

- Q5 A) Differentiate between Non-Blocking and Blocking message passing operations. 8
- Q5 B) What is bitonic sort? State parallel formulation of bitonic sort on a hypercube with  $n=2^d$  processes. 6

Final Year B.Tech. (Computer Science and Engineering)

MAKEUP EXAMINATION SEM. I APRIL/MAY-2017

PROFESSIONAL ELECTIVE V HIGH PERFORMANCE COMPUTING (2CS411)

Exam Seat Number: \_\_\_\_\_

Date and Time: Saturday, 13/05/2017, 02.00pm to 05.00pm

Max Marks: 100

**IMP: Verify that you have received question paper with correct course, code, branch etc.**

- Instructions:**
- All questions are compulsory. Writing question number is compulsory. The answers may not be assessed if question number is not written.
  - Figures to the right of question text indicate full marks.
  - Assume suitable data wherever necessary.
  - Write the answers with neat handwriting.

		Marks
On the right of marks indicates course outcomes (only for faculty use).		
Q1 A)	Explain omega network with an interconnection pattern. Show a perfect shuffle interconnection of omega network for eight inputs and outputs.	8 CO1
Q1 B)	Show a cost optimal way of computing the sum of 16 numbers using four processing elements. Calculate parallel runtime and prove how parallel system is cost optimal.	8 CO2
Q1 C)	Derive the expression for communication cost of Store-and-Forward Routing and Packet Routing.	6 CO1
Q2 A)	Why decomposition techniques are necessary? Explain any one in detail.	8 CO2
Q2 B)	Calculate Diameter, Bisection Width, Arc Connectivity, Cost for Completely connected Network, Hypercube Network and 2D wraparound mesh.	6 CO1
Q2 C)	Define isoefficiency function. The serial runtime of an algorithm is $n^3$ and the parallel runtime of an algorithm is $n^3/p + n^2/\sqrt{p} + p \log p$ . What is the isoefficiency of the parallel algorithm?	6 CO2
Q3 A)	What do you mean by scalable parallel system? Explain with scaling characteristics.	6 CO2
Q3 B)	Write a MPI program for simulation of broadcast operation among N processes.	8 CO3
Q3 C)	Write a CUDA program for vector multiplication using N blocks and N threads.	8 CO3
Q4 A)	How Dijkstra's algorithm of single-source shortest paths can be parallelised?	8 CO3
Q4 B)	Prove that Cannon's algorithm is memory-efficient version of the simple matrix multiplication algorithm.	8 CO3
Q4 C)	Describe CUDA memory model in detail.	6 CO1

- Q3 B) What do you mean by exploratory decomposition? Show the scenario where work performed by the parallel algorithm is greater than that performed by the serial algorithm.  
--OR--  
What do you mean by partitioning output data? Illustrate output decomposition method by the example of matrix-vector multiplication.
- Q3 C) Discuss in detail clauses of OpenMP that do not have implicit barrier.



Day, Date and Time: Tuesday, 11/10/2022, 03.00 pm to 04.30 pm

Max Marks: 30

**IMP: Verify that you have received question paper with correct course, code, branch etc.**

- Instructions:
- All questions are compulsory.
  - Writing question number on answer book is compulsory otherwise answers may not be assessed.
  - Assume suitable data wherever necessary.
  - Figures to the right of question text indicate full marks.
  - Mobile phones and programmable calculators are strictly prohibited.
  - Except PRN anything else writing on question paper is not allowed.
  - Exchange/Sharing of stationery, calculator etc. not allowed.

Text on the right of marks indicates course outcomes (only for faculty use).

- |   | Marks |
|---|-------|
| Q1 A) Consider an SMP with a distributed shared-address-space. Consider a simple cost model in which it takes 10ns to access local cache, 100 ns to access local memory and 400 ns to access remote memory. A parallel program is running on this machine. The program is perfectly load balanced with 80% of all accesses going to local cache, 10% to local memory and 10% to remote memory. What is the effective memory access time for this computation? If the computation is memory bound, what is the peak computation rate? Now consider the same computation running on one processor. Here the processor hits the cache 70% of the time and local memory 30% of the time. What is the effective peak computation rate for one processor? What is the fractional computation rate of a processor in a parallel configuration as compared to the serial processor? | 4 CO2 |
| Q1 B) Why it is difficult to construct true shared-memory computer ? What is the minimum number of switches for connecting p processors to a shared memory with b words( Where each word can be accessed independently) ?   | 3 CO1 |
| Q1 C) If the diameter of an interconnection network is 4, would you prefer a store-and-forward network with $ts=10$ and $tw=0.1$ or a cut through network with $ts=10$ and $tw=0.5$ . Justify your answer   | 3 CO1 |
| Q2 A) Show cost optimal way of computing the sum of 16 numbers using four processing elements. Show that parallel system is cost optimal by deriving cost of serial and parallel version.   | 4 CO2 |
| Q2 B) The sequential version of a program executes in 10000 hours. Executing on 1000 processors it executes in 20 hours. What is the speedup and efficiency? How long would the program take to execute if the efficiency when executing on 1000 processors was 100%?   | 3 CO2 |
| Q2 C) Prove the following statement<br>Multistage interconnection network is more scalable than the bus in terms of performance and more scalable than crossbar in terms of cost.   | 4 CO1 |
| Q3 A) Which speed up could be achieved according to Amdahl's Law for infinite number of processes if 5% of a program is sequential and the remaining part is ideally parallel ?   | 3 CO2 |

- D) OpenMP API defines which of the following Work sharing constructs?  
 a) Loop constructs  
 b) Section constructs  
 c) Single constructs
- E) What are the major issues with non-buffered blocking send? Draw the handshaking procedure in blocking non-buffered send-receive operation considering all possible scenarios.

- Q3** A) If you launch a CUDA kernel with following parameters

```
dim3 block_dim (128,1,1);
dim3 grid_dim (10,1,1);
kernel<<<grid_dim,,block_dim>>>(...)
```

Then what ranges you would effectively have for following variables?  
`threadIdx.x , blockIdx.x , blockDim.x , gridDim.x`

- B) Explain Device memory management in CUDA.  
 C) Write a CUDA C/C++ Program to reverse the contents of 1D array of size N using multiple CUDA Blocks.

- Q4** A) How in a Multi-processor system, cache coherence is maintained using invalidate protocol? Explain with the example of parallel program execution with the same protocol.  
 B) Show the steps of All-to-all broadcast on a Hypercube and analyze the cost of the operation.

-OR-

Show the steps of Prefix-sum operation on a Hypercube and analyze the cost of the operation.

- C) What are the performance metrics of parallel systems? Derive equation for each parameter.

.....End of question paper.....



Day & Date: Saturday, 17/12/2022

Time: 3.00 pm to 5.00 pm

PRN: \_\_\_\_\_

Max Marks: 50

**IMP: Verify that you have received question papers with correct course code, branch etc.**

- Instructions**
- a) All questions are compulsory.
  - b) Writing question number on answer book is compulsory otherwise answers may not be assessed.
  - c) Assume suitable data wherever necessary.
  - d) Figures to the right of question text indicate full marks.
  - e) Mobile phones, smart gadgets and programmable calculators are strictly prohibited.
  - f) Except PRN anything else writing on question paper is not allowed.
  - g) Exchange/Sharing of stationery, calculator etc. not allowed.

Text on the right of marks indicates course outcomes (Only for faculty use) Marks

**Q1 A)** How to form a Bitonic Sequence from a random input? Convert the following sequence 6 CO2  
to a Bitonic sequence: 3, 7, 4, 8, 6, 2, 1, 5.

**B)** Why parallel run-time of the Cannon's matrix-matrix multiplication is 6 CO3

$$T_p = \frac{n^3}{P} + 2\sqrt{P}t_s + 2t_w \frac{n^2}{\sqrt{P}} ?$$

**C)** Why in parallel formulation of Prim's algorithm it is hard to parallelize outer loop? 4 CO3  
Explain the data partitioning approach taken in above algorithm.

**Q2 A)** Which of the following statement is/are true about MPI\_INIT and 1 CO1  
MPI\_FINALIZE?

- a) MPI\_INIT must be called before any other MPI function and called exactly once per process
- b) No further MPI function can be called after MPI\_FINALIZE
- c) Both a) and b)
- d) None of the above

**B)** Which function Gathers data from all tasks and distribute the combined data to all tasks? 1 CO1  
a) MPI\_Gather b) MPI\_Allgather c) MPI\_Reduce d) MPI\_Allreduce

**C)** Which of the following statement about OpenMP is incorrect? 1 CO1

- a) OpenMP is an API that enables explicit multi-threaded parallelism.
- b) The primary components of OpenMP are compiler directives, runtime library, and environment variables.
- c) OpenMP is designed for distributed memory parallel systems and guarantees efficient use of memory.
- d) OpenMP supports UMA and NUMA architectures.

Exam Seat Number:

Date and Time: Tuesday, 27/11/2018, 10.00am to 12.00noon

Max Marks:

50

**IMP: Verify that you have received question paper with correct course, code, branch etc.**

**Instructions:** i) All questions are compulsory. Writing question number is compulsory. The answers may not be assessed if question number is not written. Assume suitable data wherever necessary.

ii) Figures to the right of question text indicate full marks.

iii) Mobile phones and programmable calculators are strictly prohibited.

iv) Except Exam Seat Number writing anything on question paper is not allowed.

**Exchange/Sharing of stationery, calculator etc. not allowed.**

**Note:** The figure on the right of marks indicates course outcomes (only for faculty use).

- Q1 A) Show the communication steps for one-to-all broadcast on a two-dimensional square mesh with  $v_p$  rows and  $v_p$  columns. Derive the expression for total time taken for this procedure. Marks 5 CO3
- Q1 B) What is collective communication? Explain following MPI routines.  
 1. MPI\_Scatter    2. MPI\_BARRIER    3. MPI\_Reduce    4. MPI\_Bcast 5 CO1

- Q2 A) Define bitonic sequence. Draw +BM[16] bitonic merging network for following input numbers. 3,5,8,9,10,12,14,20,95,90,60,40,35,23,18,0 4 CO3  
 -OR-

What are the advantages of Buffered Send and Receive Operations? Is there any possibility of deadlock in this operation? If yes write a simple code fragment to prove this.

- Q2 B) How to parallelize Prim's algorithm? Derive parallel run time of the algorithm. 4 CO2  
 Q2 C) How to perform Matrix-vector multiplication with block 2-D partitioning where  $n \times n$  matrix is partitioned among  $n^2$  processes with each process owns a single element. Calculate the parallel run time of this algorithm. Is this algorithm cost optimal? 5 CO3

-OR-

The overhead function of the cost-optimal parallel implementation of matrix-vector multiplication with block 2-D partitioning and using fewer than  $n^2$  process is as follows.  $T_0 = ts p \log p + tw n\sqrt{P} \log P$  Derive asymptotic isoefficiency function of scalability.

- Q3 A) Differentiate between CPU and GPU. 3 CO1  
 Q3 B) Write a program for vector addition on a GPU using CUDA API. 5 CO3  
 Q3 C) Draw CUDA Memory Architecture and elaborate on all memory types. 4 CO1

- Q4 A) What is Omega network? Is this a blocking network? For eight inputs and outputs show perfect shuffle interconnection pattern. 5 CO1  
 Q4 B) Consider the problem of adding  $n$  numbers on  $p$  processing elements such that  $p << n$  and both  $n$  and  $p$  are powers of 2. For values of  $n=16$  and  $p=4$ , show the cost-optimal way of computing the sum. Also, derive parallel runtime. 5 CO2

- Q4 C) What is recursive decomposition? Draw task-dependency graph for finding the minimum of eight numbers. Numbers are as follows. 4,9,1,7,8,11,2,12 5 CO2

(P.T.O.)

Q4 C)

Write a CUDA program for calculation of dot product of two vectors. Analyze the program and comment on the essential statements present in the code in view of getting correct result of dot product.



Exam Seat Number: \_\_\_\_\_  
Date and Time: Wednesday, 04/12/2019, 10.00am to 12.00Noon

Max Marks: **50**

- IMP: Verify that you have received question paper with correct course, code, branch etc.**  
Actions: i) All questions are compulsory. Writing question number is compulsory. Assume suitable data wherever necessary.  
ii) Figures to the right of question text indicate full marks.  
iii) Mobile phones and programmable calculators are strictly prohibited.  
iv) Except Exam Seat Number writing anything on question paper is not allowed. Exchange/Sharing of stationery, calculator etc. not allowed.

		Marks
on the right of marks indicates course outcomes (only for faculty use).		
A)	i) If a problem of size $W$ has a serial component $W_s$ , prove that $W/W_s$ is an upper bound on its speedup, no matter how many processing elements are used.  ii) If serial component is 0.1 and number of processing elements( $n$ ) used are 10 calculate the speedup. Calculate the speedup for values of $n=100, n=1000, n=100000$ and state your observation.	5 CO2
B)	Show the parallel formulation of matrix-vector multiplication using fewer than $n$ processes in which the matrix is 1-D block partitioned along the rows and the vector is equally partitioned among all the processes. State the parallel run time of the algorithm. Is the algorithm cost-optimal?  —OR— How will you parallelize Prim's minimum spanning tree algorithm? Justify the parallelism in the algorithm. State the parallel run time of the formulation.	6 CO3
A)	a) The serial runtime of an algorithm is $n^3$ and the parallel runtime of an algorithm is $n^3/p + n^2/\sqrt{p} + p \log p$ . What is the isoefficiency of the parallel algorithm?  b) A second parallel algorithm has a runtime of $(n^3 \log p)/p + n / \sqrt{p}$ . Would you prefer the parallel algorithm in part (a) above or part (b) for a very large number of processors?	6 CO2
B)	What do you mean by a cluster? Draw cluster computer architecture and elaborate on some prominent components of cluster computer.	4 CO1
C)	What is collective communication? Explain the following MPI routines. 1. MPI_Scatter    2. MPI_BARRIER    3. MPI_Reduce    4. MPI_Bcast	5 CO1
A)	Show how one-to-all broadcast and all-to-one reduction on an eight-node ring can be executed optimally. Analyze the cost of both the operations.	5 CO3
B)	Show how the input sequence $\{5,3,6,2,4,8,7,1\}$ will be sorted by Bitonic Sort. What is the parallel run time of Bitonic sort?	6 CO2
A)	How many different kinds of memories are in a GPU? If all the various types of device memory were to race in terms of speed, how the race would turn out?	4 CO1
B)	What is a warp and how it is different from thread or block in CUDA?	3 CO1

**WALCHAND COLLEGE OF ENGINEERING, SANGLI.**  
*(An Autonomous Institute)*



Final Year B.Tech. (CSE)

END SEMESTER EXAMINATION NOV./DEC.-2016  
 PROFESSIONAL ELECTIVE V HIGH PERFORMANCE COMPUTING (2CS411)

ESE

Exam Seat Number: \_\_\_\_\_

Day, Date and Time: Tuesday, 06/12/2016, 03.00pm to 05.00pm

Max Marks: 50

**IMP: Verify that you have received question paper with correct course, code, branch etc.**

- Instructions: i) All questions are compulsory. Writing question number is compulsory. The answers may not be assessed if question number is not written.  
 ii) Figures to the right of question text indicate full marks.  
 iii) Assume suitable data wherever necessary.  
 iv) Write the answers with neat handwriting.

Text on the right of marks indicates course outcomes (only for faculty use).

			Marks
Q1	A)	Consider the problem of multiplying an $n \times n$ matrix A with an $n \times 1$ vector X on an $n \times n$ mesh of nodes to yield an $n \times 1$ result vector Y. Draw a diagram of possible mapping of the matrix and the vectors on to mesh architecture. Which communication operations are required for generation of resultant vector? Show communication operations clearly in above mapping.	6 CO2
Q1	B)	Compare between Blocking and Non-Blocking message passing operations. Explain MPI_send and MPI_Recv functions alongwith with meaning of parameters. -OR- Elaborate OpenMP Programming Model with all parallel directives.	6 CO2
Q2	A)	Given two vectors (i.e. arrays), write a CUDA program to add them together in a third vector. For example: A = {0, 2, 4, 6, 8} B = {1, 1, 2, 2, 1} Then A + B = C = {1, 3, 6, 8, 9}.	6 CO3
Q2	B)	Draw CUDA Memory Architecture. Discuss all CUDA memory types, their usage and limitations.	6 CO1
Q3	A)	Consider two $n \times n$ matrices A and B partitioned into p blocks $A_{i,j}$ and $B_{i,j}$ ( $0 \leq i, j < \sqrt{p}$ ) of size $(n/\sqrt{p}) \times (n/\sqrt{p})$ each. These blocks are mapped onto a $\sqrt{p} \times \sqrt{p}$ logical mesh of processes. The processes are labeled from $P_{0,0}$ to $P_{\sqrt{p}-1, \sqrt{p}-1}$ . Multiplication of above two $n \times n$ dense, square matrices A and B is to be carried out to yield the product matrix $C = A * B$ . What is the memory efficient way to achieve this? Justify how it is memory efficient by showing communication steps. Also estimate its parallel runtime.	8 CO3
Q3	B)	What is bitonic merge? Show merging of following 16-element bitonic sequence through a series of log 16 bitonic splits to sort the sequence . 3 5 8 9 10 12 14 20 95 90 60 40 35 23 18 0 -OR- How parallel formulation of Prim's minimum spanning tree algorithm is possible? Estimate parallel runtime of this formulation.	4 CO3
Q4	A)	Derive the expression for isoefficiency function of scalability.	4 CO2
Q4	B)	What is exploratory decomposition? Why the work performed by the parallel formulation can be either smaller or greater than that performed by the serial algorithm in exploratory decomposition ? Give proper example for justification.	5 CO2
Q4	C)	What is Omega network? Is this a blocking network? For eight inputs and outputs show perfect shuffle interconnection pattern.	5 CO1

			8	C0
Q5	A)	Given two vectors (i.e. arrays), write a CUDA program to add them together in a third vector. For example: A = {0, 2, 4, 6, 8} B = {1, 1, 2, 2, 1} Then A + B = C = {1, 3, 6, 8, 9}.	6	CO
Q5	B)	Draw CUDA Memory Architecture. Discuss all CUDA memory types, their usage and limitations.	6	CO
Q6	A)	What is bitonic merge? Show merging of following 16-element bitonic sequence through a series of $\log 16$ bitonic splits to sort the sequence . 3 5 8 9 10 12 14 20 95 90 60 40 35 23 18 0	6	CO
Q6	B)	Consider two $n * n$ matrices A and B partitioned into p blocks $A_{i,j}$ and $B_{i,j}$ ( $0 \leq i, j < \sqrt{p}$ ) of size $(n/\sqrt{p}) * (n/\sqrt{p})$ each. These blocks are mapped onto a $\sqrt{p} * \sqrt{p}$ logical mesh of processes. The processes are labeled from $P_{0,0}$ to $P_{\sqrt{p}-1, \sqrt{p}-1}$ . Multiplication of above two $n * n$ dense, square matrices A and B is to be carried out to yield the product matrix $C = A * B$ . What is the memory efficient way to achieve this? Justify how it is memory efficient by showing communication steps. Also estimate its parallel runtime.	8	CO
Q6	C)	How to parallelize Prim's algorithm? Derive parallel run time of the algorithm.	4	C



# WALCHAND COLLEGE OF ENGINEERING, SANGLI.

(An Autonomous Institute)

MakeUp

Final Year B.Tech. (Computer Science and Engineering)

MAKEUP EXAMINATION: SEMESTER I MAY-2019

PROFESSIONAL ELECTIVE V HIGH PERFORMANCE COMPUTING (3CS411)

Exam Seat Number: \_\_\_\_\_  
Day, Date and Time: Saturday, 11/05/2019, 02.00pm to 05.00pm

Max Marks: 100

**IMP: Verify that you have received question paper with correct course, code, branch etc.**

Instructions: i) All questions are compulsory. Writing question number is compulsory. The answers may not be assessed if question number is not written. Assume suitable data wherever necessary.

ii) Figures to the right of question text indicate full marks.

iii) Mobile phones and programmable calculators are strictly prohibited.

iv) Except Exam Seat Number writing anything on question paper is not allowed. Exchange/Sharing of stationery, calculator etc. not allowed.

Text on the right of marks indicates course outcomes (only for faculty use).

		Marks	
Q1	A)	4	CO1
Q1	B)	6	CO1
Q1	C)	6	CO1
Q2	A)	6	CO2
Q2	B)	4	CO2
Q2	C)	6	CO2
Q3	A)	5	CO2
Q3	B)	7	CO2
Q3	C)	6	CO2
Q4	A)	6	CO1
Q4	B)	6	CO1
Q4	C)	6	CO1

(P.T.O.)



## WALCHAND COLLEGE OF ENGINEERING, SANGLI.

(An Autonomous Institute)

MSE

Final Year B.Tech. (CSE)

MID SEMESTER EXAMINATION SEPTEMBER / OCTOBER-2016

PROFESSIONAL ELECTIVE V HIGH PERFORMANCE COMPUTING (2CS411)

Exam Seat Number: \_\_\_\_\_

Day, Date and Time: Saturday, 01/10/2016, 03.00pm to 04.30pm

Max Marks:

30

**IMP: Verify that you have received question paper with correct course, code, branch etc.**

Instructions: i) All questions are compulsory. Writing question number is compulsory. The answers may not be assessed if question number is not written.

- ii) Figures to the right of question text indicate full marks.
- iii) Assume suitable data wherever necessary.
- iv) Write the answers with neat handwriting.

Text on the right of marks indicates course outcomes (only for faculty use).			Marks
Q1	A)	What do you mean by a cluster? Draw cluster computer architecture and elaborate on some prominent components of cluster computer.	5 CO1
Q1	B)	For a given problem size, Why efficiency of the parallel system goes down with increasing number of processing elements? Justify w.r.t.expression of efficiency. The serial fraction of a parallel program is 1%. What is the highest speedup you can achieve for this program on a 1000 processor machine?	5 CO2
Q2	A)	The serial runtime of an algorithm is $n^3$ and parallel run time of an algorithm is $n^3/p + n^2/\sqrt{p} + p \log p$ . What is the isoeficiency of the parallel algorithm? A Second parallel algorithm has a runtime of $(n^3 \log p)/p + n/\sqrt{p}$ . Which parallel algorithm Would you prefer? Why?	5 CO2
Q2	B)	What is diameter of interconnection network? If the diameter of an interconnection network is 4, would you prefer a store-and-forward network with $ts=10$ and $tw=0.1$ or a cut-through network with $ts=10$ and $tw=0.5$ . Justify your answer.	5 CO1
Q3	A)	How will you ensure that partitioning step while designing parallel algorithm has been carried out successfully ?	5 CO2
Q3	B)	Draw a complete state diagram of a simple three-state cache coherence protocol. Show example of parallel program execution.	5 CO1



**WALCHAND COLLEGE OF ENGINEERING, SANGLI.**  
*(An Autonomous Institute)*

MSE

**Final Year B.Tech. (Computer Science and Engineering)  
MID SEMESTER EXAMINATION SEMESTER- I · SEPTEMBER-2018  
PROFESSIONAL ELECTIVE V HIGH PERFORMANCE COMPUTING (3CS411)**

Exam Seat Number:

Date and Time: Saturday, 22/09/2018, 03.00pm to 04.30pm

Max Marks: **30**

**IMP: Verify that you have received question paper with correct course, code, branch etc.**

**Instructions:** i) All questions are compulsory. Writing question number is compulsory. The answers may not be assessed if question number is not written. Assume suitable data wherever necessary.

ii) Figures to the right of question text indicate full marks.

iii) Mobile phones and programmable calculators are strictly prohibited.

iv) Except Exam Seat Number writing anything on question paper is not allowed.

**Exchange/Sharing of stationery, calculator etc. not allowed.**

(A mark to the right of marks indicates course outcomes (only for faculty use).)

**Marks**

Q1 A)	Prove the following statement Multistage interconnection network is more scalable than the bus in terms of performance and more scalable than crossbar in terms of cost.	4	CO1
Q1 B)	Derive the expression for the total communication time for Store-and-Forward routing.	4	CO1
Q1 C)	Draw and explain the state diagram of a simple three-state cache coherence protocol.	4	CO1

Q2 A)	What do you mean by exploratory decomposition? Show the scenario where work performed by the parallel algorithm is greater than that performed by the serial algorithm. -OR- What do you mean by partitioning output data? Illustrate output decomposition method by the example of matrix-vector multiplication.	3	CO2
Q2 B)	Why average degree of concurrency is a useful indicator of a parallel program's performance?	2	CO2
Q2 C)	Is there an inherent bound on how fine-grained decomposition a problem permits? If yes justify your answer with an example.	2	CO2

Q3 A)	What is meant by superlinear speed up? Is it something theoretical? How it is obtained?	4	CO2
Q3 B)	Define efficiency. Calculate the efficiency of adding N numbers on N processing elements.	4	CO2
Q3 C)	Derive the expression for overhead function $T_0$ . Also state the sources of overhead in parallel programs.	3	CO2

(P.T.O.)

Q5 B) Write a CUDA program for vector multiplication considering following two cases.

1. Several blocks with one thread each
2. One block with several threads

Q6 A) How to parallelize Prim's Minimum Spanning Tree algorithm? Derive an expression for parallel runtime, Speedup, and efficiency for parallel formulation.

Q6 B) What are the different issues in sorting on parallel computers? How are comparisons performed when each processor holds more than one element?



**Final Year B.Tech. (Computer Science and Engineering)**  
**MID SEMESTER EXAMINATION OCTOBER-2017**  
**PROFESSIONAL ELECTIVE V HIGH PERFORMANCE COMPUTING (3CS411)**

Day, Date and Time: Monday, 16/10/2017, 03.00pm to 04.30pm Exam Seat Number: \_\_\_\_\_

Max Marks: 30

**IMP: Verify that you have received question paper with correct course, code, branch etc.**

- Instructions: i) All questions are compulsory. Writing question number is compulsory. The answers may not be assessed if question number is not written.  
 ii) Figures to the right of question text indicate full marks.  
 iii) Assume suitable data wherever necessary.  
 iv) Write the answers with neat handwriting.

Text on the right of marks indicates course outcomes (only for faculty use).

		Marks	
Q1 A)	What is MIMD architecture? How is it different from SIMD architecture?	3	CO1
Q1 B)	How multi-stage interconnection network lies between crossbar and bus interconnection network in terms of cost and scalability?	4	CO1
Q2 A)	Derive the expression for total communication time for cut-through routing. Is derived communication time is an improvement over store-and-forward routing?	4	CO1
Q2 B)	What is recursive decomposition? Write a recursive algorithm for finding the minimum in an array of numbers of length n.	4	CO2
Q3 A)	Suppose you are designing a parallel algorithm by using Foster's design paradigm. Which questions will you consider before finishing the communication analysis step?	3	CO2
Q3 B)	Explain with block diagram various strategies for developing parallel applications. --OR-- Draw typical architecture of cluster computer and explain its prominent components.	4	CO2
Q3 C)	Show cost-optimal way of computing the sum of 16 numbers using four processing elements.	4	CO3
Q3 D)	Write a parallel program for addition of two vectors containing n numbers using OpenMP directives.	4	CO3



Final Year B.Tech. (Computer Science and Engineering)

MAKEUP EXAMINATION APRIL/MAY-2018

PROFESSIONAL ELECTIVE V HIGH PERFORMANCE COMPUTING (3CS411)

Exam Seat Number: \_\_\_\_\_

Day, Date and Time: Monday, 07/05/2018, 02.00pm to 05.00pm

Max Marks: 100

**IMP: Verify that you have received question paper with correct course, code, branch etc.**

- Instructions: i) All questions are compulsory. Writing question number is compulsory. The answers may not be assessed if question number is not written. Assume suitable data wherever necessary.  
ii) Figures to the right of question text indicate full marks.  
iii) Mobile phones and programmable calculators are strictly prohibited.  
iv) Except Exam Seat Number writing anything on question paper is not allowed.  
Exchange/Sharing of stationery, calculator etc. not allowed.

Text on the right of marks indicates course outcomes (only for faculty use).

Marks

Q1 A)	Explain omega network with an interconnection pattern. Show a perfect shuffle interconnection of omega network for eight inputs and outputs.	6	CO1
Q1 B)	Calculate Diameter, Bisection Width, Arc connectivity and Cost for following static network topologies: 1. Completely-connected 2. Star 3. 2-D mesh without wraparound 4. Hypercube	8	CO1
Q2 A)	What do you mean by a cluster? Draw cluster computer architecture and elaborate on some prominent components of cluster computer.	6	CO1
Q2 B)	How will you ensure that partitioning step while designing parallel algorithm has been carried out successfully?	6	CO2
Q2 C)	Show a cost-optimal way of computing the sum of 16 numbers using four processing elements. Calculate parallel runtime and prove how parallel system is cost optimal.	8	CO2
Q3 A)	Draw a complete state diagram of a simple three-state cache coherence protocol. Show example of parallel program execution.	6	CO1
Q3 B)	The serial runtime of an algorithm is $n^3$ and parallel run time of an algorithm is $n^3/p + n^2/\sqrt{p} + p \log p$ . What is the isoefficiency of the parallel algorithm? A Second parallel algorithm has a runtime of $(n^3 \log p)/p + n/\sqrt{p}$ . Which parallel algorithm Would you prefer? Why?	6	CO2
Q3 C)	What are different sources of parallel overhead? Derive the expression of isoefficiency function taking into account overhead term.	8	CO2
Q4 A)	How one to all broadcast and all to one reduction operations will be performed on an eight-node ring optimally? Show these operations diagrammatically. How many steps are required for such broadcast and reduction operation?	6	CO1
Q4 B)	What is OpenMP? Discuss OpenMP Programming model in detail.	6	CO1
Q4 C)	Write a vector addition program using OpenMP dynamic threading.	6	CO3
Q5 A)	What is CUDA? Explain CUDA Programming model and memory architecture with neat diagram.	6	CO3



**WALCHAND COLLEGE OF ENGINEERING, SANGLI.**  
(An Autonomous Institute)

ESE

**Final Year B.Tech. (Computer Science and Engineering)  
END SEMESTER EXAMINATION SEM. I NOVEMBER-2017  
PROFESSIONAL ELECTIVE V HIGH PERFORMANCE COMPUTING (3CS411)**

Exam Seat Number: \_\_\_\_\_

Day, Date and Time: Tuesday, 28/11/2017, 03.00pm to 05.00pm

Max Marks: **50**

**IMP: Verify that you have received question paper with correct course, code, branch etc.**

- Instructions:
- All questions are compulsory. Writing question number is compulsory. The answers may not be assessed if question number is not written.
  - Figures to the right of question text indicate full marks.
  - Assume suitable data wherever necessary, Write the answers with neat handwriting.
  - Only FX82 series non programmable Calculator is allowed.

Text on the right of marks indicates course outcomes (only for faculty use).

Marks

Q1	A)	What is data decomposition methodology? Explain the concept of partitioning of input, output and intermediate data with an example.	5	CO2
Q1	B)	What do you mean by superlinear speedup ? How it happens? Explain with example.	4	CO2
Q1	C)	Define following 1. Scalability of parallel systems 2. Amdahl's law	4	CO2

Q2	A)	How one to all broadcast and all to one reduction operations will be performed on an eight-node ring optimally? Show these operations diagrammatically. How many steps are required for such broadcast and reduction operation?	6	CO1
Q2	B)	Compare blocking and non-blocking message passing operations. How deadlocks occur in blocking operations? Explain with simple example.	5	CO1

Q3	A)	How will you perform matrix-vector multiplication using row-wise block 1-D partitioning? Calculate parallel runtime of this procedure. Derive the isoefficiency function for this parallel algorithm. (Note :- consider the case in which p processes are used such that $p \ll n$ ) --OR-- Prove that Cannon's algorithm is a memory-efficient version of the simple matrix multiplication algorithm. Show communication steps in Cannon's algorithm on 16 processes and derive its isoefficiency function.	8	CO3
Q3	B)	What is meant by bitonic sequence? Write a parallel algorithm for bitonic sort on a hypercube with $n=2^d$ processes.	4	CO3

Q4	A)	Write a CUDA program for vector addition considering following three cases. 1. Several blocks with one thread each 2. One block with several threads 3. Several blocks with several threads	6	CO3
Q4	B)	Define a minimum spanning tree. How will you parallelize Prim's algorithm ?	4	CO3
Q4	C)	Draw CUDA Memory Architecture and elaborate on all memory types.	4	CO1



**WALCHAND COLLEGE OF ENGINEERING, SANGLI.**  
*(An Autonomous Institute)*

MSE

Final Year B.Tech. (Computer Science and Engineering)  
MID SEMESTER EXAMINATION SEMESTER-I SEPTEMBER-2019  
PROFESSIONAL ELECTIVE V HIGH PERFORMANCE COMPUTING (3CS411)

Exam Seat Number: \_\_\_\_\_  
Day, Date and Time: Friday, 20/09/2019, 03.00pm to 04.30pm

Max Marks: 30

**IMP: Verify that you have received question paper with correct course, code, branch etc.**

- Instructions: i) All questions are compulsory. Writing question number is compulsory. The answers may not be assessed if question number is not written. Assume suitable data wherever necessary.  
ii) Figures to the right of question text indicate full marks.  
iii) Mobile phones and programmable calculators are strictly prohibited.  
iv) Except Exam Seat Number writing anything on question paper is not allowed.  
Exchange/Sharing of stationery, calculator etc. not allowed.

Text on the right of marks indicates course outcomes (only for faculty use).

		Marks
Q1	A) What do you mean by inter processor communication in parallel machine? Explain the shared memory conflicts in inter processor communication. How it can be solved?	4 CO1
Q1	B) Show the construction of hypercubes from hypercubes of lower dimension. Calculate the Diameter, Bisection Width, Arc Connectivity, Cost of hypercube-connected network.	4 CO1
Q2	A) <pre>#pragma omp parallel for private(i) for(int i=0; i&lt;100; i++) { A[i]=i; }</pre> How many iterations are executed per thread if four threads execute the above program?	2 CO1
Q2	B) State True or False <ol style="list-style-type: none"> <li>Code in an OpenMP program that is not covered by a pragma is executed by all the threads</li> <li>OpenMP is used for programming shared-memory systems</li> <li>Reduction operations are not parallelizable</li> <li>OpenMP uses compiler directives</li> <li>An omega network has <math>p/2 * \log p</math> switching nodes</li> <li>Caching is used to help overcome the memory bottleneck</li> </ol>	3 CO1
Q2	C) Let $l$ =latency and $b$ =bandwidth , $n$ =size . Which formula describes the message transmission time? i) $b+n/l$ ii) $b+l/n$ iii) $l+n/b$ iv) $l+b/n$ Also Derive the formula for store-and-Forward Routing.	5 CO1
Q3	A) Consider the problem of multiplying two $n * n$ matrices A and B to yield a matrix C. How this problem can be decomposed? Which type of decomposition is this? --OR-- Show the coarse-grained and fine-grained decomposition of dense matrix-vector multiplication.	4 CO2
Q3	B) How speculative decomposition is different from exploratory decomposition?	3 CO2
Q3	C) What is speedup and superlinear speedup? What are the causes behind achieving superlinear speedup?	5 CO2