Homework 2
CS 118: Cloud Computing
Fall '22

Assigned: 09/29/22
Due: 10/13/22 at 11:59PM Medford Time
Please submit on Gradescope in PDF format

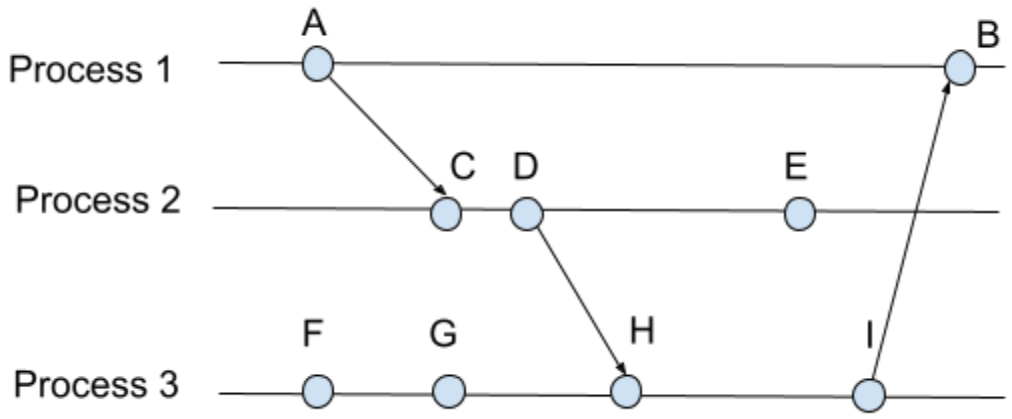
Distributed Systems

1 - [10 pts] Scatter/aggregate patterns & tail latency. Google's web search distributed service uses a scatter/aggregate communication pattern. The distributed service has a fan out of 1520 processes and these processes have a 99.9% latency of 1ms.

(a) What is the probability that a search query will take more than 1ms?

(b) Name one technique to reduce tail latency. How could it help? (General description is good enough; no need to show math.)

2 - [20 pts] Logical clocks.

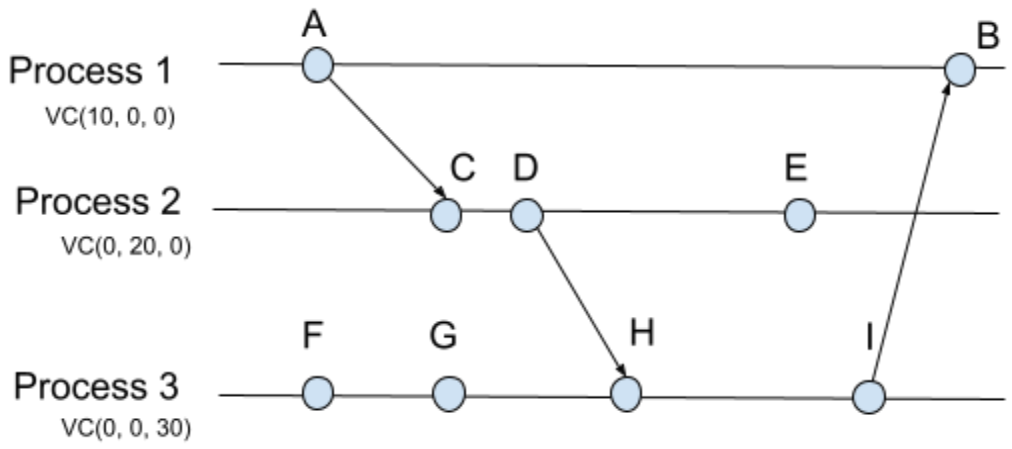


(a) State the definition of Lamport's happens-before relationship.

(b) Which events among the above processes are deemed concurrent? (I.e., are not linked by happens-before relationships?)

(c) Logical clocks are one simple way to capture the happens-before relationship (there are others). Fill the diagram in with the logical clock values that would be assigned to the events. Assume that Process 1 starts with logical clock value 0, process 2 starts with logical clock value 10, and process 3 starts with logical clock value 30.

(d) Vector clocks are another way to capture the happens-before relationship. Fill the diagram in with the vector clock values that would be assigned to the events. Assume that Process 1 starts with vector clock value (10, 0, 0), process 2 starts with vector clock value (0, 20, 0), and process 3 starts with vector clock value (0, 0, 30).



3 - [10 pts] GPS vs. happens before: Modern computers can associate GPS timestamps with individual events, solving the synchronized clock problem that logical clocks also solve. Why is the happens-before relationship still useful?