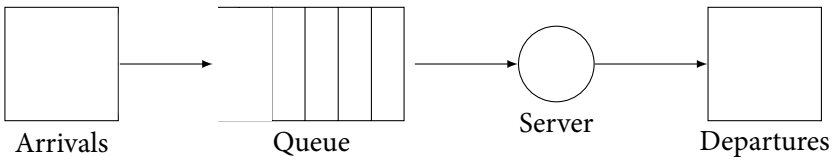
Lesson 2. Performance Measures for Queueing Systems

1 A small example

We motivate our consideration of performance measures with the following system with a first-in-first-out (FIFO) queue and a single server.



Suppose that we simulate the system with 4 customers, starting at time 0. We obtain the time measurements as shown below.

Customer	Arrival time	Time leaves queue	Departure time
1	0	0	8
2	4	8	10
3	6	10	12
4	15	15	20

2 Customer-level performance measures: delay and waiting time

Recall the definitions of delay and waiting time from SA402:

- **Delay** is the time a customer spends in the queue.
- **Waiting time** is the time a customer spends in the system from arrival to departure.

Calculate the delay and waiting time for the four customers as well as the **average delay** and **average waiting time** in the table below.

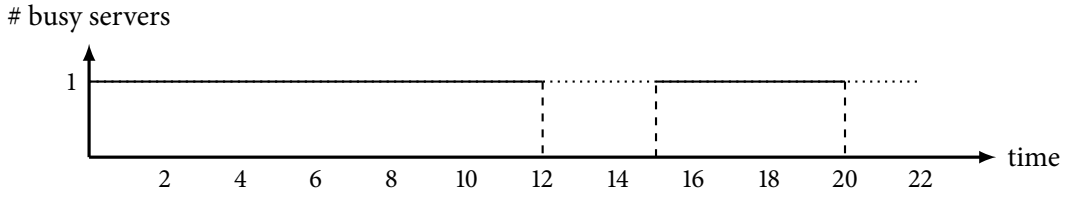
Customer	Delay	Waiting Time
1		
2		
3		
4		
Average		

3 System-level performance measures: busy servers and customers in queue

System-level performance measures describe how the overall system is performing. We define utilization and time-average number of customers in queue over a given **time horizon** T.

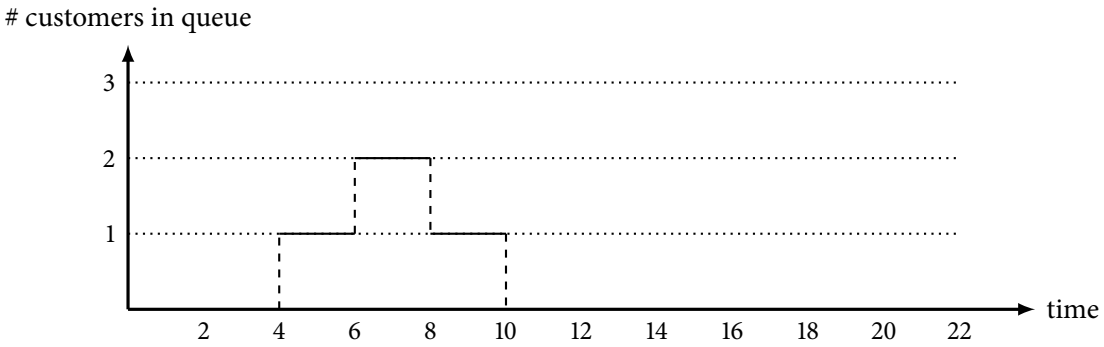
- The **time-average number of busy servers** is the integral of the number of busy servers from 0 to T, divided by T.
  - This is a weighted average of the number of busy servers. Each number is weighted by the length of time that number of servers are busy.
  - When there is one server, this is the proportion of the time horizon that the server is busy.
- The **time-average number of customers in queue** is the integral of the number of customers in queue from 0 to T, divided by T.
  - Intuitively, this is a weighted average of the number of customers in the queue. Each number is weighted by the length of time the queue contains that number.

Here is a graph depicting the busy times of the server:



Calculate the time-average number of busy servers from this graph using $T = 20$.

Now consider the following graph depicting the number of customers in queue:



Calculate the time-average number of customers in queue from the graph using $T = 20$.