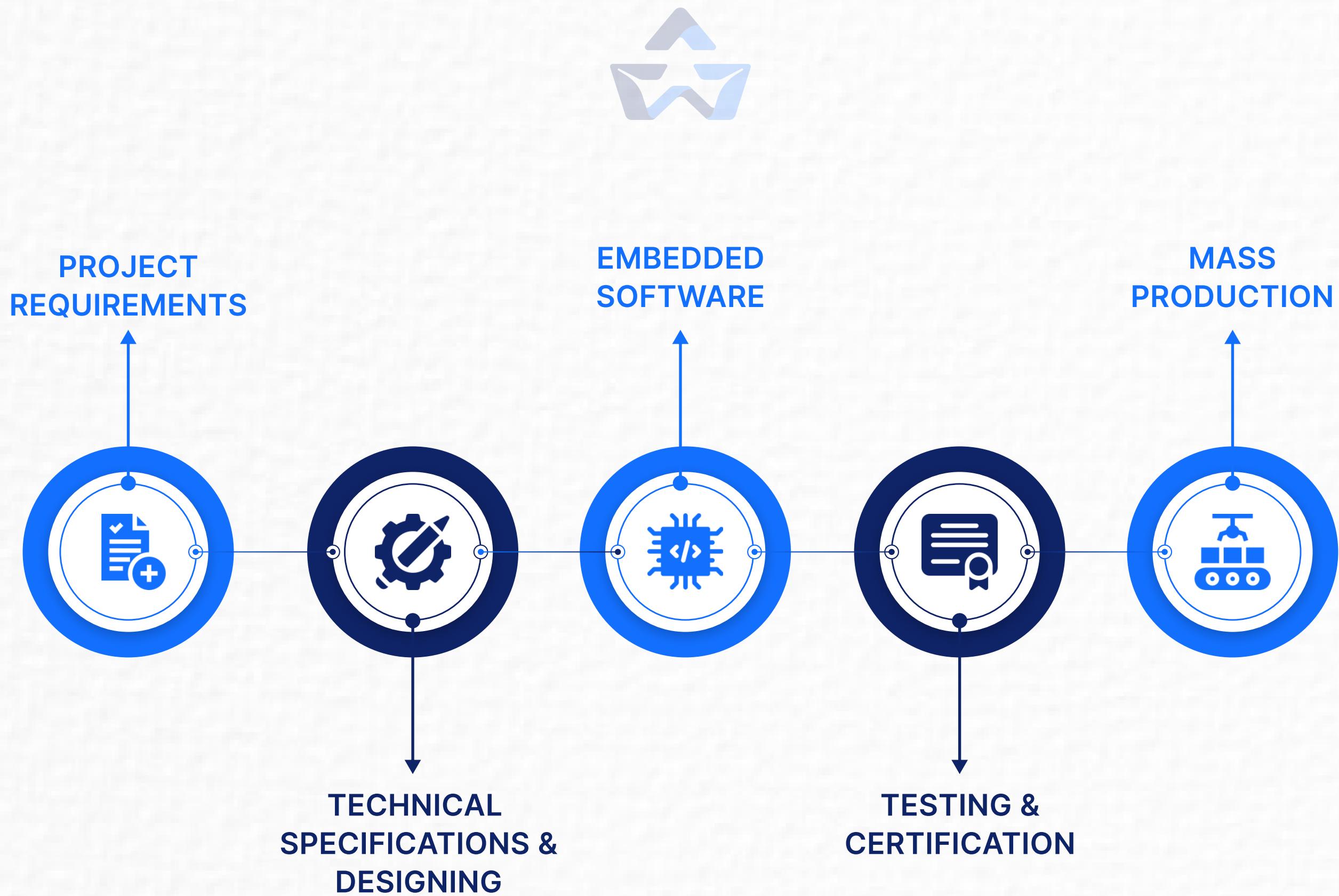


TOP 25

SYSTEM DESIGN

INTERVIEW QUESTIONS



For Data Engineering role



Disclaimer

Understanding System design for data engineering can be tough, especially for working professionals.

This document is here to guide you through the key topics you need to cover if you're already in the transition phase.

It will be a valuable resource to help you navigate your journey effectively.

QUESTION-1

How can you design a URL shortening service like Bitly?

- URL shortening services create special brief links that redirect users as they click on those links.

Steps:

- Short hash generation occurs for the initial URL.
- The matching information must be stored within a database.
- The system routes users toward the short URL destination when they interact with it

Diagram:

User → Short URL Service → Database (Store URL Mapping)
→ Redirect to Original URL

QUESTION-2

What's the approach to building a recommendation system?

- Recommendation system can work based on collaborative filtering where recommendation depends on user profile or based on content based filtering where recommendation depends on similarity between the items. It makes customized recommendations through the analysis of the consumer's behaviour and choices.

QUESTION-3

What is an efficient way to implement a caching layer?

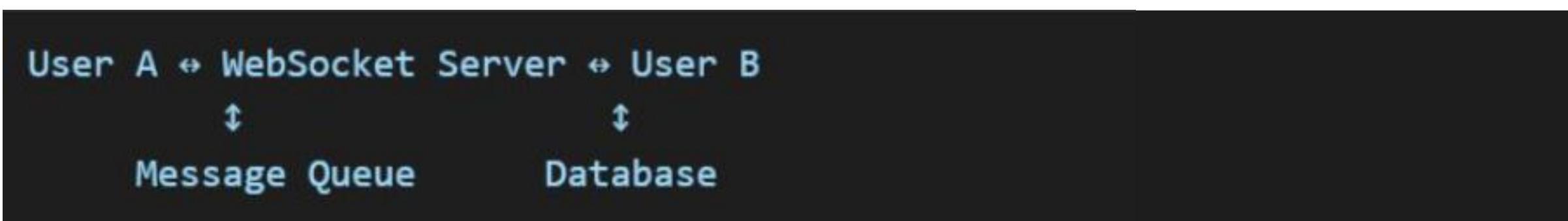
- An application cache such as Redis can be used successfully to cache commonly used data to free the database. They eliminate the issue of large pool of data and or large cache by ensuring that eviction policies (such as LRU) are employed to ensure fast access to important data and consequently enhance system performance.

QUESTION-4

How would you create a real-time chat application?

- For actual time sending and receiving of information, employ WebSockets to create the persistent connections. For scalability and handling of large messages adopt message queues and store your messages into a database (NoSQL basically).

Diagram:



QUESTION-5

How can rate limiting be implemented?

- The techniques for rate limiting are **token bucket** or **leaky bucket** algorithms. Do this to count the number of requests per user then deny access to users who have requested more than the standard allowed to utilize resources.

QUESTION-6

How would you build a job scheduling system?

- A job scheduling system involves placing jobs and then coordinating with workers to get the jobs done. When jobs fail, there should be an ability to handle priorities, and attempt to retry. If the tasks may be time consuming use a database or a message queue to schedule and coordinate the jobs.

Diagram:

```
User → Job Queue → Workers → Task Execution
```

QUESTION-7

What's the best way to design a file storage system?

- A distributed file storage system like, HDFS or S3 splits files to individual blocks, stores a block on different nodes and provides redundancy to ensure failure is handled well. This assists in dealing with large amount of data.

QUESTION-8

How do you handle database sharding?

- Partition the data in the database by sorting the data in a particular way, for instance user ID. Use a shard router to get the right requests from the partitioned database to prevent a bottleneck and to make the application scale better.

QUESTION-9

What considerations would you take for an online voting system?

- Online voting requires measures such as secure socket layer code for encoding, authentication of user and lastly, ensuring that a single user can only vote once. Use of a database to store votes and a system that will reduce the possibility of fraud.

QUESTION-10

How do you design an e-commerce checkout system?

- The checkout process divides into three sections which contain cart administration along with payment processing and inventory system updates. The system integrates real-time stock visibility functions while providing order confirmation processes which maintains smooth payment transaction flow and accurate inventory management.

QUESTION-11

How would you create a notification system?

- User notifications are managed through queues while the push service operates from Firebase. The system adjusts notifications through custom filters derived from both individual choice sets and machine availability capabilities.

QUESTION-12

What steps are involved in designing a search engine?

- Search engines operate through content collection using crawling after inverting the collected data into their index system for efficient query execution. Ranking algorithms will organize retrieved results according to relevance once a query is processed.

QUESTION-13

How would you handle session management?

- Cookies contain session IDs while session data remains contained within a database or cache system. Log out sessions will expire automatically alongside database invalidation to secure user data as well as minimize server load and processing needs.

QUESTION-14

What's the approach to building a social media feed?

- A database stores user posts with timestamp information which feeds chronological posts to users in reverse order. The system arranges posts through algorithms that utilize user interactions (likes and shares) while delivering personalized content.

Diagram:

```
User Activity → Feed Generator → Personalized Feed
```

QUESTION-15

How would you design an image upload and sharing system?

- Scalable file storage should be in cloud storage (such as S3) while image compression before upload leads to quicker load times. Share access management together with thumbnail previews functions through implementation.

QUESTION-16

How can you implement an authentication system?

- The system uses JWT tokens for authentication which create tokens after successful logins and safely stores them. While logging into the system users must enable multi-factor authentication (MFA) mechanisms to provide improved security.

QUESTION-17

What's the best way to design a video streaming platform?

- Adaptive bitrate streaming protocols (HLS as an example) will modify video quality automatically based on the current bandwidth. When stored across a CDN in chunks users get fast content delivery which leads to smooth video playback during high traffic periods.

QUESTION-18

How can you build a URL redirect service?

- URL redirect services maintain a database which redirects users from short URLs to their destination URLs. A visitor who clicks on a short URL triggers the system to check for the URL mapping information before redirecting them to the corresponding full URL.

QUESTION-19

How would you design an inventory management system?

- A real-time database system should track warehouse inventory while maintaining synchronized data between all facilities. Users should receive stock level alerts while the system should show live information about product availability.

QUESTION-20

How do you design a content management system (CMS)?

- A content management system (CMS) needs to provide users with an easy interface to create content and make changes and release those updates. A secure and scalable content management system requires version control in combination with user roles and permissions implementation.

QUESTION-21

How would you implement logging and monitoring?

- Place service logging data into one centralized system through implementation of the ELK Stack technology. You should create dashboards that show real-time system performance with quick error detection through alert notifications.

QUESTION-22

What's the process for designing an online file-sharing platform?

- Cloud storage provides scalable secure file storage which also supports versioning implementation. People can share data through secure connections while administrators use permission systems to control system access and maintain straightforward upload and download procedures.

QUESTION-23

How would you design a multi-tenant SaaS application?

- Choose between a database structure containing tenant IDs for shared data storage or maintain individual database instances for each tenant. Users should get customized features including personalized branding alongside settings tailored to each specific tenant.

QUESTION-24

How would you design an email delivery service?

- Use SMTP for sending emails. A scheduled queue for handling volume should include retry procedures limited to failed attempts and delivery tracking protocols guarantee efficient low-latency email delivery.

QUESTION-25

What steps would you take to design a content delivery network (CDN)?

- CDN content delivery functions by putting content in edge servers which minimize responses times for users. Assigning users to the closest edge server using DNS routing techniques improves response times while lightening server traffic.

Diagram:

User → Request → Edge Server (Cache) → Origin Server (If Not Cached)



WHY BOSSCODER?

 **2200+** Alumni placed at Top Product-based companies.

 More than **136% hike** for every **2 out of 3** working professional.

 Average package of **24LPA**.

The syllabus is most up-to-date and the list of problems provided covers all important topics.

Lavanya
 Meta



Course is very well structured and streamlined to crack any MAANG company

Rahul .
 Google



[**EXPLORE MORE**](#)