

Quiz - Observability

- Due Nov 25 at 12:05pm
- Points 6
- Questions 6
- Available Nov 25 at 11:40am - Nov 25 at 12:05pm 25 minutes
- Time Limit 20 Minutes

This quiz was locked Nov 25 at 12:05pm.

Attempt History

| | Attempt | Time | Score |
|--------|---------------------------|------------|------------|
| LATEST | Attempt 1 | 19 minutes | 6 out of 6 |

 Correct answers are hidden.

Score for this quiz: 6 out of 6

Submitted Nov 25 at 11:59am

This attempt took 19 minutes.



Question 1

1 / 1 pts

A distributed service is partially degraded for some users but continues serving others. The most appropriate reliability measure here is:

- Ratio of successful to total requests
- Percentage of servers online
- Mean time between failures
- Server uptime



Question 2

1 / 1 pts

Your batch processing system runs 100,000 jobs daily. To meet a **99.99% availability** SLO, what's the maximum allowed number of failed jobs per day?

- 100
- 10
- 1
- 5



Question 3

1 / 1 pts

You propose improving a system from 99.9% to 99.99% availability. The additional uptime yields only \$900 in value annually, but the upgrade costs \$3,000. What should the SRE team conclude?

- Proceed; costs are irrelevant to reliability
- Do not proceed; marginal gains do not justify cost
- Recalculate using mean downtime instead
- Proceed; reliability should always improve



Question 4

1 / 1 pts

Why is the Median (p50) often a better performance indicator than the mean for latency SLIs?

- Ignores outliers
- captures only the slowest requests
- exaggerates performance differences
- reflects the typical user experience without distortion from extreme values



Question 5

1 / 1 pts

The graph above shows latency measurements for a web service at three percentiles — **p50**, **p90**, and **p99** — over time.

- Explain what each percentile (p50, p90, p99) represents in terms of user experience.
- Analyze the graph to describe what is happening to latency across these percentiles.

Your Answer:

1) In terms of user experience,

- p50 is median; Half of the users experience latency below this line
- p90 means 90% of users are faster than this line, while remaining 10% are slower. These users aren't rare but are just slower than most
- p99 is the experience of the slowest 1% of users; these are the really slow requests

2) Analyzing,

- p50 is fairly stable around ~250–350 ms with small fluctuations
- p90 shows larger swings in the 450–700 ms range, indicating that the slower 10% of users see noticeably more variable performance
- p99 is highly volatile, spiking repeatedly, especially after 12:20, meaning that the slowest 1% of requests sometimes become extremely slow



Question 6

1 / 1 pts

1. Describe the trend in ride requests across time and explain what the sharp dip and recovery might indicate.
2. Which metric(s) would you correlate this with to confirm a system-level incident?

Your Answer:

- 1) Around 8:40 PM, ride request counts (ongoing, completed, upcoming) stay high and fairly stable. Then after 8:40PM suddenly all lines drop sharply to very low levels, this sharp drop could be because of system failure for eg backend outages, API failures etc. Then again rise back up to normal high levels possibly because the system has been restored/recovered, so normal request volumes returned.
- 2) To validate this was an outage we can correlate with metrics like API error rate, traffic drop, database errors, load balancer error or connection issue.

Quiz Score: 6 out of 6