LLM Apps Microservices Architecture

Problem

- Hosting an LLM Application in production.
- The Monolithic Architecture presents several problems:
 - Difficult to maintain.
 - Difficult debugging.
 - Difficult to scale.
 - Inefficient resource use.

Solution

- Microservices Architecture.
 - Host the program in separate parts.
 - o Docker.
 - Microservices.
 - Kubernetes.

Process

- It's not simple.
- Usually not the responsibility of the LLM Application developer.
- In this chapter, we will limit ourselves to this conceptual approach.

Other

• For more info, see Docker Documentation.