Log In   **Join**

Back To Course Home

Grokking Modern System Design Interview for Engineers & Managers
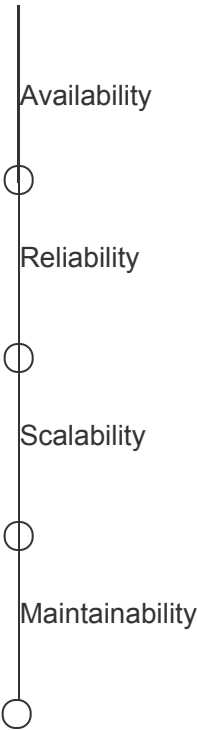0% completed

System Design Interviews

Introduction

Abstractions

Non-functional System Characteristics

Availability

Reliability

Scalability

Maintainability

Fault Tolerance

Back-of-the-envelope Calculations

Building Blocks

Domain Name System

Load Balancers

Databases

Key-value Store

Content Delivery Network (CDN)

Sequencer

Distributed Monitoring

Monitor Server-side Errors

Monitor Client-side Errors

Distributed Cache

Distributed Messaging Queue

Pub-sub

Rate Limiter

Blob Store

Distributed Search

Distributed Logging

Distributed Task Scheduler

Sharded Counters

Concluding the Building Blocks Discussion

Design YouTube

Design Quora

Design Google Maps

Design a Proximity Service / Yelp

Design Uber

Design Twitter

Design Newsfeed System

Design Instagram

Design a URL Shortening Service / TinyURL

Design a Web Crawler

Design WhatsApp

Design Typeahead Suggestion

Design a Collaborative Document Editing Service / Google Docs

Spectacular Failures

Concluding Remarks

Course Certificate

Mark Course as Completed

Availability

Learn about availability, how to measure it, and its importance.

We'll cover the following

- What is availability?
  - Measuring availability
  - Availability and service providers

What is availability?#

Availability is the percentage of time that some service or infrastructure is accessible to clients and is operated upon under normal conditions. For example, if a service has 100% availability, it means that the said service functions and responds as intended (operates normally) all the time.

Measuring availability#

Mathematically, availability, A, is a ratio. The higher the A value, the better. We can also write this up as a formula:

A (in percent) = $\frac{(Total\ Time\ -Amount\ Of\ Time\ Service\ Was\ Down)}{Total\ Time} * 100$

We measure availability as a number of nines. The following table shows how much downtime is permitted when we're using a given number of nines.

The Nines of Availability

Availability Percentages versus Service Downtime			
Availability %	**Downtime per Year**	**Downtime per Month**	**Downtime per Week**
90% (1 nine)	36.5 days	72 hours	16.8 hours
99% (2 nines)	3.65 days	7.20 hours	1.68 hours
99.5% (2 nines)	1.83 days	3.60 hours	50.4 minutes
99.9% (3 nines)	8.76 hours	43.8 minutes	10.1 minutes
99.99% (4 nines)	52.56 minutes	4.32 minutes	1.01 minutes
99.999% (5 nines)	5.26 minutes	25.9 seconds	6.05 seconds
99.9999% (6 nines)	31.5 seconds	2.59 seconds	0.605 seconds
99.99999% (7 nines)	3.15 seconds	0.259 seconds	0.0605 seconds

Availability and service providers#

Each service provider may start measuring availability at different points in time. Some cloud providers start measuring it when they first offer the service, while some measure it for specific clients when they start using the service. Some providers might not reduce their reported availability numbers if their service was not down for all the clients. The planned downtimes are excluded. Downtime due to cyberattacks might not be incorporated into the calculation of availability. Therefore, we should carefully understand how a specific provider calculates their availability numbers.

Back

The Spectrum of Failure Models

Next

Reliability

Mark as Completed

Report an Issue