

Parallel and Distributed Computing (CS3006)

Date: Feb 27th 2024

Course Instructors:

Dr. Abdul Qadeer

Dr. Rana Asif Rehman

Dr. Zeeshan Ali Khan

Mr. Danyal Farhat

Ms. Namra Absar

Sessional-I Exam

Total Time: 1 Hours

Total Marks: 35

Total Questions: 05

Semester: SP-2024

Campus: Lahore

Dept: Computer Science

Student Name

Roll No

Section

Student Signature

Vetted by

Vetter Signature

-
- Instruction/Notes**
- Attempt all questions on the provided space in the question paper.
 - Even if you do use rough sheets, they should NOT be attached with final paper.
 - If you find any ambiguity in a question, you can make your own assumption and answer the question accordingly by stating your assumption.
-

Question #	1	2	3	4	5	
Total Marks	8	5	6	8	8	35
Obtained Marks						
CLO #	1	1	1	3	1/3	

CLO # 1: Learn about parallel and distributed computers

Q1: Each MCQ is of 6 points. One or more choices can be right. Select more than one for such cases.

[8 marks]

i. How we can classify the type of computer architecture on a typical, consumer-grade multi-core laptop computer released in the year 2024 out of the following?

- A) SISD
- B) SIMD
- C) MIMD
- D) MISD

Correct answer: (A), (B), and (C) because if we just use a single code it will be SISD, if we use all cores concurrently, it will be MIMD. And all modern computers have vector instructions, hence SIMD.

ii. Karp-Flatt metric can give us information for which of the following?

- A) Serial code in terms of number of lines-of-code
- B) Time proportion spent executing serial parts of the code
- C) Any parallelization overheads
- D) Number of processors used for parallelization

Correct answer: (B) and (C)

iii. A network with high connectivity is desirable, because it:

- A) lowers the contention for communication resources.
- B) decreases the energy consumption.
- C) increases the contention for communication resources.
- D) All of the above

Correct answer: (A)

iv. Interpretation of Moore's law assert that:

- A) Computing power will eventually level off
- B) Computing power doubles every 18 months
- C) Transistors decrease in size 50% every 2 years
- D) Data storage costs decrease by 50% every 18 months

Correct answer: (B)

v. Different computers possess the same version of operating system when connected in the form of:

- A) Cluster
- B) Cloud
- C) Network of workstations
- D) None of the given option

Correct answer: (A)

vi. In Bus network, the number of shared connections is equal to:

- A) 1
- B) n
- C) $\log_2 n$
- D) n^2

Correct answer: (B)

vii. Which one is true about Amdahl's law formulation:

- A) It does not account for communication overheads
- B) It provides an upper-bound on the achievable speedups
- C) If the number of processors approaches infinity, Speedup is bounded by $\frac{1}{\text{sequential fraction}}$
- D) All (a), (b), and (c) are correct
- E) Only (b) and (c) are correct

Correct answer: (D)

viii. Choose correct statement/s regarding NUMA architecture:

- A) Access time may vary depending on the data location
- B) Different processors interact with each other using message passing
- C) The same logical address on different processors maps to the same physical memory location
- D) The same logical address on different processors maps to the different physical memory locations

Correct answer: (A) and (C)

CLO # 1: Learn about parallel and distributed computers

Q2: Consider the following equation in the context of Karp-Flatt metric.

[1+1+1+2 marks]

$$f = \frac{1/s - 1/p}{1 - 1/p}$$

(a) If p is the number of processors in the above equation, what is s?

s is speed-up achieved.

(b) What is f in the above equation?

f is the Karp-Flatt metric.

(c) What are the units of f?

It's a unitless entity being a time fraction.

National University of Computer and Emerging Sciences

(d) Following data was gathered experimentally on some parallel computer A.

p	Time (seconds)	s	f
1	2.17	---	---
2	1.11	1.95	0.024
3	0.754	2.88	0.021
4	0.577	3.76	0.021
8	0.312	6.96	0.021

What can we infer about the quality of parallelism by looking at the data provided? Please point to specific data to support your answer.

- (1) We are getting speed-up with increasing number of processing elements.
- (2) Karp-Flatt metric is almost the same across runs with increasing number of processing elements. That implies that there might not be any overheads for parallelization though for further improvement we need to target the serial proportion. There is limited amount of parallelism available in the problem.

CLO # 1: Learn about parallel and distributed computers

Q3: Derive generalized expressions for arc-connectivity, diameter, bisection-width, and link cost of a 2D mesh (without wraparound) having M rows and N columns. **[1.5+1.5+1.5+1.5 marks]**

Solution:

(i) \hookrightarrow Arc-Connectivity = 2

(ii) Diameter = Row distance + Column distance
[diameter] [diameter]
 $= (N-1) + (M-1) \Rightarrow M+N-2$

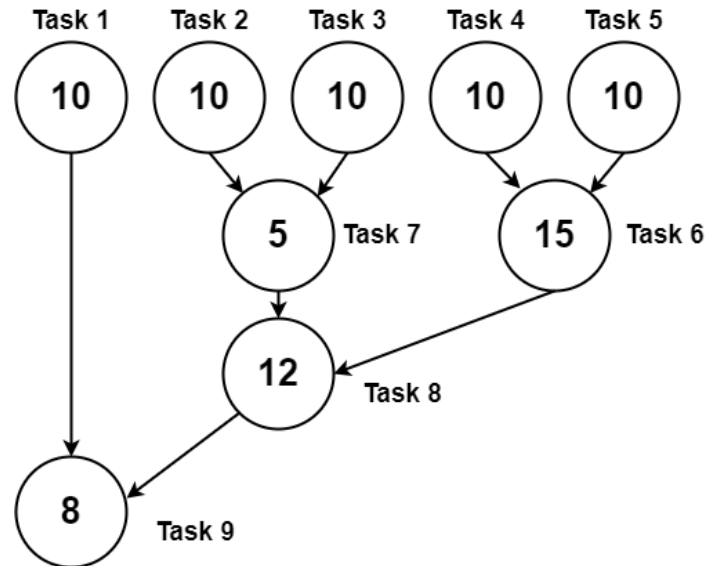
(iii) Bisection width = Min of (M, N)
 \hookrightarrow Minimum number of links that must be removed to partition in equal halves

(iv) Cost = (Links in a row) (Total rows) + (Links in a column) (Total columns)
 $= (N-1)(M) + (M-1)(N)$
 $= MN - M + MN - N$
 $= 2MN - (M+N)$

CLO # 3 - Analytical modeling and performance evaluation of parallel programs

Q4: Consider the following task-dependency graph:

[6+2 marks]



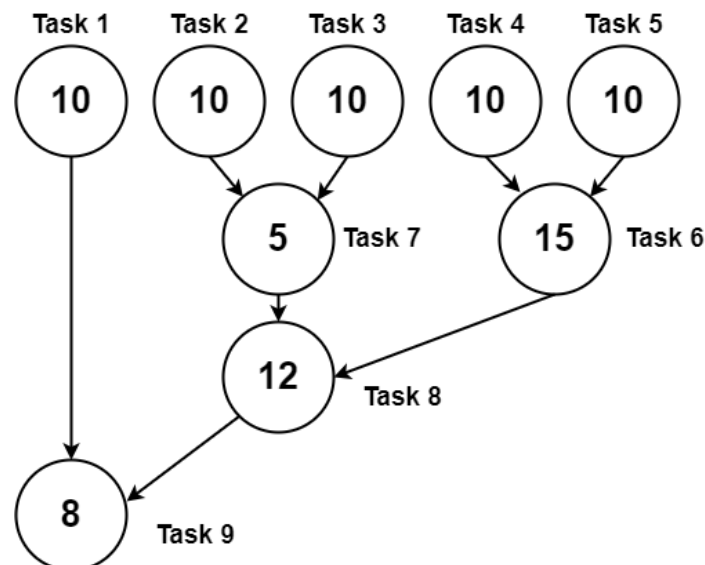
- (a) What is the Maximum degree of concurrency, Critical Path Length, and Average Degree of Concurrency of the task graph?

Maximum degree of concurrency = 5

Critical Path Length = $10 + 15 + 12 + 8 = 45$

Average degree of concurrency = $(10 + 10 + 10 + 10 + 10 + 5 + 15 + 12 + 8)/45 = 90/45 = 2.0$

- (b) Map the task dependency graph on five processes namely, Pa, Pb, Pc, Pd and Pe, to reduce the idling time for the task processing. Please answer using the graph provided below and add a short explanation.



One of the possible solutions is as follows:

Map Tasks 1 and 9 on Pa.

Map Tasks 2 and 8 on Pb.

Map Tasks 3 and 7 on Pc.

Map Tasks 4 and 6 on Pd.

Map Task 5 on Pe.

The mapping will be able to reduce the idling time in the task processing using the maximum degree of concurrency of the task graph.

CLO # 3 - Analytical modeling and performance evaluation of parallel programs

Q5 [For all sections except section BDS-6B]: Recall how merge sort work. We have provided its pseudocode to jog your memory.

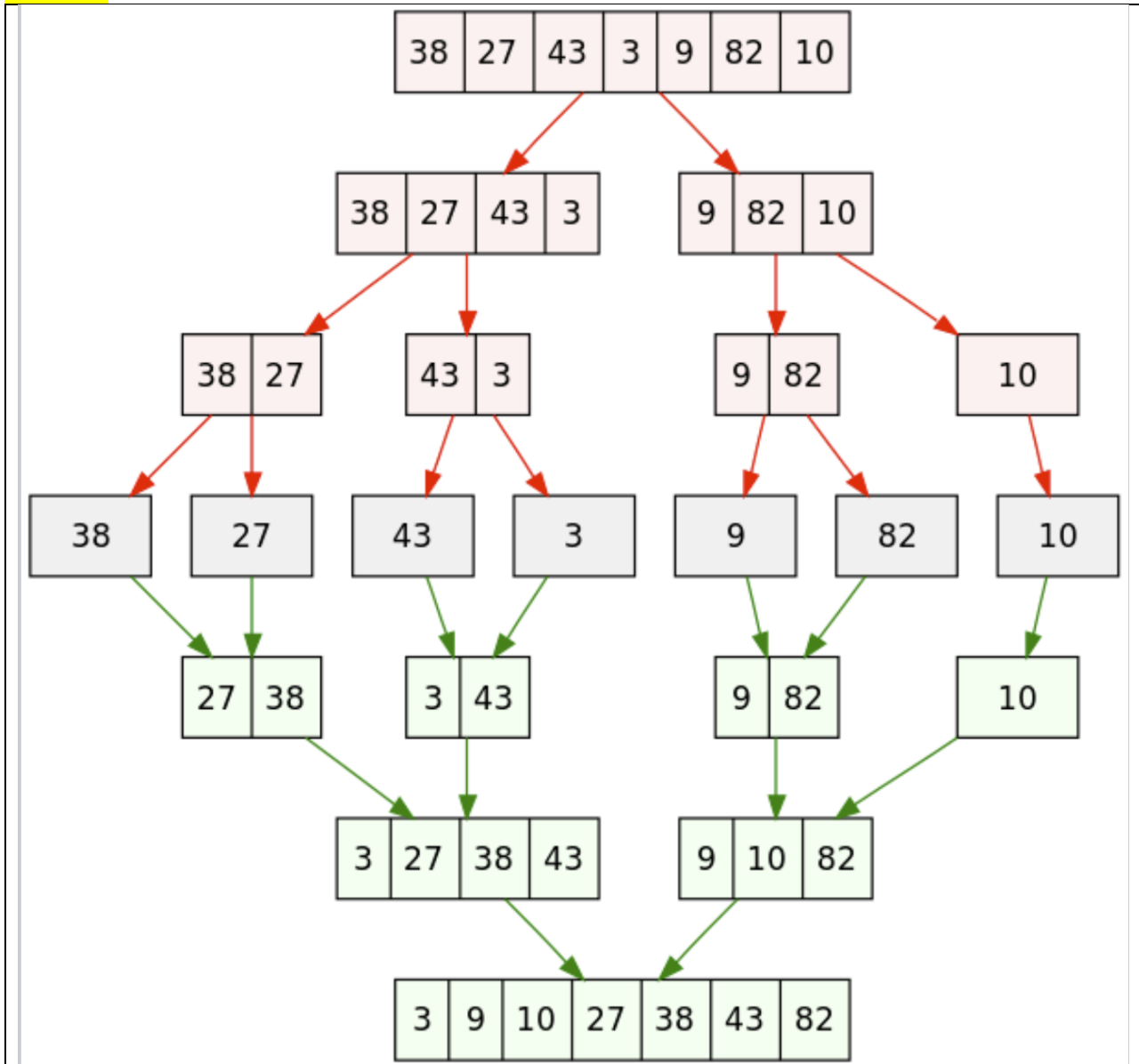
[2+2+2+2 marks]

```
MERGE-SORT( $A, p, r$ )
1  if  $p < r$ 
2     $q = \lfloor (p + r) / 2 \rfloor$ 
3    MERGE-SORT( $A, p, q$ )
4    MERGE-SORT( $A, q + 1, r$ )
5    MERGE( $A, p, q, r$ )
```

(a) Using the following example data provide a recursive decomposition.

	38	27	43	3	9	82	10	
--	----	----	----	---	---	----	----	--

Solution:



(b) Using the same data as above provide a data-based decomposition.

Divide the data based on the input and equally divide reasonable (but almost equal sized) data to each task. In this example, data is small we can just use two tasks. One will do [38,27,43,3], and the other will do [9,82,10]. They will independently run merge sort and will finally combine the result.

(c) In terms of potential for better concurrency (and better speedup) which one of the above two are a better decomposition approach? (Just name it.)

Data-based decomposition and then combine recursively.

(d) Provide rationale for your above answer logically and succinctly.

For our small example, there seems to be a lot of interactions among tasks in the case of recursive decomposition as the merging starts (for example to bring data from another processor and to merge it, which is a serial work). On the contrary, the data based decomposition initially divides data for processors and then they don't have to interact until they have sorted their part. Since data size is almost same, they will finish about the same time. We could use parallel tasks for merging (for example if we had 4 sorted pieces, we could use 2 processes to merge, and finally one processor to merge).

CLO # 1: Learn about parallel and distributed computers

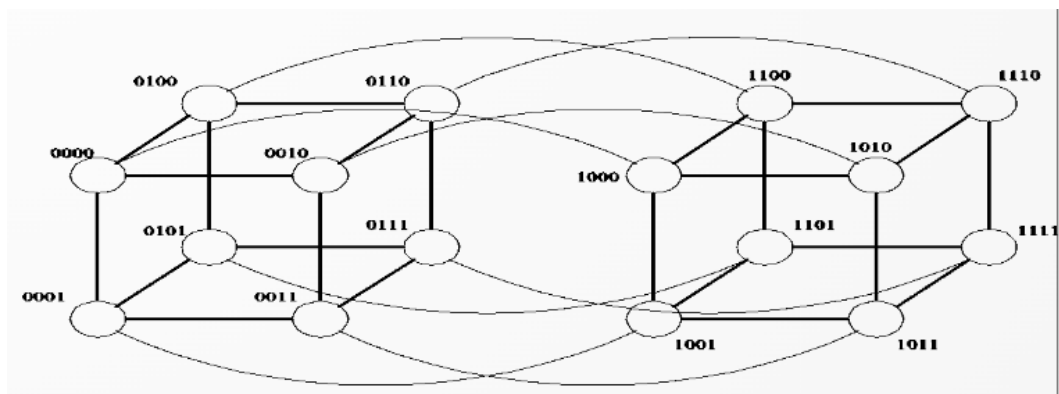
Q5 [For section BDS-6B ONLY]:

(a) Derive the total communication cost for a message of size w words to traverse d communication links using cut-through routing. (Describe each parameter involved.)

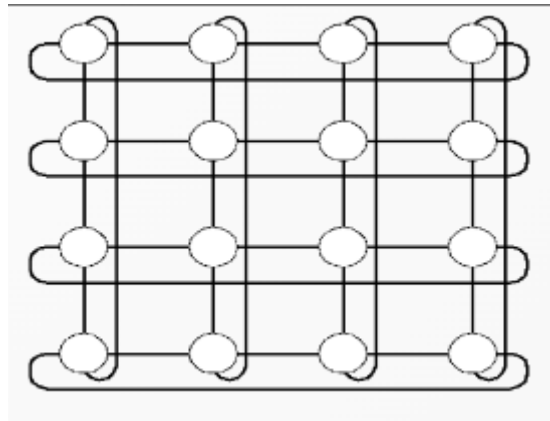
[4 marks]

$$T_{\text{comm}} = T_s + wT_{\text{word}} + dT_h$$

(b) Consider the following two graphs



Graph 1



Graph 2

Compute and compare the diameter and arc connectivity of the given graphs. What does it tell about their performance? **[2+2 marks]**

Diameter of graph1 = $\log p = \log 8 = 3$

Diameter of graph 2 = $2(\sqrt{p}/2) = 4$

arc connectivity of graph 1 = $\log p = 3$

arc connectivity of graph 2 = 4

Graph 1 is more prone to link failure. In case of any failure, there are fewer other routes to the node.