

Project Proposal Template

- **Project Name:** Load Modeling and Load Balancing Simulation Platform
- **Project Proposer:** Shripad Nadgowda / snadgowda@meta.com
- **Open Source:** Yes
- **Mentors:** Shripad Nadgowda / snadgowda@meta.com

Preferred Experience

List technical skills and mark each as:

- Required (for most or all team members)
 - Programming: (any) Python, C, Go
- Required (at least one team member)
 - Container /Microservice development
 - Familiarity with KV (Key-Value) Stores (e.g. redis)
 - Familiarity with SQL databases (e.g. postgres)
- Valuable
 - CLI development experience
- Nice to have
 - Basic understanding / experience with Load Balancers

Project Background

Load balancing is one of the foundational functions for any cloud platform. Basically, it allows selection of a particular server (from a set of eligible ones) to serve the request. There are various factors that affect that decision making including current load, machine-type, capacity-weights and many more. For a large scale cloud platform, these factors become more prominent, which makes it hard to come up with and maintain optimal load balancing policy. Also, it makes it difficult to simulate the “what-if” effect of some policy change.

In this project, we will try to build such a load modeling and simulation platform which can approximately represent the steady state of request routing and enable “what-if” hypothesis queries to predict impact of some load balancing policy changes.

Project Description

This project roughly entails following outline:

- **Research and Survey:** Read/Learn about load modeling concepts, existing simulation frameworks.
- **Load Modeling:** First we will be building a load modeling system, where we can simulate the load. For instance, if we have 100K servers with different types (x86, ARM) what will be the CPU utilization of them when they are serving 1B requests/sec. We will be modeling more resources like this (memory, latency, etc.)
- **Load Balancing(LB) Simulator:** Build the LB simulator which can read the simulated load (from load modeling stage) and make LB decisions. Scope only a couple of LB policies.
- **Dashboard / Query Interface:** Build a dashboard that shows the load modeling and LB in action. Also, provide a query interface for “what-if” analysis where users can ask queries like “what-if I increase the request load from 1B/s to 1.2B/s ?” or “what-if I change the LB policy from round-robin to least-connection?”

Learning Outcomes

What skills or knowledge will students gain?

- Load balancing in-practice at large scale (We will not delve into kernel or package-level routing, but, primarily at application-level(L7) routing)
- Databases (KV, SQL) operations
- API request modeling (when a request is being served by a server, what system properties are important)
- Simulation framework - Things to consider when building a simulation framework