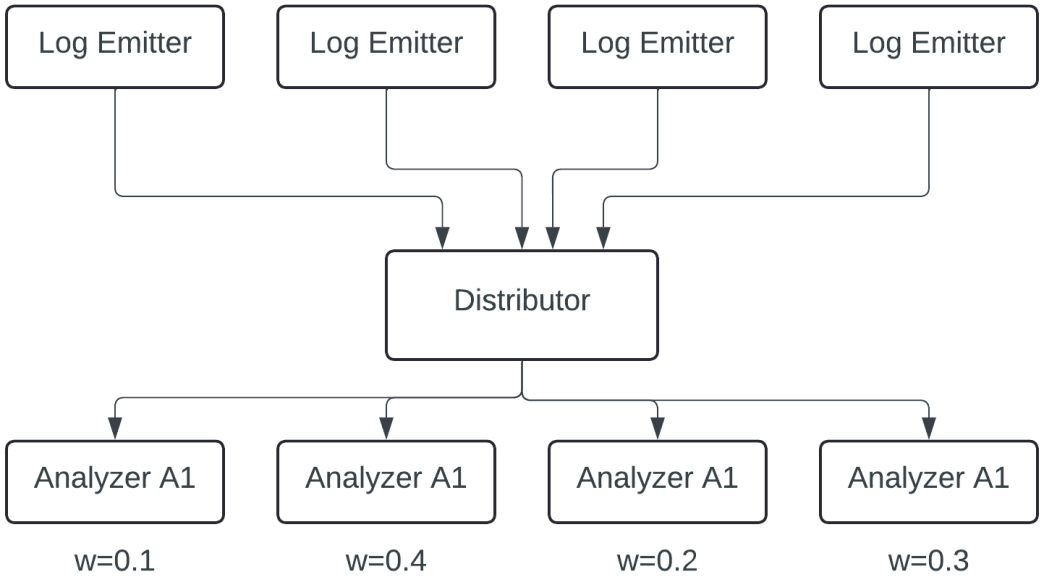
Logs Collector/Distributor Design

Background



The goal is to design a high-throughput logs distributor that will act as an initial receiver of packets of log messages, where each packet could have multiple log messages. The distributor receives log message packets from a number of agents that collect and transmit application/infrastructure logs. The distributor fronts several analyzers, each analyzer being assigned a relative weight (e.g. 0.4, 0.3, 0.1, 0.2) - if it helps you can assume that the weights add up to 1.0. The distributor should route log message packets to analyzers, so that eventually each analyzer analyzes a fraction of log messages roughly proportional to their relative weight.

Task

Your task is to:

- Define a data model for a log message and packet
- Develop a multi-threaded web server (e.g. java springboot but feel free to use your own framework) that accepts a post request for a log packet
- Design the logic for high-throughput non-blocking thread-safe distribution
- Handle the condition of one or more analyzers going offline for a while, and then coming back online
- Setup a working demo (e.g. docker compose w/ jmeter, but feel free to use whatever you would like) to show:
    - The distribution respects the weights

    - The distributor can handle an analyzer failing and coming back online
    - The system can handle high throughput
- Include clear instructions on how to run the demo locally (e.g. we should be able to run your demo)
- Give me a 1-page write-up on what other conditions you might want to handle or improvements you might want to add given more time, what would be your testing strategy, etc.

Feel free to pick any programming language of your choice.

Expectations

- We expect you will use chatgpt and/or github copilot heavily to enhance your productivity
- Feel free to make any assumptions and state those
- You are welcome to ask clarifying questions

Timeline

- We expect you to not spend more than 8-10 hours max on the problem
- You will have roughly a week to work on the problem, we can accommodate for existing plans etc
- You are welcome to ask clarifying questions. Your contact will collaborate with you on this doc to answer any async questions, and can also hop on to 1-2 calls during the week if needed