How to scale message consumption

single consumer

messages are processed in order

unreliable

multiple consumers

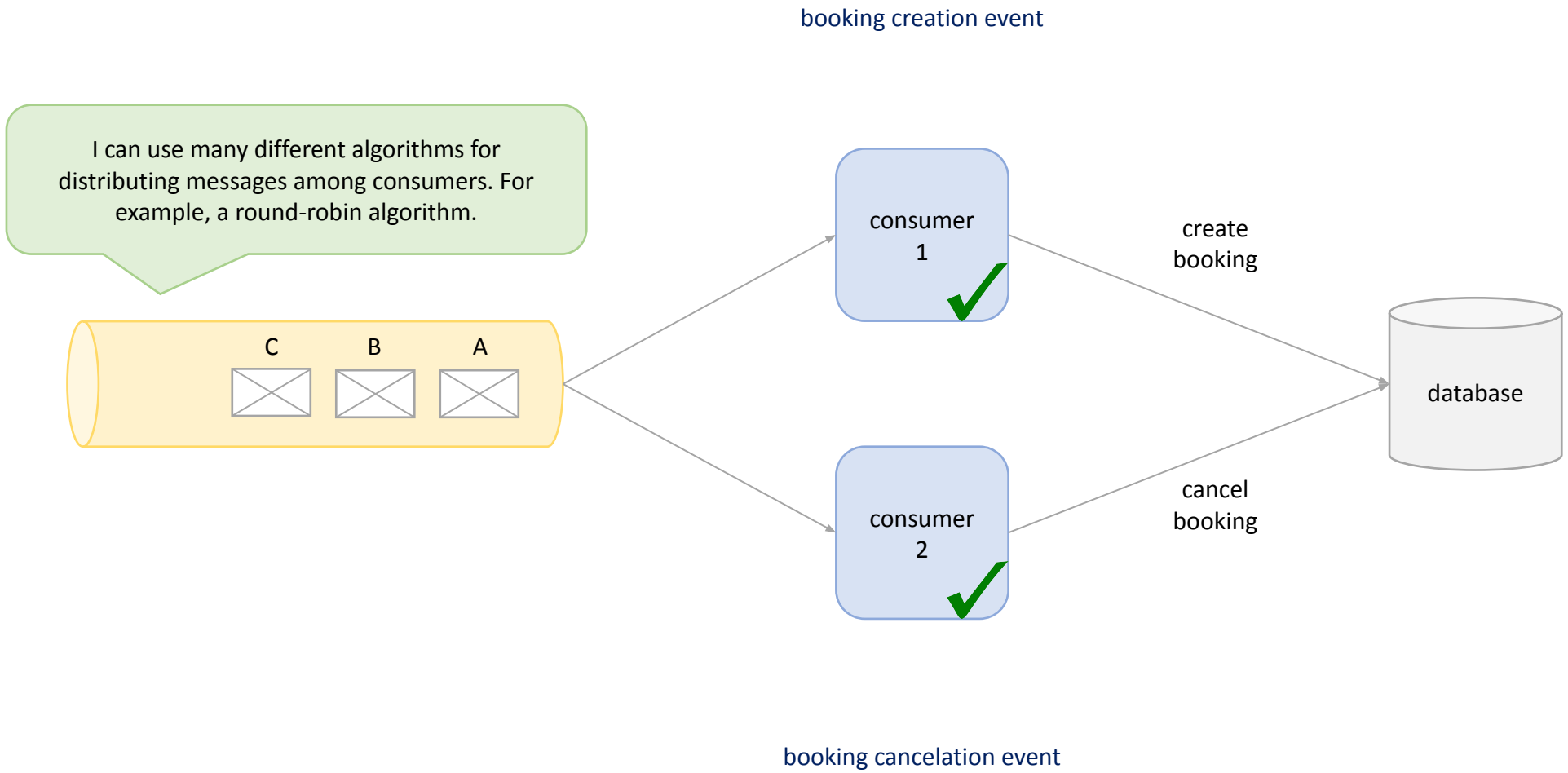
high availability

messages can be processed out of order

higher chance of processing the same message multiple times

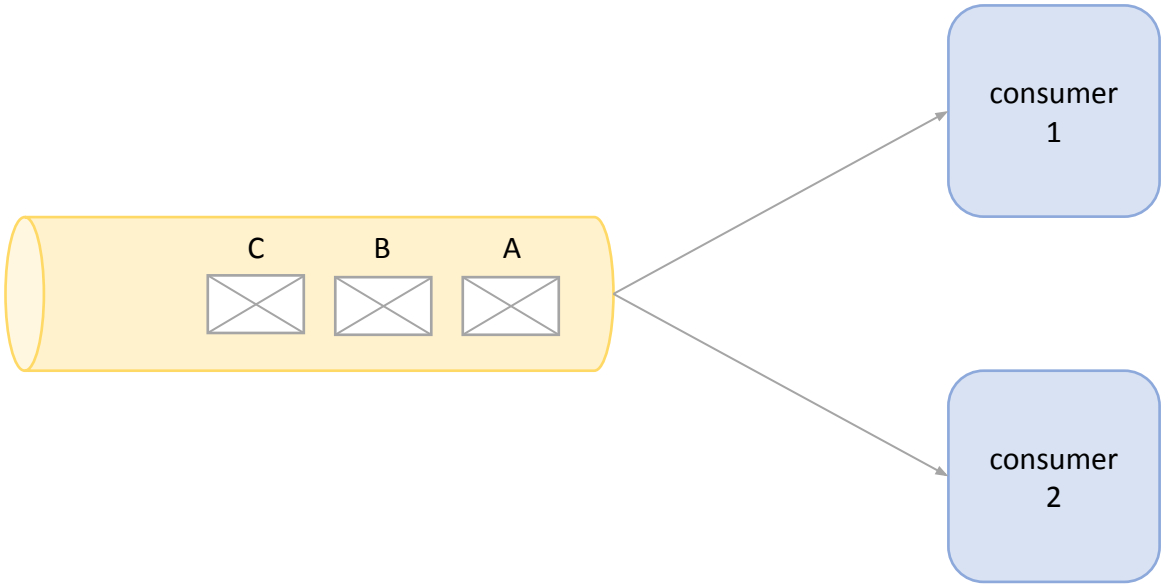
How to scale message consumption

competing consumers break the order of message processing



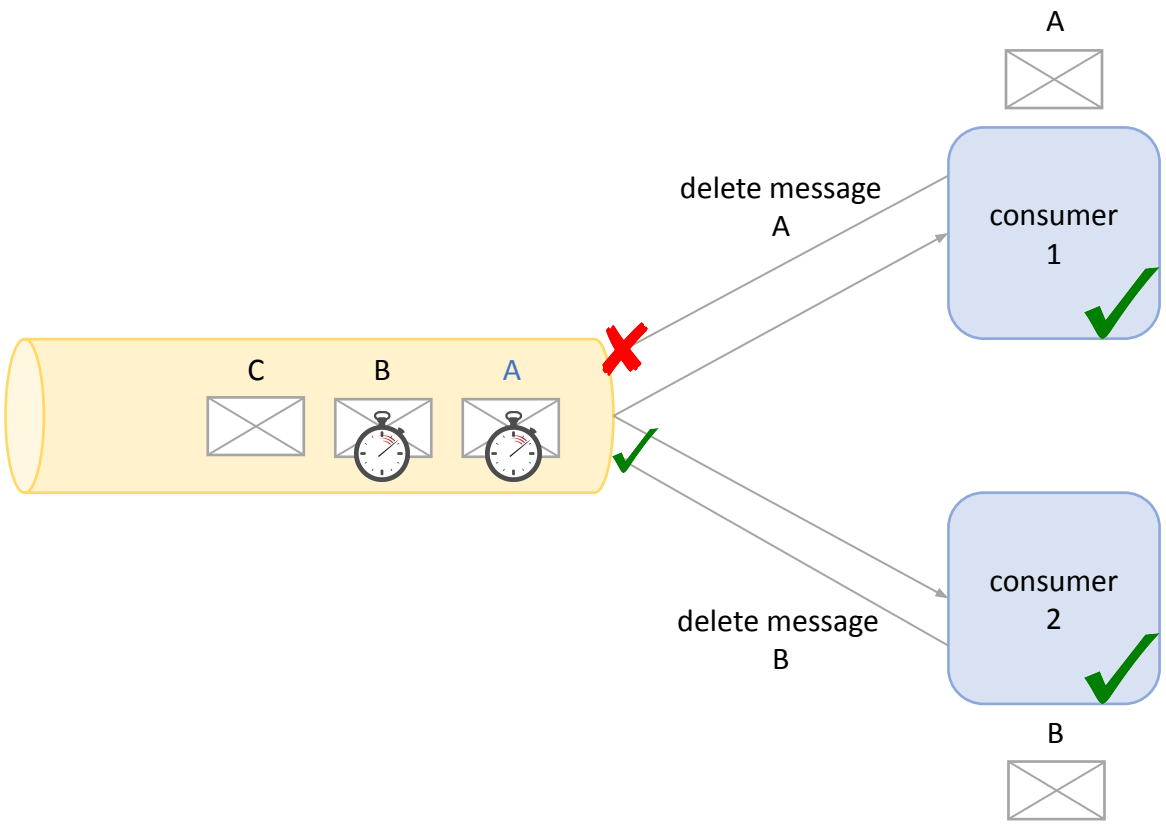
How to scale message consumption

competing consumers increase the chances of processing the same message multiple times



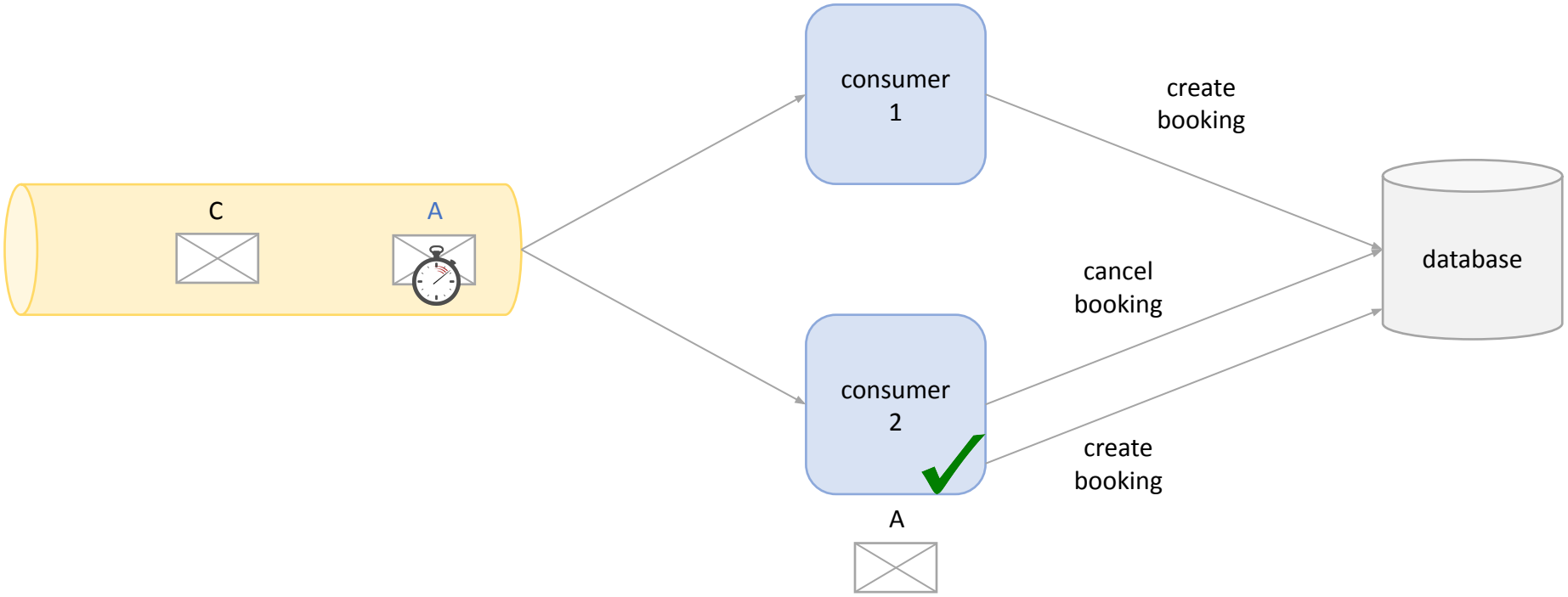
How to scale message consumption

competing consumers increase the chances of processing the same message multiple times



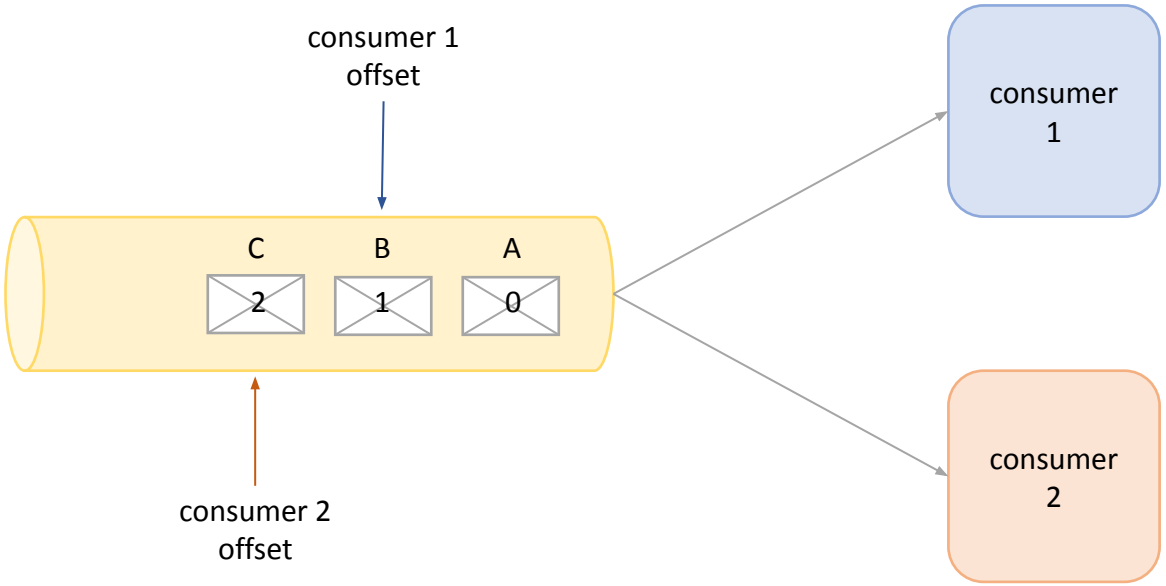
How to scale message consumption

competing consumers increase the chances of processing the same message multiple times



How to scale message consumption

multiple consumers in a log-based messaging system



How to scale message consumption

to process messages in parallel we need multiple consumers

but

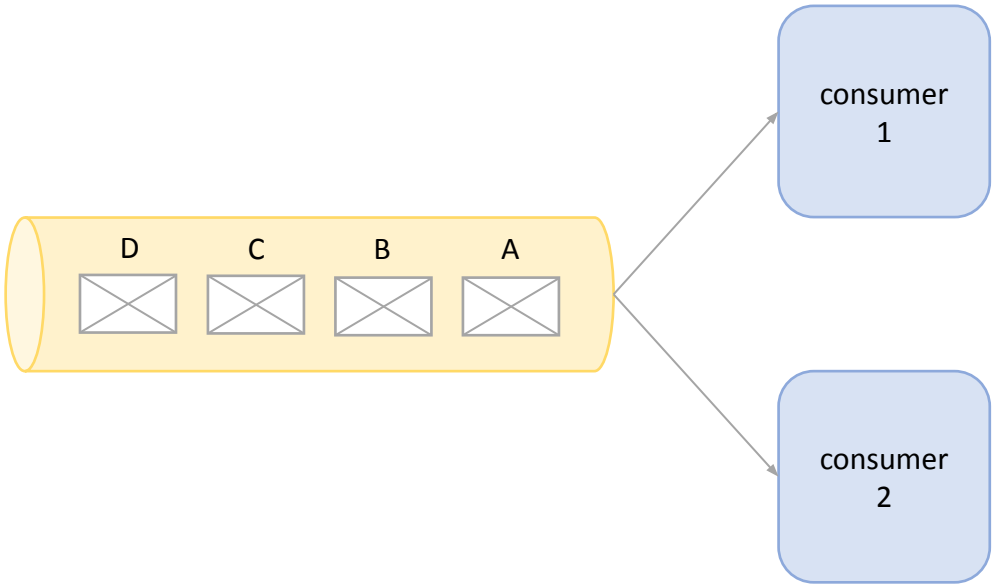
- order of message processing is lost
- this can increase the number of duplicates in the system
- log-based messaging systems do not even support having competing consumers on a single queue

solution

data partitioning (sharding)

How to scale message consumption

instead of this



let's have this