IX. Disposability

What is fast startup?

- Processes should strive to minimize startup time.
- This is defined as the time from when the launch command is executed until the process is up and ready to receive requests.
- Short startup time provides more agility for the release process and scaling up;
 and it aids robustness.

IX. Disposability

What is graceful shutdown?

- For a <u>web process</u>, graceful shutdown is achieved by **ceasing to listen** on the service port (thereby refusing any new requests), allowing any **current requests to finish**, and then exiting. This implies that **HTTP requests should be short**.
- For a <u>worker process</u>, graceful shutdown is achieved by manually **returning the current job to the work queue**. Other queueing systems return the job automatically to the queue whenever a worker disconnects. For this to happen, jobs are assumed to be **reentrant**, which is typically achieved by wrapping the results in a **transaction** (to allow rollbacks).