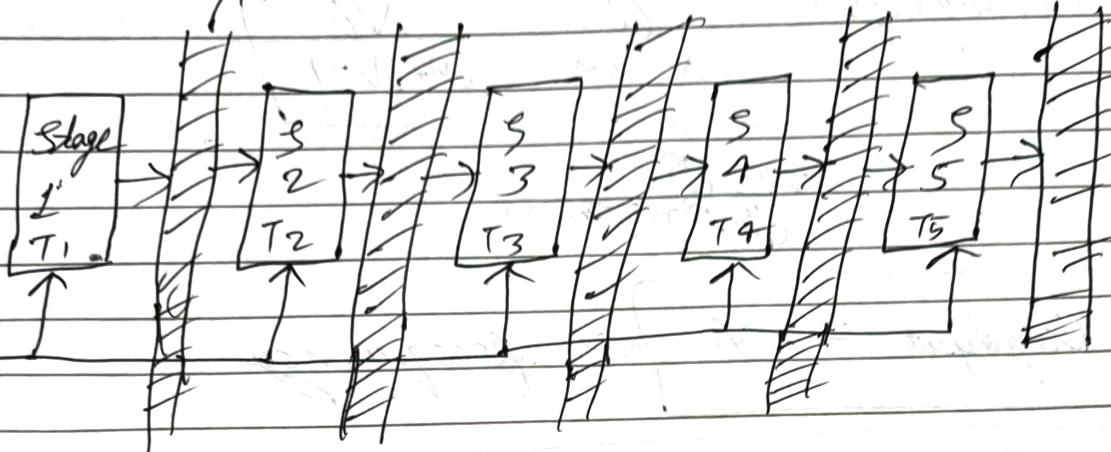
$$\text{Max. Speedup} = M \rightarrow \text{No. of segment}$$

Max. the no. of segments \rightarrow max. the freq.

because time taken to complete one stage decreases.



batch



$$f = \frac{1}{\text{batch time} + \max(T_1, T_2, \dots, T_5)}$$

time to put & fetch data from latch (memory)

⇒ PCR = Performance Rate Cost Rates

$\text{PCR} = \frac{\text{Performance}}{\text{Cost}}$

more info on memory is on next page

notes on disposal of waste will be added

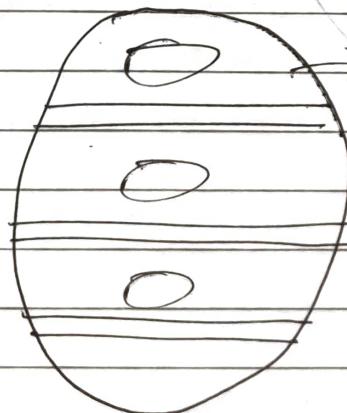
c = cost of digital gates

h = latch cost

k = no. of stages

$$PCR = \frac{f}{c + kh} = \frac{1}{(c + kh)(t/k + d)}$$

t = time taken by H/w without pipeline.



d = latch time

$$\text{Time Period} = \frac{t + d}{k} = t/k + d/k$$

$$\text{freq} = \frac{1}{d + (t/k)} = \frac{1}{(t/k - t) + d/k}$$

$$PCR = \frac{1}{(c + kh)(t/k + d)}$$

$$PCR = \frac{1}{(c+kh)(t/k+d)}$$

$$\frac{d(PCR)}{dk} = 0$$

$$(c+kh) \frac{(-1)}{(c+kh)^2} = 0$$

$$k_0 = \sqrt{\frac{tc}{dh}}$$

$$k_0^2 = \frac{tc}{dh}$$

$$\frac{-1}{(c+kh)^2(t/k+d)^2} \times \left[(c+kh)\left(\frac{-1}{k^2}\right) + (t/k+d)(h) \right] = 0$$

$$\left(\frac{t+d}{k}\right)h = \frac{(c+kh)t}{k^2}$$

$$(t+dk)h = \frac{(c+kh)t}{k}$$

$$k(t+dk)h = (c+kh)t$$

~~$$(c+kh)$$~~

$$ct + dk^2 h = ct + kh t$$

$$dk^2 h + (t - ht)k = ct = 0$$

$$k = \frac{-(t - ht) \pm \sqrt{(t - ht)^2 + 4dhct}}{2dh}$$



$$\Rightarrow \frac{-1}{(c+kh)^2(t/k+d)^2} \left[(c+kh)\left(-\frac{t}{k^2}\right) + (t/k+d)(h) \right] = 0$$

$$\Rightarrow \left(\frac{t}{k}+d\right)h = \frac{t}{k^2}(c+kh)$$

~~$$\frac{th}{k} + dh = \frac{tc}{k^2} + \frac{th}{k^2}$$~~

~~$$K^2 = \frac{tc}{dh}$$~~

$$K = \sqrt{\frac{tc}{dh}}$$

$$\frac{d^2(P(R))}{dk^2} = \frac{d^2}{dk^2} \left[\left(\frac{t}{k}+d\right)h - \frac{t}{k^2}(c+kh) \right]$$

$$= \left(-\frac{t}{k^2}\right)h - \left[(c+kh)\left(-\frac{2t}{k^3}\right) + \frac{t}{k^2}(h)\right]$$

~~(-ve)~~ ~~(-ve)~~ = ~~-ve~~
 ↓
 maximum

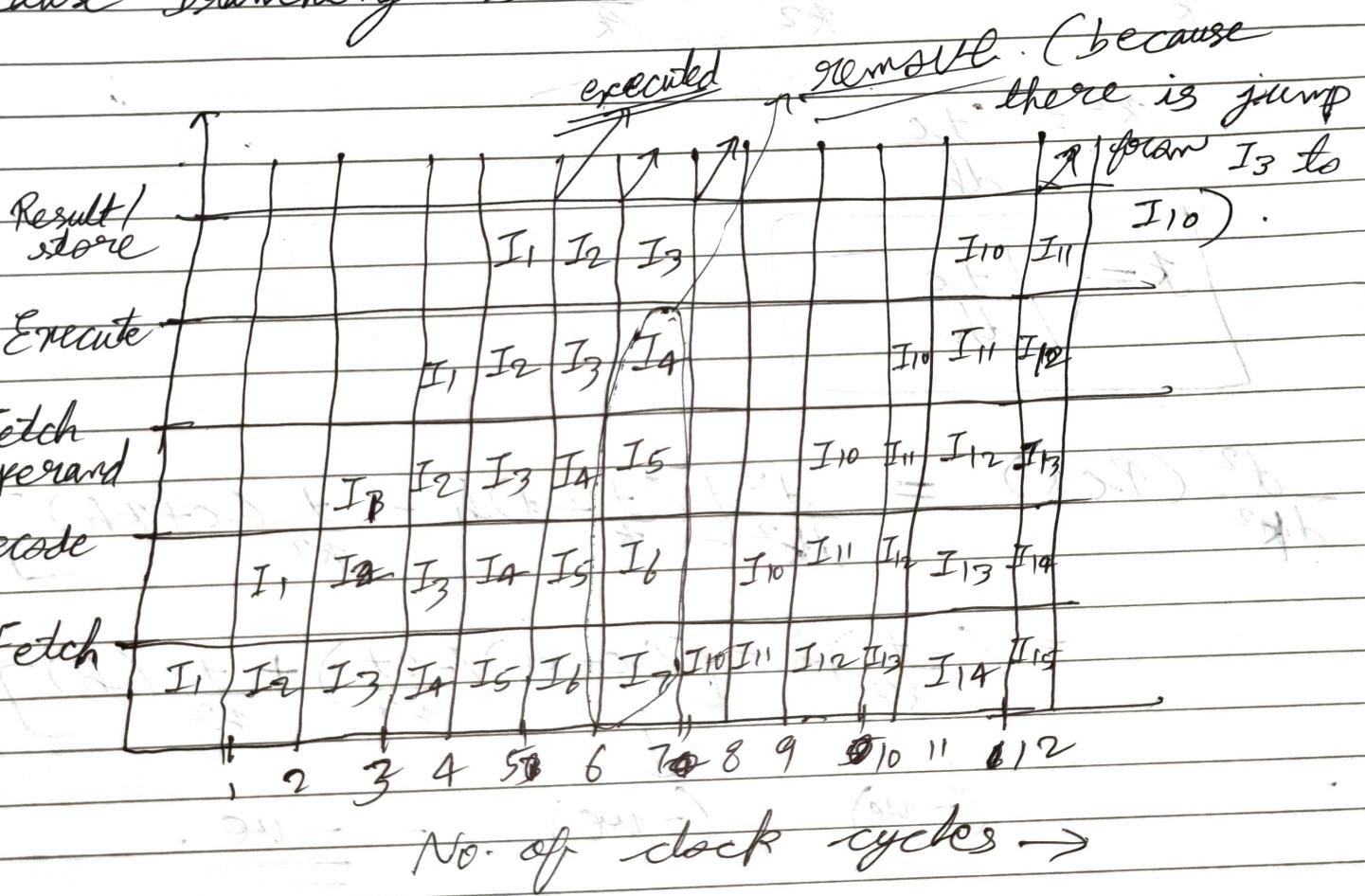
A. S.

⇒ Instruction Pipeline :-

No. of instructions = N

" " segment = M

Probability that a given instruction will cause branching = s





No. of extra clock due to 1 branch construction = $(M-1)$

Avg. no. of branching = $\sum_{j=1}^N 1 \cdot s + 0 \cdot (1-s)$

$$\text{Time} = M+N-1 + NS(M-1)$$

clock cycles

$$\text{Speed up} = \frac{MN}{M+N-1 + NS(M-1)}$$

$$\text{Throughput} = \frac{Nf}{M+N-1 + NS(M-1)}$$

$$\text{Max. speed up} = \lim_{N \rightarrow \infty} \frac{MN}{M+N-1 + NS(M-1)}$$

$$\text{Max. speed up} = \frac{M}{1 + s(M-1)}$$

$$\text{Max. Throughput} = \frac{f}{1 + s(M-1)}$$



Q:- $M = 5$, $f = 10 \text{ MHz}$

S	max speed up	Throughput
0	5	10×10^6
0.1	3.5	7.14×10^6
0.2	2.8	5.55×10^6
0.3	2.27	4.54×10^6

2) 1 Instruction Cycle (IC) = M clock cycles.

$$\text{Avg. no. of Instructions executed per IC} = \frac{N}{M+N-1+Ns(M-1)}$$

$$= \frac{NM}{M+N-1+Ns(M-1)}$$

⇒ N = No. of instructions.

Probability that a given inst^m is unconditional Branch inst^m = P

" " " " " " " conditional " " " = Q

" " a conditional branch will cause branching = R

No. of segment of pipeline = M

Frequency of pipeline = f

i) Total time taken to execute N instruction = ?

ii) Throughput = ?

iii) Speed up = ?

iv) Avg. No. of inst^m executed per Ic = ?

Avg. No. of branching = $NP + NQR$

Total time = $M + N - 1 + (NP + NQR)(M - 1)$

clock cycle

Throughput = $\frac{N \cdot f}{M + N - 1 + N(P + QR)(M - 1)}$

Max. throughput = $\frac{f}{M + (P + QR)(M - 1)}$



$$\text{Speedup} = \frac{NM}{(M+N-1) + N(P+QR)(M-1)}$$

$$\text{Avg. No. of inst}^m \text{ exec. per I.C} = \frac{NM}{(M+N-1) + N(P+QR)(M-1)}$$

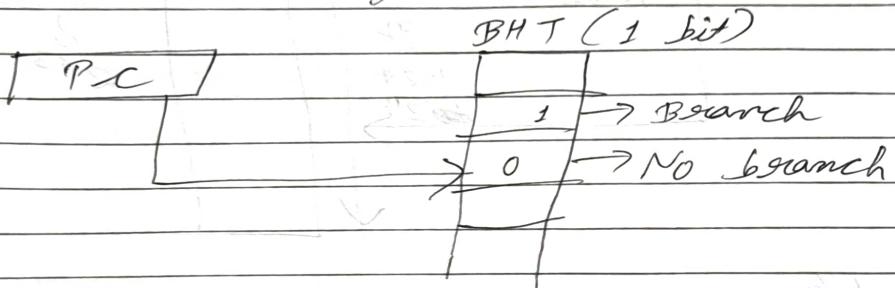
$$\text{Max. speed up} = \frac{M}{1 + (P+QR)(M-1)}$$

⇒ Branch



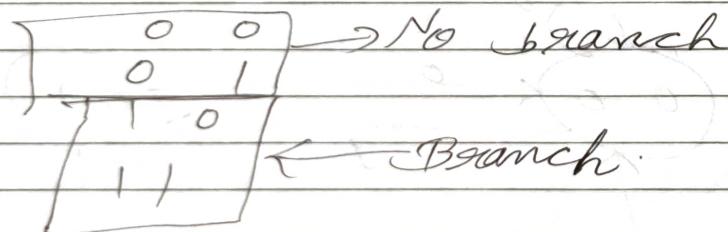
Branch Prediction :-

• Branch History Table :- (BHT) :-



1-bit Branch predictor

• 2-bit Branch :-

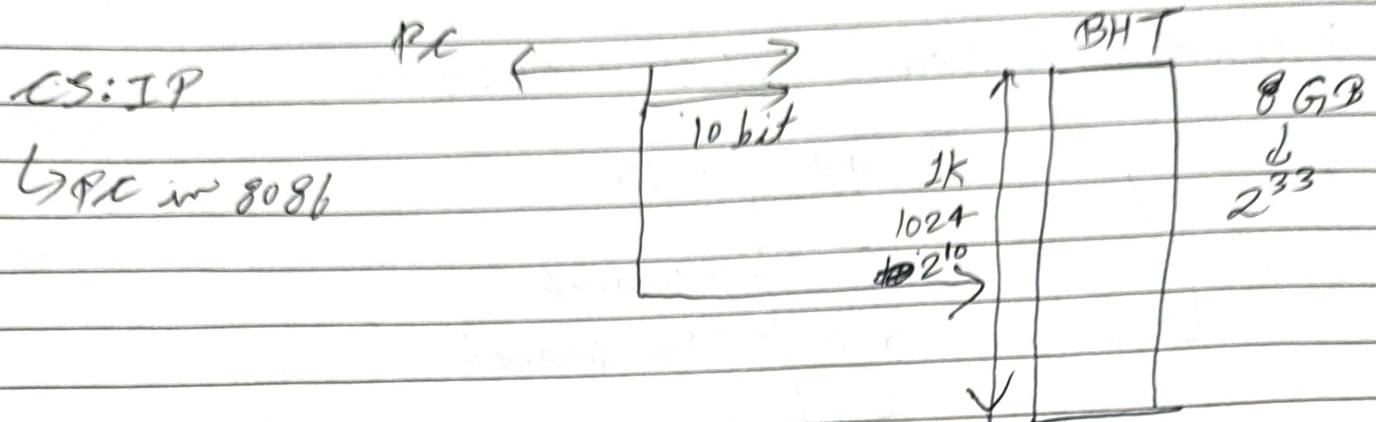


m-bit :-

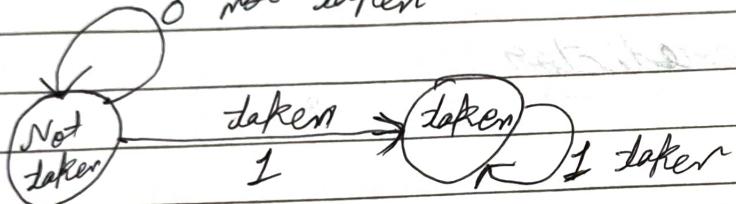
$0 - 2^{m-1} \rightarrow \text{No branching}$

$2^{m-1} - 2^m - 1 \rightarrow \text{Branching}$

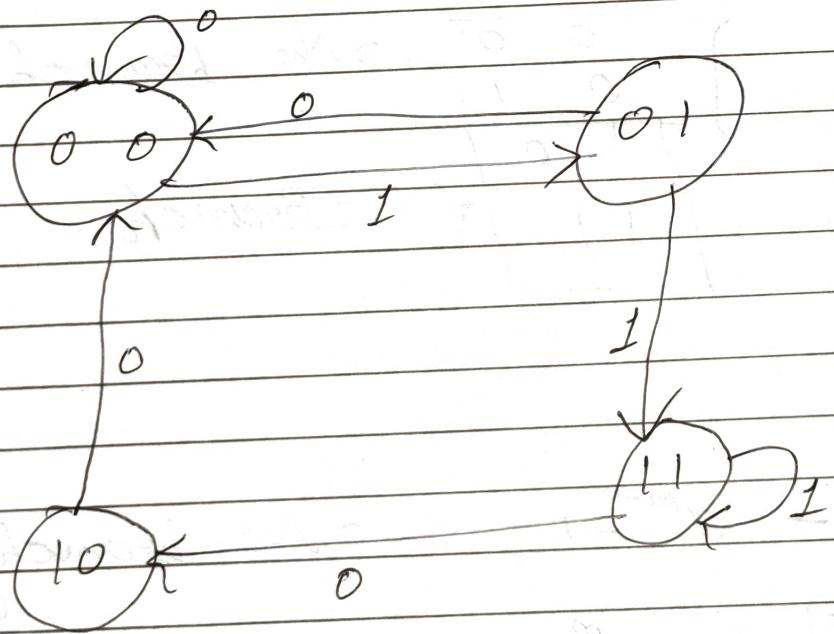
Address Instruction (8K-32K)



1-bit :- 0 not taken



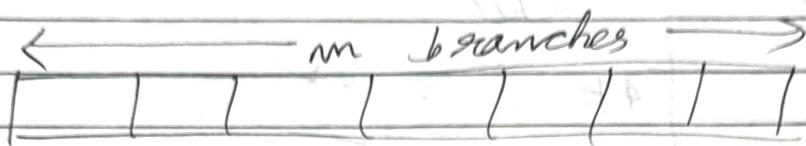
2-bit :-



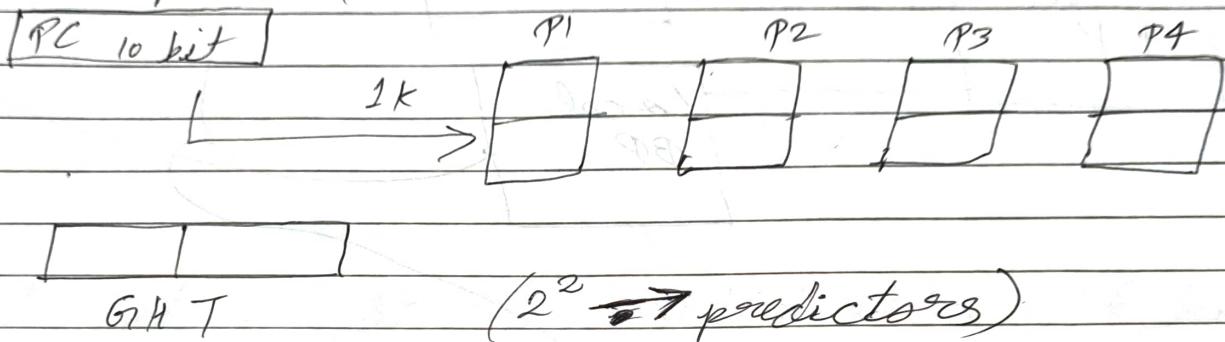
→ correlating branch predictors



→ (m, m) branch predictor :-



$(2, 2)$ branch predictor



Q1:- How many bits are in $(0, 2)$ branch predictors with 4K entries.

Q2:- How many entries are in $(2, 2)$ branch predictors with same no. of bits.

Ans 1:- $2^0 \Rightarrow 1$ predictor

$$4K = 2^3 \times 2^0 = 8K$$

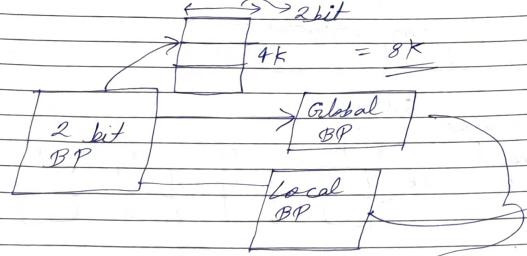
12 bits.

$$\text{bit in 1 predictor} = 2$$
$$2 \times 4K = 8K \text{ bits}$$

$$\text{No. of entries} \times 2^2 \times 2 = 8K$$

$$\text{No. of entries} = \frac{8K}{2^3} = 1K$$

⇒ Tournament Branch Predictor :-



Date 12/09/22
Page No. _____

Date _____
Page No. _____

$$\text{Total Bits} = 8k + 8k + 10k + 3k = 29k \text{ bits}$$

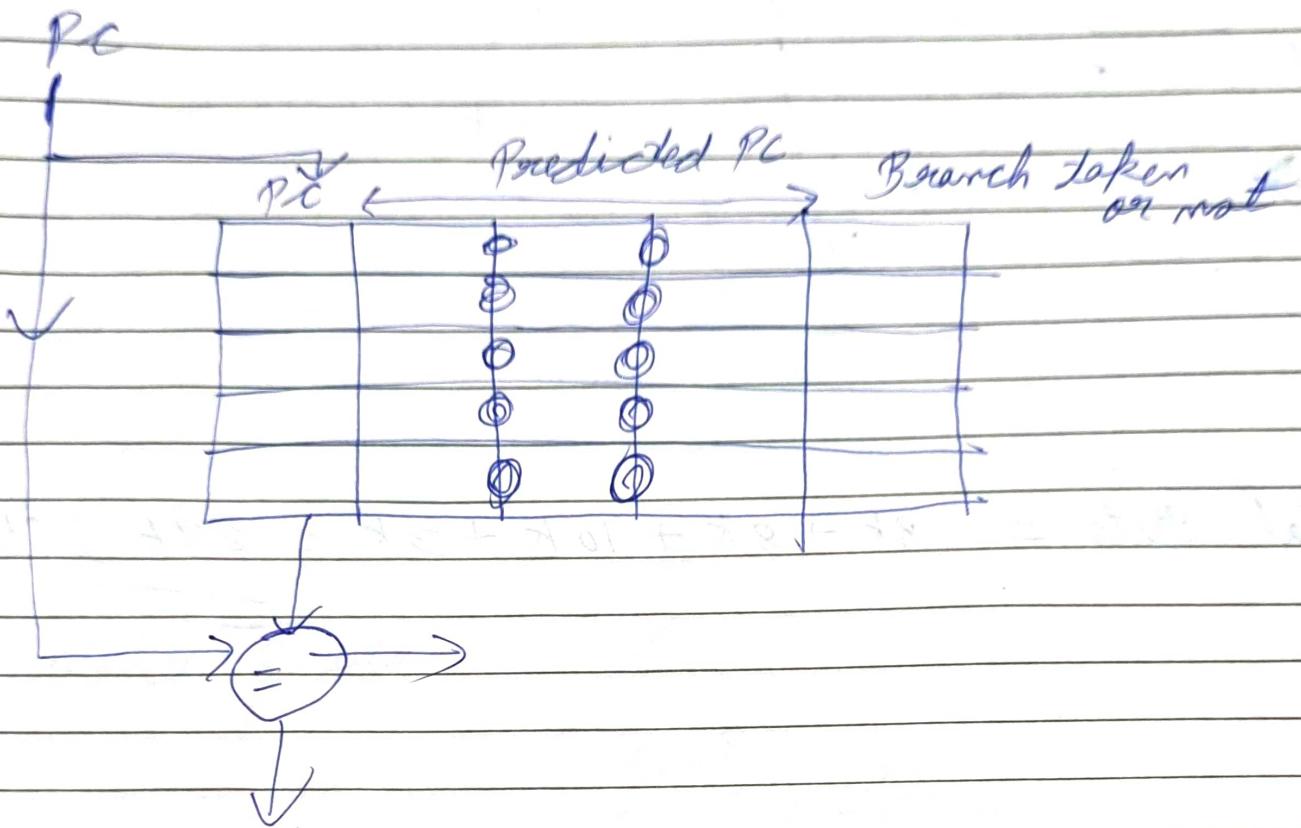
* Alpha 21264



12 Branch



⇒ BTB (Branch Target Buffer) :-



Ques:- Determine the total branch penalty for a branch target buffer assuming the penalty cycle for individual mispredictions from the following table.

0.18
0.2
0.38



~~Ans:- 800~~

Instruction in Buffer	Prediction	Actual prediction	Penalty cycle
-----------------------------	------------	----------------------	------------------

Yes	Taken	Taken	0
-----	-------	-------	---

Yes	Taken	Not Taken	2
-----	-------	-----------	---

No		Taken	2
----	--	-------	---

No		Not Taken	0
----	--	-----------	---

+ 2 = insertion with delay

Prediction Accuracy = 90% (for instruction in buffer).

Hit rate = 90% (for branches predicted taken in buffer)

Ans:- Branch Penalty = (BTB hit but has wrong prediction) $\times 2$

+ (Miss in BTB but actually taken) $\times 2$

$$= 0.90 \times 0.1 \times 2 + 0.1 \times 2$$

$$= 0.38$$



⇒ Collision force scheduling for y :-

	1	2	3	4	5	6	
S_1	X				X		4
S_2			X	∅			
S_3	X	∅	X	∅	X		2,4

⇒ Forbidden latencies, Permissible latency ~~2,4~~

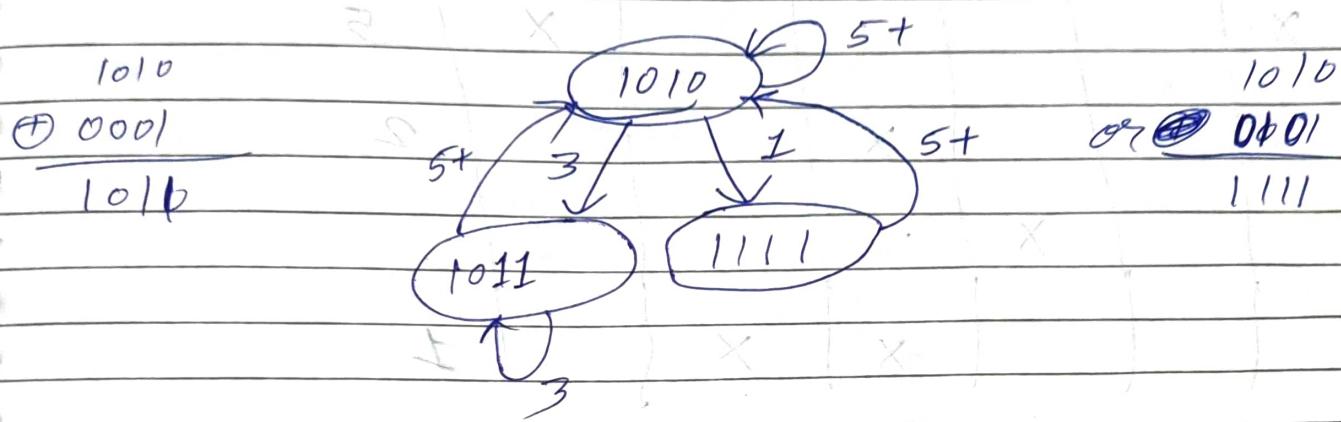
Forbidden latencies = 2, 4

Permissible latencies = 1, 3, 5, 6+

⇒ Initial collision vector :-

1010

⇒ State Transition Diagram :-



(a) ω_{eff} is inverted with respect to (a)

2) Simple Cycles :-

$$(5), (3,5), (3), (1,5)$$

\Rightarrow Greedy Cycle:- (cycle containing minimal outgoing edges from each state).
 $(3), (1, 5)$.

\Rightarrow McAL :- 3



2)	1	2	3	4	5	6
S1	X				X	5
S2		X		X		2
S3		X				
S4			X	X		1

⇒ Forbidden latencies = (1, 2, 5)

⇒ Permissible latencies = (3, 4, 6+)

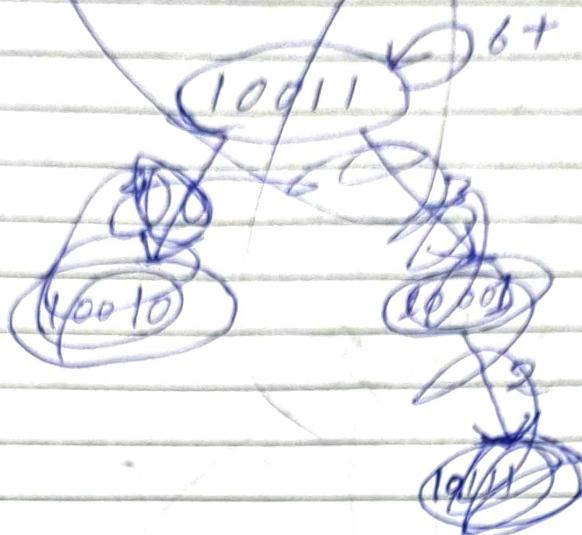
⇒ Initial collision vector:-

10011



⇒ State Transition Diagram :-

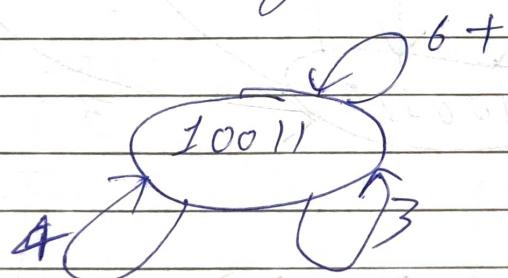
10011
 $\oplus 00001$
 10010



$$\begin{array}{r} 10011 \\ + 00010 \\ \hline 10001 \end{array}$$

$$\begin{array}{r} 10011 \\ + 00100 \\ \hline 10011 \\ + 00010 \\ \hline 10011 \end{array}$$

⇒ State Transition Diagram :-



$$\begin{array}{r} 10011 \\ + 00010 \\ \hline 10011 \end{array}$$

⇒ Simple Cycle :- (3), (4), (6)

⇒ Greedy Cycle :- (3)

⇒ MAFL :- (3) ~~Ans~~



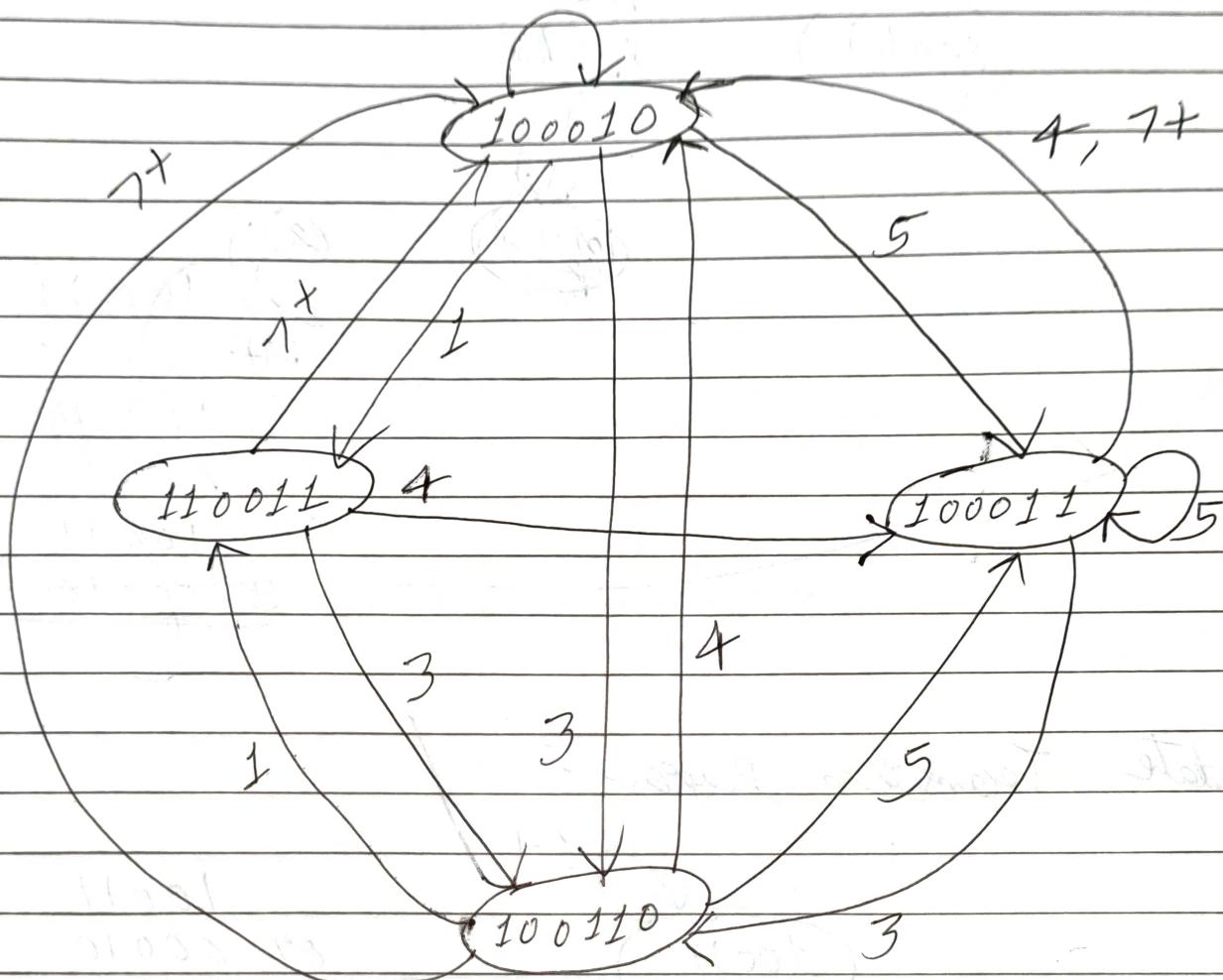
100010

000110

100110

$$2 \leq M AL \leq 4$$

4, 7+



⇒ Simple cycles = (4), (7), (4, 5), (5), (5, 7),
 (3, 4), (3, 5), (1, 3), (1, 7),
 (1, 3, 7); (1, 3, 5, 4)

⇒ Greedy cycle = (1, 3)

⇒ $M AL = \frac{1+3}{2} = 2$

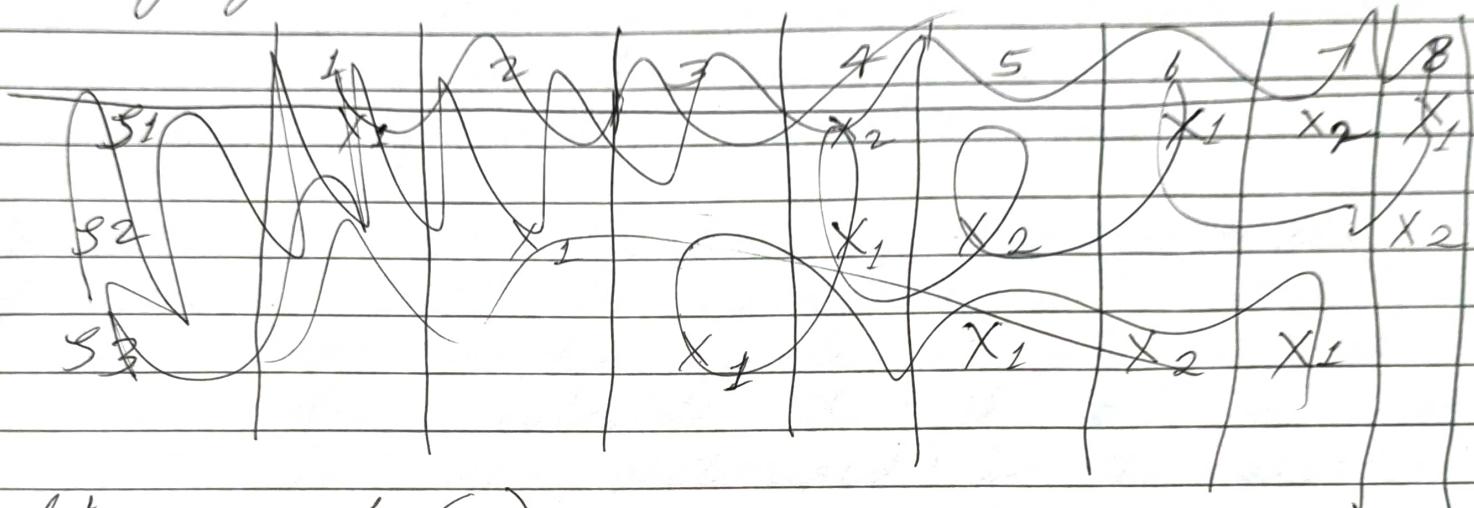


⇒ Latency cycle (3)

$$\Rightarrow \text{Throughput} = \frac{1}{M.d.L}$$

$$= \frac{1}{2} = 0.5 \text{ units}$$

⇒ Latency cycle = 3



Latency cycle (3)

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21
S1	1	2	1	3	1	2	4	2	3	5	3	4	6	4	5	5	6				
S2	1	1	2	2	3	3	4	4	5	4	5	5	6	6							
S3	1	1	2	1	2	3	2	3	4	3	4	5	4	5	6	5	6	6			

$$\text{efficiency} = \frac{0.88}{9} \times 100 = 88.9\%.$$

Af

(1,8) cycle

Date _____

Page No. _____



⇒ Performance laws:-

Q:- A uni-processor computer can operate in either scalar or vector mode. In vector mode computation can be done 9 times faster than scalar mode. A certain benchmark program took time T to run on this computer. Further, it was found that 25% of T was attributed to vector mode. For the remaining time machine operated in scalar mode.

i) Calculate the effective speedup under the above condition as compared with condition when the vector mode is not used at all. Also calculate α (the % of code that has been vectorised) in the above program.

$$\text{Ans} \quad \text{speedup} = \frac{\text{time without enhancement}}{\text{time with enhancement}}$$

$$= \frac{0.75 T + 0.25 T \times 9}{T}$$

$$= 0.75 + 2.25$$

$$\boxed{\text{speedup} = 3.00}$$

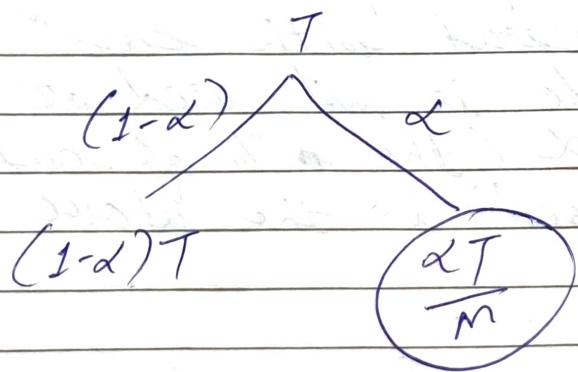


M

$$\Rightarrow \alpha = \frac{0.25 T \times 9^3}{3 T} \\ = 0.75$$

$$\boxed{\alpha = 75\%} \quad \boxed{\text{Ans}}$$

\Rightarrow Series | Parallel



$$\boxed{\text{Time enhancement} = (1-d)T + \frac{dT}{m}}$$

$$\Rightarrow \text{speedup} = \frac{T_{without \text{ enhancement}}}{T_{with \text{ enhancement}}}$$

$$\boxed{\text{speedup} = \frac{T}{(1-d)T + \frac{dT}{m}} = \frac{1}{(1-d) + \frac{d}{m}}}$$



$$\text{Q) } 3 = \frac{1}{1-\alpha + \frac{\alpha}{9}}$$

$$3(1-\alpha + \frac{\alpha}{9}) = 1$$

$$3 - 3\alpha + \frac{3\alpha}{9} = 1$$

$$\alpha = \frac{48\alpha}{3}$$

$$\boxed{\alpha = \frac{3}{4}}$$

enhancement 1

enhancement 2

enhancement m

$$\text{overall speedup} = \frac{1}{d_1 + d_2 + \dots + d_m + \frac{d_1}{S_1} + \frac{d_2}{S_2} + \dots + \frac{d_m}{S_m}}$$

↳ Generalized Amdahl's law.



~~reduced execution~~

⇒ Fraction of [^]time _{when} enhancement is used =

$$\frac{1 - (\alpha_1 + \alpha_2 + \dots + \alpha_m)}{1 - (\alpha_1 + \alpha_2 + \dots + \alpha_m) + \frac{\alpha_1}{S_1} + \frac{\alpha_2}{S_2} + \dots + \frac{\alpha_m}{S_m}}$$

ii) Suppose we double the speed ratio b/w the vector mode and the scalar mode by hardware improvement. Cal. the effective speedup that can be achieved.

iii) Suppose the same speedup obtained in part (ii) must be obtained by compiler improvement instead of hardware improvement. What could be new vectorization ratio α that should be supported by the vectorizing compiler to achieve the same enhancement.

~~Ans~~ \rightarrow speedup = $\frac{1}{1 - \alpha + \frac{\alpha}{S}}$

$$= \frac{1}{1 - 0.75 + \frac{0.75}{18}}$$

$$\boxed{\text{speedup} = 3.43}$$



$$\text{Quesiii) } 3.43 = \frac{1}{1-\alpha + \frac{\alpha}{9}}$$

$$1-\alpha + \frac{1}{9} = \frac{1}{3.43}$$

$$1 - \frac{8\alpha}{9} = \frac{1}{3.43}$$

$$1 - \frac{1}{3.43} = \frac{8\alpha}{9}$$

$$1 - 0.2915 = \frac{8\alpha}{9}$$

$$0.7084 = \frac{8\alpha}{9}$$

$$\boxed{\alpha = 0.79}$$

Ques:- let α be the fraction of program code that can be executed simultaneously by m processors in a computer system.

Assume that remaining code must be executed sequentially by a single processor.

Each processor has an execution rate of n MIPS and all processors are assumed to be equally capable.
Derive an expression for effective MIPS rate.

Ans:- Let no. of instructions to be executed
= K millions.

$m n \rightarrow$ Parallel Mode

$$\alpha K$$

sequential mode

$$(1-\alpha)K$$

$$T = \frac{\alpha K}{mn} + \frac{(1-\alpha)K}{n}$$

$$\boxed{T = \frac{K}{mn} [\alpha + (1-\alpha)m]}$$

$$\Rightarrow \text{Effective MIPS} = \frac{K}{K[\alpha + (1-\alpha)m]} \cdot mn$$

$$\boxed{\text{effective MIPS} = \frac{mn}{\alpha + (1-\alpha)m}}$$



Eg. 3.4

Ques 1. Consider a corp which can execute a program in 2 ~~operational modes~~ mode regular mode and enhanced mode with probability distributions $\{\alpha, 1-\alpha\}$ respectively If α varies b/w $a \& b$.

$$[0 \leq a < b \leq 1]$$

Derive an expression for average speedup factor using harmonic mean concept.

$\Theta \Rightarrow$ Program:-

$= =$

Unbalanced

PAR

parallel

$\text{for } (l=1; l <= 32; l++)$

$\{ \text{for } (i=(l-1) \times 32 + 1; i <= l \times 32; i++)$

$\text{sum}[i] = 0;$

$\text{for } (j=1; j <= i; j++) \{$

$\text{sum}[i] = \text{sum}[i] + j$

$\}$

$\}$

Balanced:-

These 2
will run
parallel.

PAR $\text{for } (l=1; l <= 32; l++) \{$

$\{ \text{for } (i=(l-1) \times 16 + 1; i <= l \times 16; i++) \{$

$\text{sum}[i] = 0;$

$\text{for } (j=1; j <= i; j++) \{$

$\text{sum}[i] = \text{sum}[i] + j;$

$\}$

$\} \text{for } (i=(64-l) \times 16 + 1; i <= (64-l+1) \times 16; i++) \{$

$\text{sum}[i] = 0;$

$\text{for } (j=1; j <= i; j++) \{$

$\text{sum}[i] = \text{sum}[i] + j;$

$\}$

Full
group
parallel
with 1st
box.



⇒ MPI (Message Passing Interface):-

```
main (int argc, char * argv){  
    int myrank, size;  
    MPI_Init (&argc, &argv);
```

g

```
    MPI_Finalize();
```

~~MPI_Comm_rank MPI_Comm_size~~
MPI_Comm_size (MPI_COMM_WORLD, & size);
MPI_Comm_rank (MPI_COMM_WORLD, & myrank);



MPI-Send (buf, count, datatype, dest, tag, communicator);

MPI-Recv (buf, count, datatype, source, tag, communicator, status);

int n, y myrank, msgtag = 1; status;

MPI_Status status;

MPI_Comm_rank (MPI_COMM_WORLD, &myrank);

if (myrank == 0) {

 MPI_Send (&n, 1, MPI_INT, 1, msgtag,
 MPI_COMM_WORLD);

}

else if (myrank == 1) {

 MPI_Recv (&y, 1, MPI_INT, 0, msgtag,
 MPI_COMM_WORLD, &status);

}



7) # include <mpi.h>

```
main (int argc, char *argv) {
```

```
    char message [20];
```

```
    int i, rank, size, type = 99;
```

```
    MPI_Status status;
```

```
    MPI_Init (&argc, &argv);
```

```
    MPI_Comm_Size (MPI_COMM_WORLD, &size);
```

```
    MPI_Comm_Rank (MPI_COMM_WORLD, &rank);
```

```
    if (rank == 0) {
```

```
        strcpy (message, "Hello, world");
```

```
        for (i = 1; i < size; i++) {
```

```
            MPI_Send (message, 13, MPI_CHAR, i, i,  
                      type, MPI_COMM_WORLD);
```

```
}
```

```
else {
```

```
    MPI_Recv (message, 13, MPI_CHAR, 0, type,
```

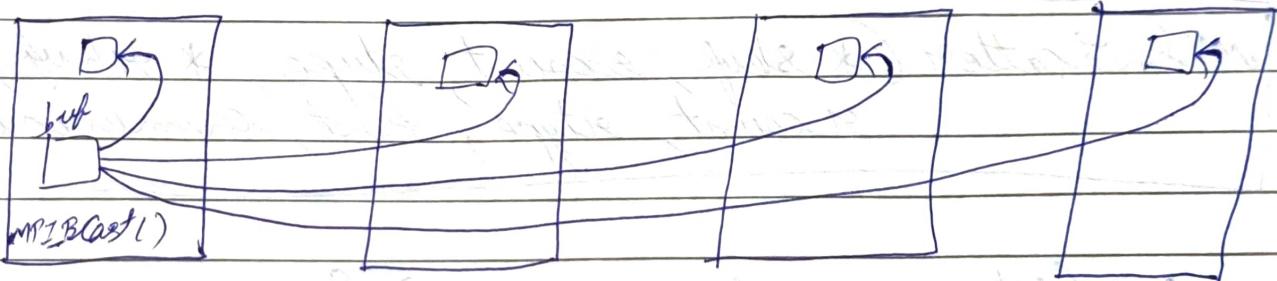
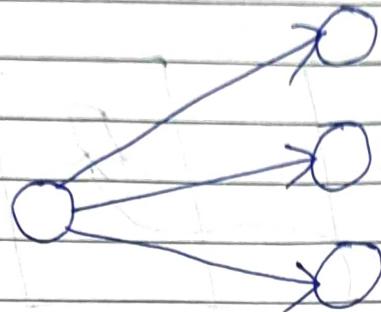
```
              MPI_COMM_WORLD, &status);
```

```
    printf ("Message from process = %d : %s\n", rank,  
           message);
```

```
    MPI_Finalize ();
```

```
}
```

⇒ MPI Broadcast :-



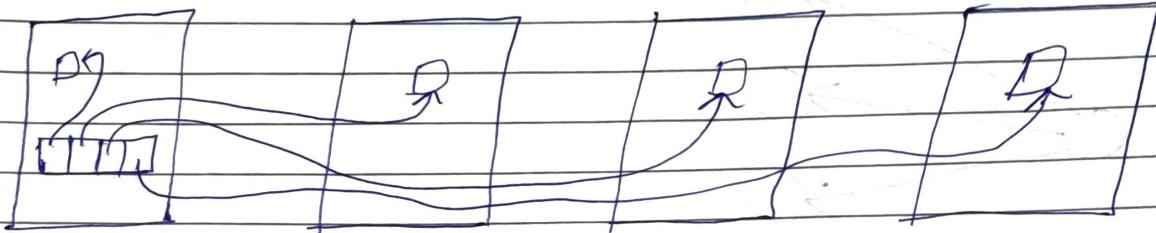
[MPI_Broadcast(*buf, count, datatype, root, comm);]

double A[N][N];

MPI_Broadcast(A, NXN, MPI_DOUBLE, 0, MPI_COMM_WORLD);



⇒ MPI-Scatter:-



~~MPI-0~~

[MPI-Scatter(* sbuf, scount, stype, * rbuf,
rcount, rtype, root, communicator);]

main (int argc, char *argv []) {
 int size, *sendbuf, recvbuf[100];

MPI_Init (&argc, &argv);

MPI_Comm_size (MPI_COMM_WORLD, &size);
 sendbuf = (int *) malloc (size * 100 * sizeof (int));

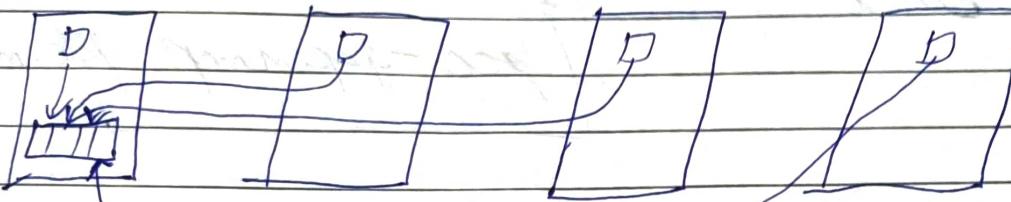
MPI_Scatter (sendbuf, 100, MPI_INT, recvbuf, 10,
 MPI_INT, 0, MPI_COMM_WORLD);

MPI_Finalize ();

3



⇒ MPI-Gather :-



`MPI_Gather(*sbuf, scount, stype, *rbuf, rcount,
rdtype, root, comm);`

Program:- change MPI-Scatter to MPI-Gather
in previous program.

⇒ MPI-Reduce :-

`MPI_Reduce(*sbuf, *recvbuf, count, datatype,
op, root, communicator);`

↳ MPI-MAX
MPI-MIN
MPI-SUM
MPI-PROD



→ OpenMP

• Thread Based

→ To compile

[gcc -fopenmp file-name]

#include <omp.h>

#pragma omp parallel

#include <omp.h>

#include <stdio.h>

int main (int argc, char *argv[]) {

#pragma omp parallel

printf ("Hello Minhaaz from thread id %d of
%d \n", omp_get_thread_num(),
omp_get_num_threads());

}

~~Expt~~
Program of π
i) thread did
ii) it's the road

MINTHA AZ
Date _____
Page No. _____



pragma omp barrier

pragma omp parallel for reduction (+: sum)

```
for (k=0; k<100; k++) {  
    sum = sum + funct(k);  
}
```

to get final sum
+ sum
* prod.
^ min
max

& logical AND
|| " OR
& bitwise AND
| " OR

② # pragma omp critical {}

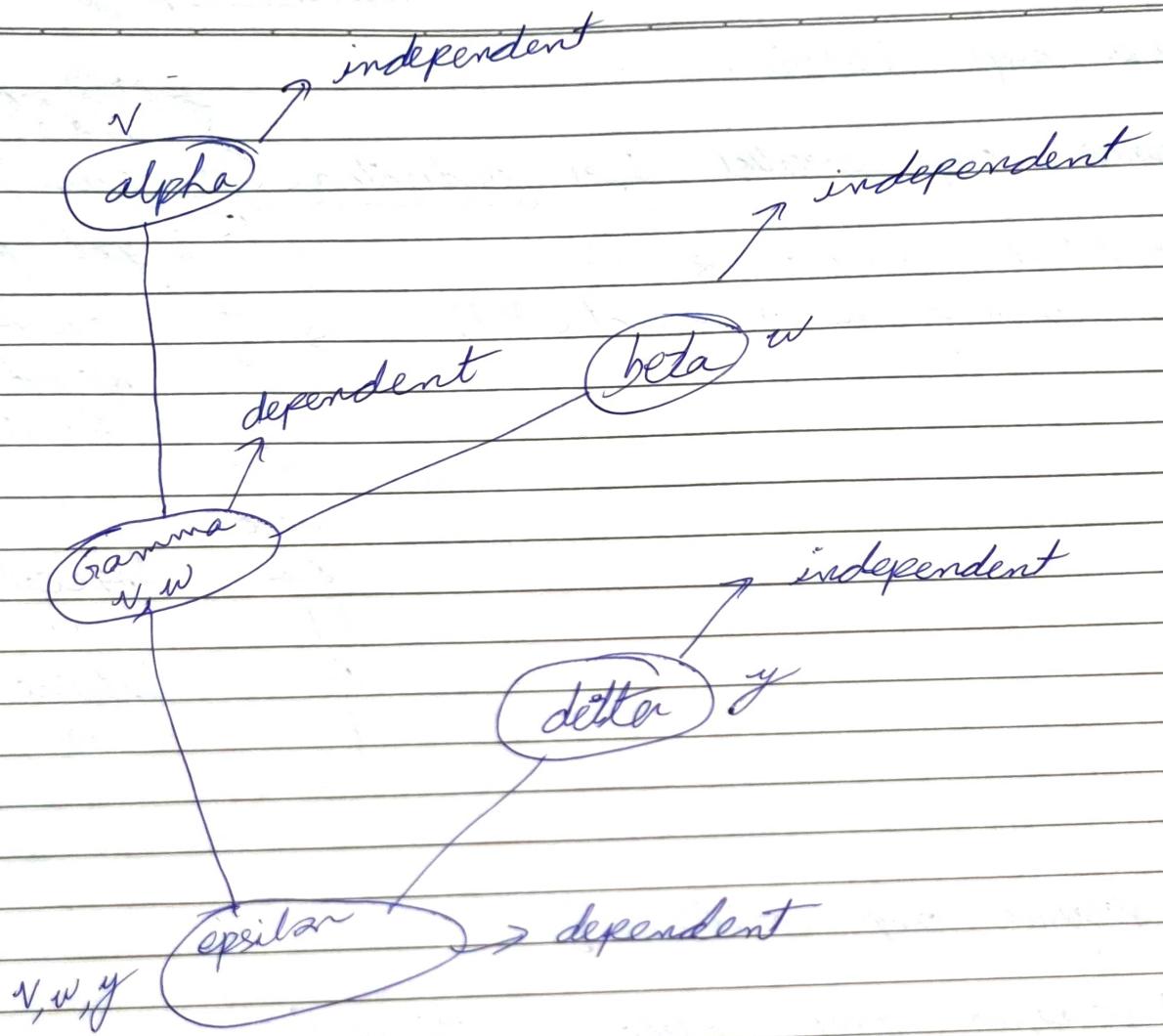
③ # pragma omp parallel sections {}

pragma omp section
v = alpha();

pragma omp section
w = beta();

pragma omp section
y = delta();

3





pragma omp parallel {

pragma omp sections {

pragma omp section
 $v = \alpha();$

pragma omp section
 $w = \beta();$

}

pragma omp sections {

pragma omp section
 $x = \gamma(v, w)$

pragma omp section
 $y = \delta();$

3

2) int omp_get_num_procs(); gives no. of physical processors.
 omp_set_num_threads(t);



⇒ Combination of MPI & openMP (Hybrid Programming)

