Part 2: As a thought exercise, please describe how you would accomplish the following:

• The size of the URL list could grow infinitely. How might you scale this beyond the memory capacity of the system?

Reply:

1. Use amazon dynamodb which is self scalable database service.
2. Add indexing.

• Assume that the number of requests will exceed the capacity of a single system, describe how might you solve this, and how might this change if you have to distribute this workload to an additional region, such as Europe.

Reply:

1. Use message queue to queue up requests.
2. Do caching or master-slave replica for read heavy system.
3. Sharding DB (horizontal and vertical) for write heavy system.

• What are some strategies you might use to update the service with new URLs? Updates may be as much as 5 thousand URLs a day with updates arriving every 10 minutes.

Reply:

1. If this URL has been verified before, return results directly.
2. If not, analyze it.
3. Queue up all requests.

• You’re woken up at 3am, what are some of the things you’ll look for?

Reply:

1. Check all pipelines and high priority tickets.
2. Find right resources and allocate tickets to different resources.
3. Fix tickets.

• Does that change anything you’ve done in the app?

Reply:

1. Yes, add a detection system to monitor requests.
2. Set threshould for alarm.

• What are some considerations for the lifecycle of the app?

Reply:

1. Well-defined requirement.
2. Well-documented online-docs.
3. Estimation based on development team's research.
4. Development and integration.
5. Fully tested.
6. Deploy by phases.
7. Maintain.

• You need to deploy a new version of this application. What would you do?

Reply:

1. Deploy with little changes once a time.
2. Deploy to a small zone first, and then, after fixing all issues from this small zone, then deploy to another zone.