

Name - Ashutosh soni

Id - 2018UCPI505

Ques \Rightarrow

(a)

Ans \Rightarrow

Parallel prefix problem is about finding sum of consecutive element of an array coming earlier to that index.

Let suppose,

we are given n values stored in array as

$a_1, a_2, a_3, \dots, a_n$.

the prefix sum problem is to compute n quantities.

$a_1, a_1 + a_2, a_1 + a_2 + a_3, \dots, a_1 + a_2 + \dots + a_n$.

► Parallel solution for this problem.

→ first we divide our initial array to subarrays equal to no. of threads. Each thread then linearly calculate the prefix sum for the assigned sub-array. This prefix sum will be less than the actual sums because elements before the start of a particular sub-array are ignored. The last element then stored in another array. Again the new array is formed now we calculate the prefix sum for this newly created array simply as we calculate prefix sum for earlier subarrays.

Example Given :-

5 1 0 2.

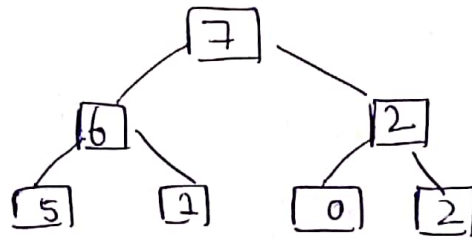
Pass - 1

1) $\boxed{5}$ $\boxed{1}$ $\boxed{0}$ $\boxed{2}$

2) 5 $\boxed{6}$ 0 $\boxed{2}$

3) 5 6 0 $\boxed{8}$

Part 2: Down Tree for earlier.

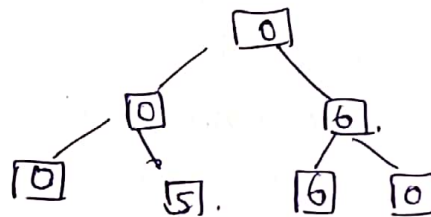


clear from that

4) 5, 6, 0, [0] } → clear

5) 5, [0], 0, [6] }

6) 0, [5], [6], [6]



(b)

Bubble sort

procedure BUBBLE_SORT(n)

begin

id := process's level.

for i := 1 to n do.

if i is odd and id is odd then.

compare-exchange_min(id+1);

else.

compare-exchange_max(id-1);

Name - Ashutosh Soni

Id - 2018 UCP 1505

if i is even and id is even then.
compare-exchange-max($id+1$);

else

compare-exchange-min($id-1$);

end for.

end BUBBLE-SORT.

above is the algorithm for Bubble sort.

we are dividing it into two phases even phase and odd phase. Even phase or even numbered processes exchange number with the number adjacent to the right and same with odd phase or odd numbered phase.

ex:

9 7 3 8 5.

7 9 3 8 5

7 3 9 5 8

3 7 5 9 8

3 5 7 9 8 ~ finalized ascending array

Performance:

Parallel Bubble sort sort the array in ~~ascending~~ ascending or descending order in $O(n)$.

whereas sequential Bubble sort sort the array in $O(n^2)$. Hence from looking at order of both we can clearly see the parallel bubble sort is far better than sequential quick sort.

Que 2:

(a)

Ans: General Semaphore:-

A semaphore, whose integer component can take arbitrary non-negative values is called a general semaphore.

Mathematically,
for Semaphore.

$$S.V = K \quad \text{where } K \geq 0$$

$$\& \cdot S.L = \phi$$

is known as general semaphore.

(b)

Readers-Writers Problem using Monitors..

Algorithm for Readers-Writers problem using Monitors:-

monitor Reader-Writer.

integer readers $\leftarrow 0$

integer writers $\leftarrow 0$.

condition OKtoRead, OKtoWrite

operation. StartRead.

if writers $\neq 0$ or not empty (OKtoWrite).

wait(OKtoRead).

readers \leftarrow readers + 1.

signal(OKtoRead)

operation. EndRead.

readers \leftarrow readers - 1;

if readers = 0.

signal(OKtoWrite).

operation, start write.

if writers $\neq 0$ or readers $\neq 0$
wait (OK to Write).

writers \leftarrow writers + 1

operation End write.

writers \leftarrow writers - 1.

if Empty (OK to Read).

then signal (OK to write).

else signal (OK to Read).

reader

p1: RW, start Read.

p2: read the data.

p3: RW, end Read.

writer.

a1: RW, start write.

a2: write the data.

a3: RW, end write.

above is the algorithm for readers writers problem.
where we declare two condition variable

1) OK to Read

2) OK to Write

and 2 integer value which will contain Readers numbers
& writers numbers & there are four operations

1) Start Read

2) end Read

3) Start Write

4) end Write.

• Readers are processes which require to exclude writers only.
NO condition with Readers, whereas writers are the
process which require to exclude both readers and the
writers. Accordingly we use wait and signal to
get desired result mentioned above.

Que 3:

Ans →

(a)

In OPENMP,

Data Scope Attribute clauses are used to define the scope of variable. By using data-scope Attribute we limit the scope of variable like is visible outside or not like that.

There are many Data scope Attribute classes listed below.

- 1) PRIVATE
- 2) FIRSTPRIVATE
- 3) LASTPRIVATE
- 4) SHARED
- 5) DEFAULT
- 6) REDUCTION
- 7) COPYIN

Data-Attribute clauses in OPENMP is used in conjunction in several directives to control the scoping of the variable. These constructs provide the ability to control the data environment during execution of parallel constructs.

They are only effective within lexical / ~~ext~~ static extent.

- PRIVATE clause → In this clause the variable declared is private to the thread.

private (list);

- SHARED clause → In this clause the variable declared is shared to all the threads in the team.

shared (list);

- DEFAULT clause → This clause allows the user to specify a default scope for all variables in the lexical extent of any parallel region.

~~def~~ default (shared | none);

- FIRSTPRIVATE: This clause combines the behavior of the PRIVATE clause with automatic initialization of the variables in its list.

firstprivate (list)

- LASTPRIVATE: This clause combines the behaviour of the PRIVATE clause with a copy from the last loop iteration, or selection to the original, variable object

lastprivate (list).

- REDUCTION → This clause performs a reduction on the variable that appears in its list. A private copy of each variable created for each thread.

reduction (operator : list).

- COPYIN: specifies that the data in the master thread, of the team to be copied to the thread private, copies of the common block, at the beginning of the parallel region

copyin (list).

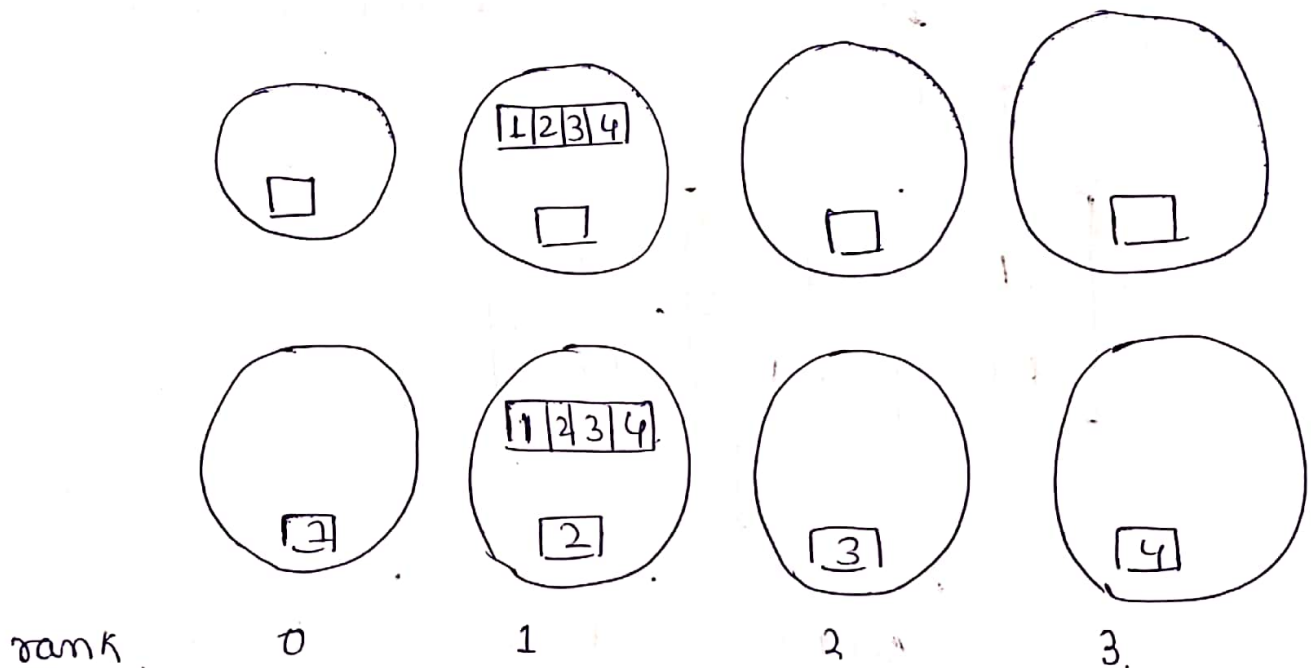
(b)

(i) Scatter:

It is a type of one to all communication in which different data sent to each process in the communicator in (rank order):

Syntax:

```
int MPI_Scatter(void* sendbuf, int sendcount,
MPI_DATATYPE sendtype, void* recvbuf,
int recvcount, MPI_Datatype, recvtype,
int root, MPI_Comm comm);
```

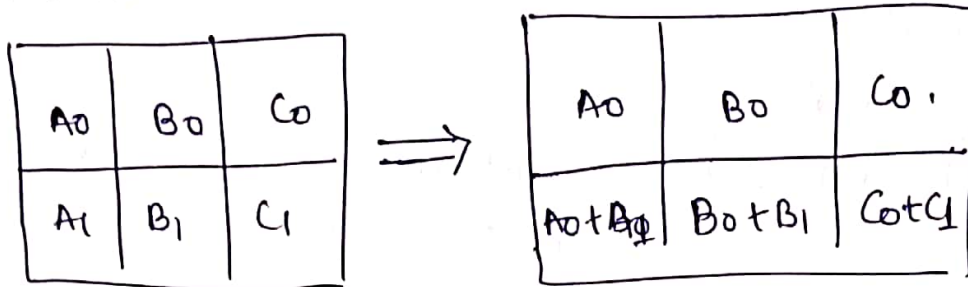


this is the example of Scatter clause.

(ii) Scan \rightarrow Parallel Prefix.

Use the value of result from previous process, in the next process.

MPI_Scan.



(iii) All to All Reduction.

In this there is no Root process & all get Result.

A_0	B_0	C_0
A_1	B_1	C_1

\Downarrow

$A_0 + A_1$	$B_0 + B_1$	$C_0 + C_1$
$A_0 + A_1$	$B_0 + B_1$	$C_0 + C_1$

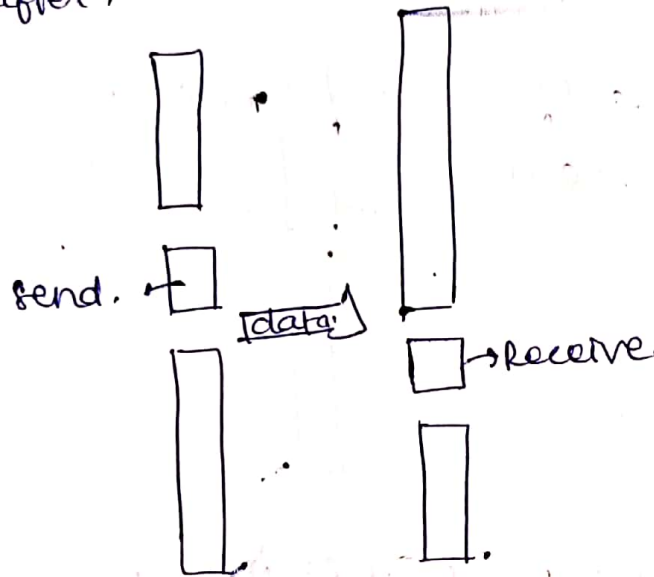
(c).

Ans →

send & receive in MPI for both blocking & non-blocking operation.

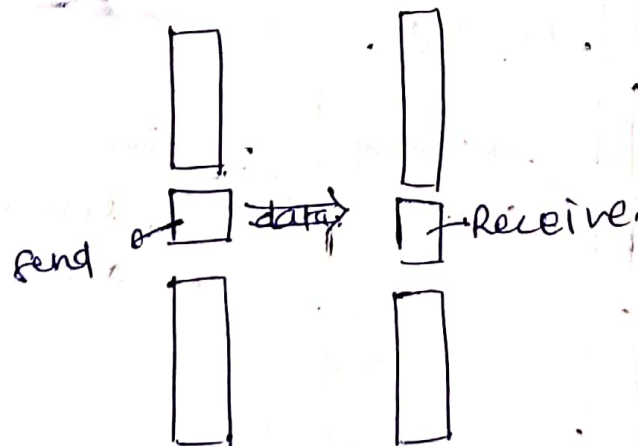
1) for Blocking & Buffered operation.

→ In this sending process returned after data has been copied into communication buffer.



2) for Non buffered blocking operation

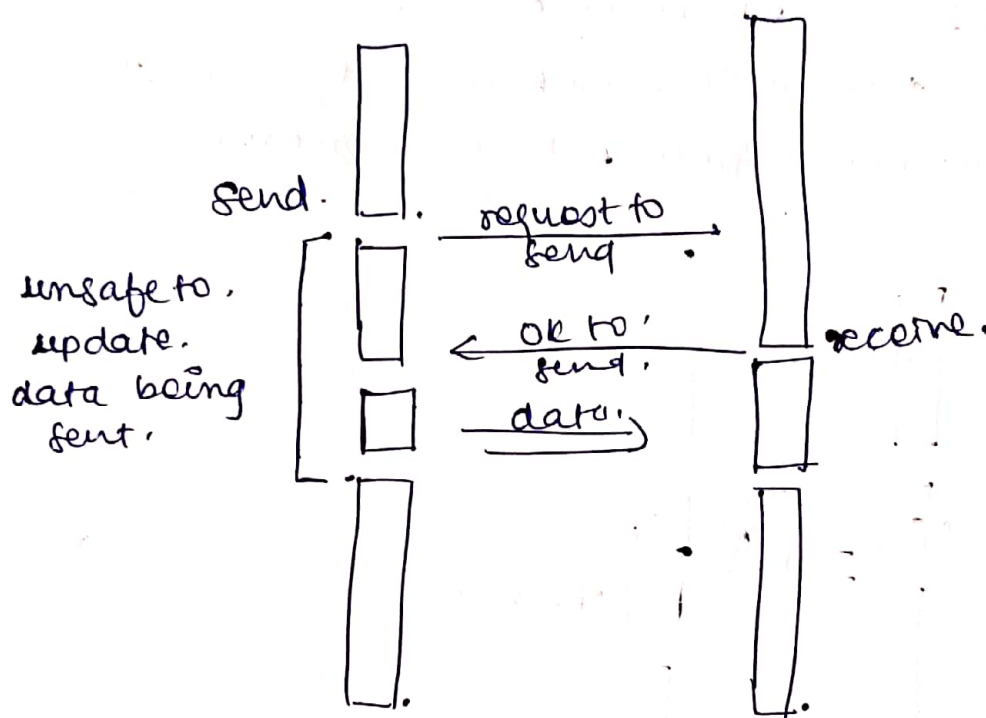
sending process blocks until ~~block~~ matching receive operation has been encountered.



before that process has blocked until rec. for that send get encountered.

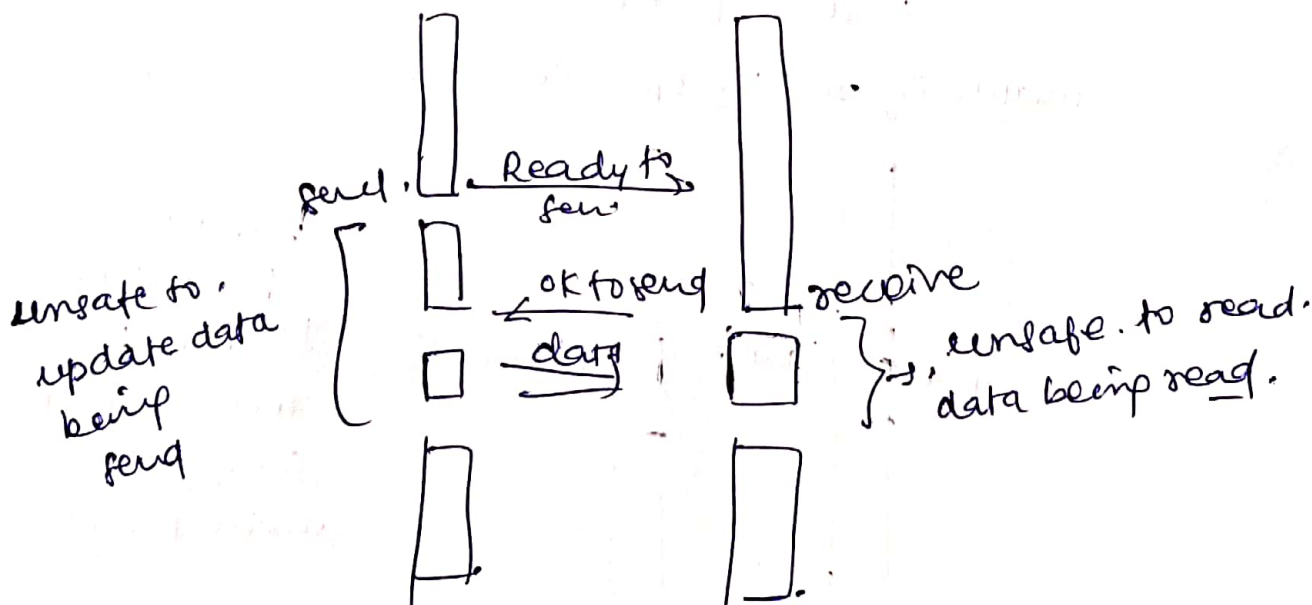
(c). Buffered Non-Blocking operation.

→ Sending process starts after initializing DMA transfer to buffer. This operation may not be completed on return.



(d). Non-Buffered Non-Blocking operation.

→ In this there is no buffer stored for sending & receiving.



In ~~Blocker~~ Buffered case there is buffer to store data, whereas not in case of Non-Buffered operation case.

- Send & Receive semantics assured by corresponding operation in Blocking operation
- Programmer must explicitly ensure semantics by polling to verify ~~open~~ completion in case of Non-Blocking operation.

Name - Ashutosh Sone

Id - 2018UCPI505

Ques 4:

(a)

Ans: memory hierarchy terminology in CUDA and OpenCL.

OPENCL.

CUDA

local \rightarrow within a thread.

private \rightarrow within a work-item.

shared \rightarrow shared between threads in a thread block

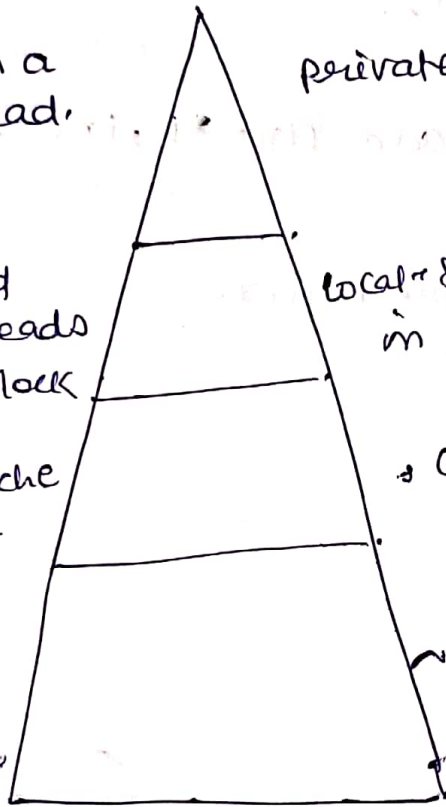
local \rightarrow shared between work-item in a work group.

constant \rightarrow a cache for constant memory

constant \rightarrow a cache for constant memory.

device \rightarrow shared between all thread blocks

Global \rightarrow shared between all work groups.



Name - Ashutosh Soni

Id - 2018UCPI505

(b)

Ans:

Program for addition of two square matrix in CUDA.

```
#include <stdio.h>
```

```
#include <cuda.h>
```

```
--global-- void matAdd (int A[][N], int B[][N],  
                        int C[][N]) {
```

```
    int i = threadIdx.x;
```

```
    int j = threadIdx.y;
```

```
    C[i][j] = A[i][j] + B[i][j];
```

```
}
```

```
int main() {
```

```
    int A[N][N] = { {1, 2}, {4, 5} };
```

```
    int B[N][N] = { {4, 5}, {5, 6} };
```

```
    int C[N][N] = { {0, 0}, {0, 0} };
```

```
    int *pA[N], *pB[N], *pC[N];
```

```
    cudaMalloc (pA, (N*N)*sizeof(int),
```

```
               cudaMemcpyHostToDevice);
```

```
    cudaMalloc
```

Name - Ashutosh Soni

Id - 2018UCP1505

```
cudaMalloc ((void**)&PA, (N*N)*(sizeof(int)));  
cudaMalloc ((void**)&PB, (N*N)*(sizeof(int)));  
cudaMalloc ((void**)&PC, (N*N)*(sizeof(int)));
```

```
cudaMemcpy (PA, A, (N*N)*sizeof(int), cudaMemcpyHostTo,  
            Device);
```

```
cudaMemcpy (PB, B, (N*N)*sizeof(int), cudaMemcpyHostTo  
            Device);
```

```
cudaMemcpy (PC, C, (N*N)*sizeof(int), cudaMemcpyHost  
            ToDevice);
```

```
int numBlocks = 1;
```

```
dim3 threadsPerBlock(N, N);
```

```
MatAdd <<<numBlocks, threadsPerBlock>>>
```

```
(A, B, C);
```

```
cudaMemcpy (C, PC, (N*N)*sizeof(int),  
            cudaMemcpyDeviceToHost);
```

```
int i, j;
```

```
printf (" ");
```

Name - Ashutosh Soni

Id - 2018UCP1505

```
for (int i=0; i<N; i++) {  
    for (int j=0; j<N; j++) {  
        printf("%d .", e[i][j]);  
    }  
    printf("\n");  
}
```

cudaFree (PA);

Cuda Free (PB);

cuda Free (PC);

printf("\n");

return 0;

}

Performance on Shared Memory :-

As in shared Memory we,

- Partition data into subsets that fit in shared memory.
- Handle each data subset with one thread block by;
 - Loading the subset from global memory to shared memory

Name - Ashutosh Soni

Id - 2018UCP1505

- Performing the computation on the subset from shared memory, each thread on efficiently multipass over every data element
- Copying results from shared memory to global memory.

by doing this,

we can increase the performance of the system a lot.