



دانشگاه صنعتی نوشیروانی بابل
دانشکده برق و کامپیوتر

پایان نامه برای دریافت درجه‌ی کارشناسی
رشته‌ی کامپیوتر
سامانه اجرای کد از راه دور

دانشجو:
محمدعلی علی پناه

استاد راهنما:
دکتر غلامی

شهریور ۱۴۰۲

[این صفحه خالی است]

لَا إِلَهَ إِلَّا اللَّهُ مُحَمَّدٌ رَسُوْلُهُ

تقديم

به مادر ، پدر و يگانه عزيزم كه نور مسير و اميد زندگي من اند.

فهرست مطالب

۱	آشنایی با پروژه	۱
۳	مفاهیم پایه	۲
۳ Firecracker	۱-۲
۴ Fiber Go	۲-۲
۴ RabbitMQ	۳-۲
۴ PostgreSQL	۴-۲
۴ Next.js	۵-۲
۵	معماری پروژه	۳
۷ API	۱-۳
۸ VMVisor	۲-۳
۹ Frontline	۳-۳
۱۰ کلاینت	۴-۳

[این صفحه خالی است]

فهرست شکل‌ها

۱	۱-۱	لوگو پروژه
۲	۲-۱	ساختار کلی پروژه
۳	۱-۲	نمایی از ساختار Firecracker
۵	۱-۳	معماری پروژه
۷	۲-۳	لیست API
۸	۳-۳	وضعیت اجرا کد
۱۰	۴-۳	طرح برخی از صفحه‌ها
۱۱	۵-۳	لیست پروژه‌ها

[این صفحه خالی است]

فصل ۱

آشنایی با پروژه

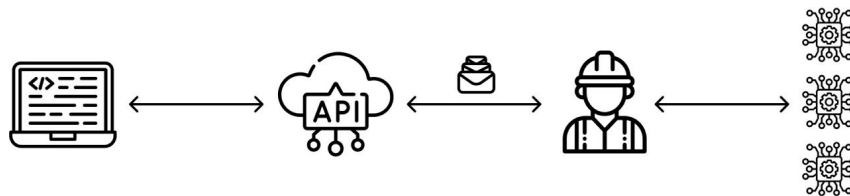


شکل ۱-۱: لوگو پروژه

YARCEE مخفف Yet Another Remote Code Execution Engine یک سرویس ابری اجرای کد می‌باشد. معماری این پروژه مایکروسرویس است و سرویس‌ها با زبان Go نوشته شده‌اند. برخی از ویژگی‌های کلیدی پروژه شامل:

- اجرای کد زبان‌های مختلف مانند python - nodejs - c - c++ و ...
- عدم دسترسی پروژه‌ها به یکدیگر به دلیل اجرا در micro-vm
- ادیتور آنلاین و ویرایش در مرورگر
- قابلیت ساخت ، ادامه و تغییر نام پروژه
- همروند بودن اجرای کد توسط صف پیام RabbitMQ

این پروژه یک سرویس ابری است که شامل ویژگی های یک محصول واقعی است. کاربر امکان ثبت نام و ورود دارد. می تواند پروژه جدید بسازد و آن را ویرایش و حذف بکند. در فصل جمع بندی نگاهی به آینده این محصول می کنیم اما قبل از آن به صورت خیلی خلاصه نحوه عملکرد این سیستم را بررسی می کنیم.



شکل ۱-۲: ساختار کلی پروژه

پس از ثبت نام و ورود به داشبورد، از قالب های از پیش تعریف شده یکی را به دلخواه انتخاب می کنیم. هر قالب شامل اسمی تصادفی، کد اولیه زبان و اطلاعاتی در مورد اجرای آن می باشد. پس از ویرایش کد آن را اجرا می کنیم. در پشت صحنه درخواستی برای دریافت خروجی کد به سرویس API زده می شود. وظیفه این سرویس ارتباط مستقیم با پایگاه داده و گوش دادن به وضعیت اجرای کد و خروجی آن می باشد.

این درخواست روی صف RabbitMQ ارسال می شود و توسط سرویسی از صف دریافت می شود. این سرویس سپس درخواست را برای پردازش یکی از vm هایی که در حالت آماده باش قرار دارد می فرستد و خروجی اجرای کد را دریافت می کند.

مجموعه از vm ها در حالت آماده باش قرار دارند که پس از اجرای کد حذف و سپس جایگزین به مجموعه اضافه می شود. متصل به هر vm فایل سیستمی از پیش ساخته شده که است که باینری کامپایلرهای مختلف درون آن وجود دارد.

وابستگی اصلی این پروژه به Firecracker می باشد که برای مدیریت micro-vm ها به کار می رود. برای ارتباط با این ابزار از sdk زبان GO آن استفاده شده است. از مزیت های استفاده از روش micro-vm می توان سرعت بسیار بالاتر نسبت به vm اشاره کرد. همچنین این روش امنیت بالاتری نسبت به اجرای کد در محیط های کانتینری مانند docker دارد زیرا micro-vm ها به طور کل از یکدیگر جدا هستند.

بخش مهم دیگر این پروژه محیط وب و ادیتور آنلاین آن می باشد. کلاینت با زبان TypeScript نوشته شده است و از فریمورک Next.js استفاده می کند. کتابخانه ui زبان به طبع React می باشد. در بخش ادیتور امکان ویرایش کن و درخواست اجرا وجود دارد. با فشردن اجرا کلاینت هر از ۱۰۰ میلی ثانیه به سرور درخواست میزد تا خروجی کد درخواست شده را دریافت کند. پس دریافت stdout و stderr این چرخه را متوقف می کند. در ادامه به معرفی هر سرویس خواهیم پرداخت.

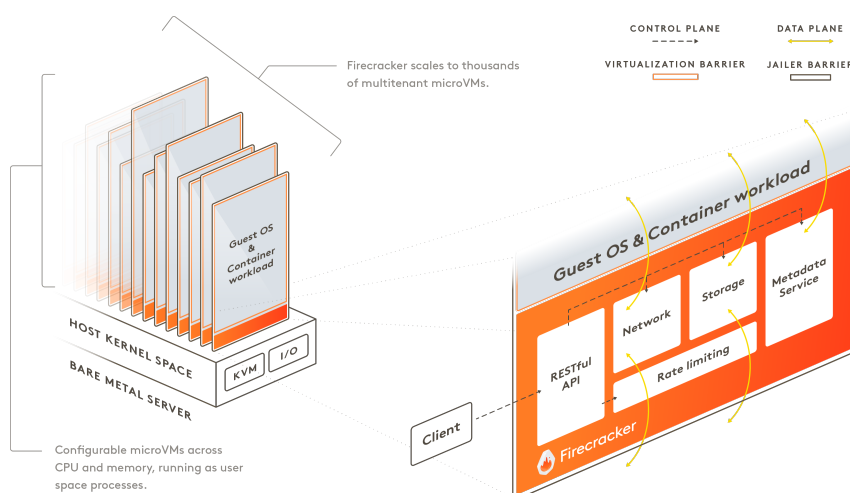
فصل ۲

مفاهیم پایه

در این فصل به معرفی تکنولوژی و فریمورک های به کار رفته در پروژه میپردازیم.

۱-۲ Firecracker

در قلب پروژه Firecracker قرار دارد. Firecracker یک مانیتور ماشین مجازی است که از KVM استفاده می کند و وظیفه اش ساخت و مدیریت ماشین های مجازی است. Firecracker توسط تیم وب سرویس آمازون توسعه داده شده و در پروژه Fargate و Lambda این شرکت نیز استفاده شده است.



شکل ۱-۲: نمایی از ساختار Firecracker

۲-۲ Fiber Go

زبان استفاده شده در میکروسرویس ها Go می باشد که توسط گوگل توسعه داده شده و از فریمورک Fiber استفاده شده که برای ساده تر شدن routing و middleware استفاده شده است. از دلایل استفاده از Go میتوان به سادگی و سرعت بالا اشاره کرد. همچنین این زبان در مسائل concurrency ابزارهای low-level زیادی در دسترس کاربر قرار می دهد.

۳-۲ RabbitMQ

RabbitMQ یک نرم افزار برای انتقال پیام بین سیستم ها است. در این پروژه درخواست اجرا کد در صف وارد می شود و توسط سرویسی پردازش می شود. دلیل استفاده از event-driven بلاک نشدن درخواست ها است. مزیت استفاده از RabbitMQ آسنکرون شدن سیستم است. پیام ها و وضعیت اجرای کد پشت یکدیگر بلاک نمی شوند. همچنین سیستم ها از وجود یکدیگر بی خبر هستند و وابستگی شان بهم کم می شود. به این معماری loosely-coupled می گویند.

۴-۲ PostgreSQL

دیتابیس اصلی استفاده شده PostgreSQL می باشد که از نوع رابطه ای است. جداول این نرم افزار شامل کاربران و کدها می باشد. دلیل انتخاب PostgreSQL مورد اطمینان بودن و سادگی این پایگاه داده بوده است. برای ارتباط و کوئری زدن از کتابخانه gorm زبان Go استفاده شده که ORM محبوبی است.

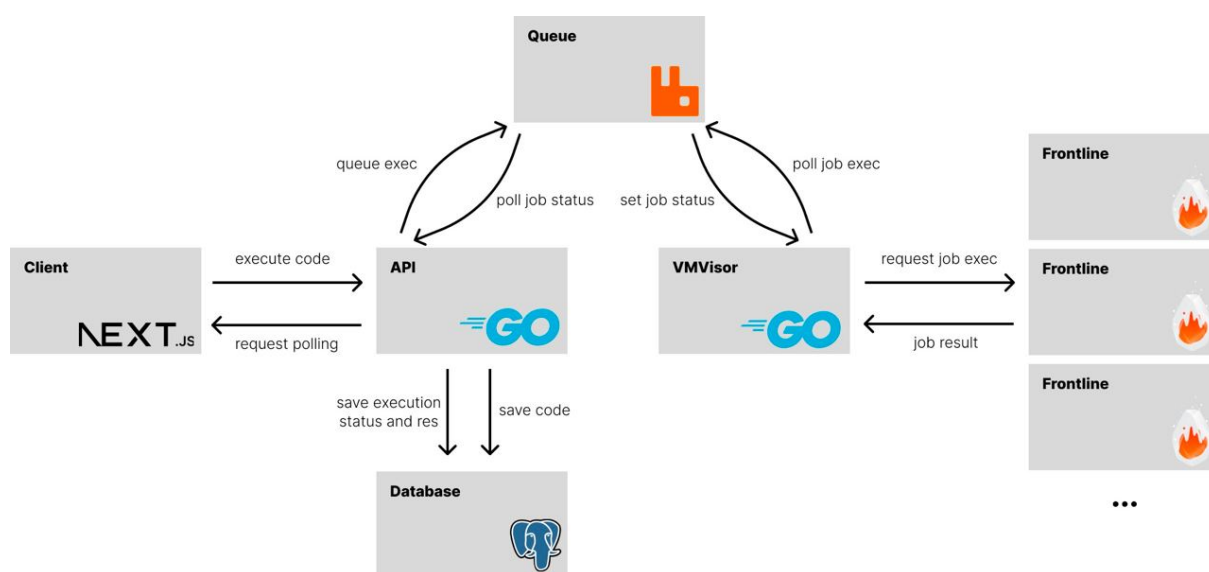
۵-۲ Next.js

فریمورک استفاده شده سمت کلاینت Next.js می باشد که از کتابخانه React برای رندر روی مرورگر استفاده میکند. دلیل استفاده از React ساده کردن پیاده سازی رابط کاربری توسط هوک ها و کامپوننت محور بودن آن است. زبان برنامه نویسی سمت کلاینت TypeScript است که به واسطه کامپایلر تبدیل به JavaScript می شود. دلیل استفاده از TypeScript اضافه شدن شی گرایی و تایپ در زمان کامپایل است.

فصل ۳

معماری پروژه

همانطور که در ۱-۳ می بینید این پروژه از معماری میکروسرویس بهره می برد. در این فصل ابتدا به معرفی کلی هر سرویس و سپس نگاهی دقیق به عملکرد هر سرویس می اندازیم.



شکل ۱-۳: معماری پروژه

همانطور که در شکل و جدول مشاهده می‌کنید پروژه از ۴ سرویس اصلی تشکیل شده است. درخواست از کلاینت شروع شده پس از طی کردن API و VMVisor به Frontline می‌رسد. پیش تر اشاره کردیم که Frontline به API Web است که درون vm در حال اجراست. Frontline در اصل یک API Web می‌باشد که با پروسه فرزند کامپایلر زبان های مختلف را فراخوانی کرده و خروجی را بر می‌گرداند.

جدول ۳-۱: لیست سرویس ها

سرویس ها	توضیحات
API	سرویس API REST می‌باشد که وظیفه صحبت با دیتابیس و پاسخ به کلاینت را دارد
VMVisor	مدیریت VM های ساخته و دریافت و تغییر وضعیت درخواست های اجرا روی صف
Frontline	درون هر VM در حال اجراست و توسط پروسه فرزند کامپایلر زبان را صدا می‌زند
Client	کلاینت وظیفه نمایش رابط کاربری و ادیتور را دارد

مهم ترین سرویس این پروژه VMVisor است که وظیفه مدیریت vm ها را بر عهده دارد. این پروژه از Firecracker استفاده می‌کند و مجموعه ای از vm ها را مدیریت و به آن ها وظیفه ای برای اجرا می‌سپارد. ارتباط بین API و VMVisor به صورت آستکرون توسط انتقال پیام در صف می‌باشد. در RabbitMQ دو صف وجود دارد. صف درخواست اجرا و صف وضعیت اجرا. در صف درخواست اجرا سرویس API کدی که کلاینت ارسال کرده را روی صف قرار می‌دهد و VMVisor این درخواست را از صف برمی‌دارد. صف دیگر وضعیت اجرا است که VMVisor خروجی Frontline را روی صف قرار می‌دهد و API به آن گوش می‌دهد و روی پایگاه داده می‌نویسد.

وظیفه API عملیات های CRUD (create update read delete) است و اولین درگاهی است که کلاینت با آن در ارتباط است. همچنین وظیفه قرار دادن درخواست اجرا در صف و گوش دادن به صف وضعیت اجرا و نوشتن آن روی پایگاه داده را برعهده دارد. از دیگر وظیفه های این سرویس مدیریت کاربران و پروژه های آن ها می‌باشد. این سرویس خود پتانسیل شکسته شدن به میکروسرویس های کوچک تر را دارد ولی در این پروژه این تصمیم گرفته نشده است.

کلاینت هم بخش مهم دیگری است که رابط کاربری سیستم با نرم افزار است. کاربر امکان ساخت پروژه جدید و ویرایش آن در مرورگر را دارد. رابط کاربری این پروژه در Figma طراحی شده و توسط کتابخانه React دیزاین سیستم طراحی شده و کامپوننت های مختلف در کنار هم قرار گرفته اند. در ادامه به معرفی دقیق تر هر سرویس و ارتباطش با سایرین می‌پردازیم.

۱-۳ API

همانطور که اشاره کردم بخش API پروژه از زبان Go و فریمورک Fiber استفاده می‌کند. این پروژه لایه ورودی ما به بخش های داخلی سیستم می‌باشد.

در عکس زیر نمایی از api های موجود در پروژه مشاهده می‌شود. این api ها در چند دسته مختلف تقسیم بندی شده اند. sandbox برای ساخت یک پروژه جدید و اجرا آن می‌باشد. بخش auth برای ثبت نام و ورود کاربر است. بخش user برای دریافت اطلاعات کاربر وارد شده است. و health برای بررسی liveness و readiness سیستم در نظر گرفته شده است. وجود این مسیر باعث می‌شود در سیستم های مدیریت کانتینر مانند kubernetes از آمادگی سرویس اطمینان حاصل کرد.

Sandbox			^
GET	/api/sandbox/	Get user sandboxes	⌵
POST	/api/sandbox/	Create sandbox	⌵
DELETE	/api/sandbox/	Delete sandbox	⌵
GET	/api/sandbox/{id}	Get sandbox	⌵
PUT	/api/sandbox/{id}	Update sandbox	⌵
POST	/api/sandbox/{id}/execute	Update sandbox	⌵
Auth			^
POST	/api/sign-in/	Sign in	⌵
POST	/api/sign-up/	Sign up	⌵
User			^
GET	/api/user/	User auth details	⌵
Health check			^
GET	/health_check/	Health check	⌵

شکل ۳-۲: لیست API

همچنین این سرویس به دو صف متصل است که به آن گوش می‌دهد و روی آن ارسال می‌کند. کلاینت پس از ساخت پروژه و ویرایش کد، درخواست اجرا آن را ثبت می‌کند. در پشت صحنه درخواستی به مسیر execute زده می‌شود. این درخواست روی صف sandbox queue ارسال می‌شود.

از طرفی سرویس VMVisor این درخواست را به یک vm می‌سپارد و خروجی کد و وضعیت اجرای آن را روی صف sandbox status queue قرار می‌دهد. سرویس API به آن صف گوش می‌دهد و آن را روی پایگاه داده می‌نویسد. کلاینت هر از ۱۰۰ میلی ثانیه در حال درخواست برای stdout و stderr است.

۲-۳ VMVisor

شاید پیچیده ترین سرویس این پروژه VMVisor باشد. این سرویس چندین وظیفه دارد شامل:

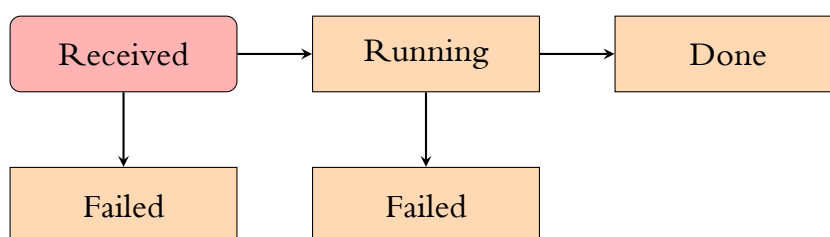
- مدیریت و ساختن مجموعه vm
- دریافت درخواست اجرا کد از صف
- ارتباط با Frontline برای اجرای کد
- ارسال تغییر وضعیت کد به صف

تعداد مجموعه vm های آماده ۱۰ عدد می باشد. به هر vm یک عدد هسته پردازنده و ۲۵۶ مگابایت حافظه RAM و ۱ گیگابایت فضا اختصاص می شود.

ماکسیمم مدت زمان اجرا ۱۰ ثانیه است. پس از آن vm حذف می شود. پس از اجرا یک vm به مجموعه اضافه می شود تا تعداد vm های آماده همان ۱۰ عدد بماند.

ارتباط VMVisor با vm از طریق سرویس به نام Frontline است. Frontline در اصل یک REST API است که درون vm در حال اجراست و درخواست را دریافت کرده و خروجی را به VMVisor بر می گرداند. این سرویس سپس وضعیت اجرا را روی صف قرار می دهد. در شکل زیر وضعیت های مختلف اجرا کد را می توانید ببینید.

شکل ۳-۳: وضعیت اجرا کد



در ادامه به آخرین تکه پازل یعنی Frontline می پردازیم.

۳-۳ Frontline

همانطور که اشاره کردیم Frontline یک Web API است که با زبان Go توسعه داده شده. این سرویس درون vm در حال اجراست و از طریق پردازنده فرزند کامپایلر زبان مورد نظر را فراخوانی می‌کند. این سرویس از طریق openrc که سرویس init برای لینوکس می‌باشد بلافاصله بعد از بوت vm در حال اجرا قرار می‌گیرد.

درون فایل سیستم هر vm تمام کامپایلرهای مورد نیاز قرار دارد. این فایل سیستم با دستور dd با فضای یک گیگابایت ساخته شده است و از طریق docker روی آن نوشته می‌شود. این پروسه توسط اسکریپتی انجام می‌شود. در این اسکریپت توسط alpine و پکیج منجر آن کامپایلر زبان های مختلف دانلود می‌شود. فایل سیستم به عنوان volume برای docker اضافه می‌شود و از این طریق امکان نصب این کامپایلرها روی آن امکان پذیر می‌شود.

هر vm نیاز به گرفتن ip دارد. VMVisor از طریق درخواست HTTP با Frontline در ارتباط است. برای بحث نتورکینگ از CNI استفاده شده است که از بحث این مقاله خارج است.

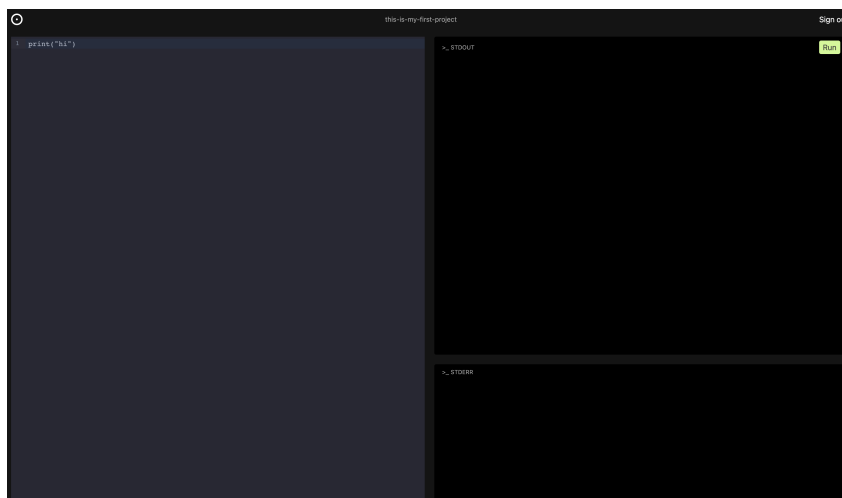
درون Frontline دو مسیر اصلی وجود دارد. health و health exec. برای بررسی آمادگی vm به کار می‌رود. پیش تر اشاره کردیم که وظیفه VMVisor آماده نگه داشتن ۱۰ vm می‌باشد. ساختن vm جدید ممکن است چندین ثانیه طول بکشد. هر ۱۰ ثانیه VMVisor درخواستی به health می‌زند و پس از دریافت ۲۰۰ آن vm را به مجموعه اضافه می‌کند.

۴-۳ کلاینت

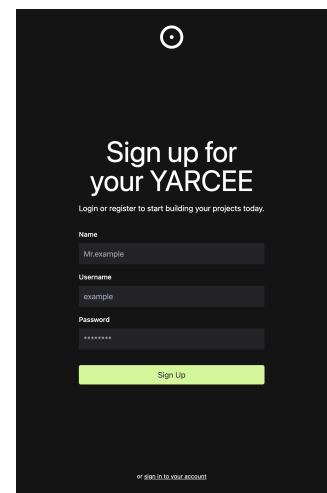
همانطور که اشاره کردیم کلاینت این پروژه از Next.js استفاده می‌کند. طراحی رابط کاربری در محیط Figma انجام شده است.

مراحل پیاده سازی به شرح زیر است:

۱. دیزاین توکن ها را استخراج و به پروژه اضافه می‌کنیم. مانند رنگ ها، فاصله ها و سایه ها
۲. المنت های دیزاین سیستم رو پیاده سازی می‌کنیم. کامپوننت هایی مانند دکمه، اینپوت و کانتینر
۳. با کنار هم قرار دادن المنت های دیزاین سیستم و کامپوننت های مخصوص هر بخش، صفحه را تکمیل می‌کنیم
۴. مراحل دریافت یا فرستان اطلاعات را انجام می‌دهیم
۵. مرحله ۳ و ۴ را برای هر صفحه تکرار میکنیم



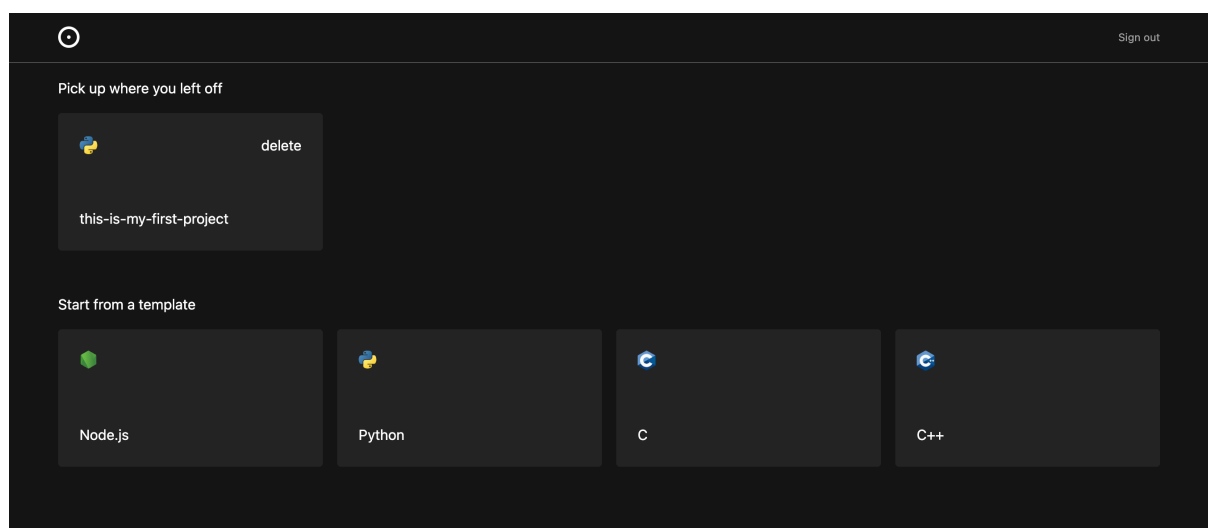
(ب) صفحه ادیت کد



(الف) صفحه ثبت نام

شکل ۴-۳: طرح برخی از صفحه ها

در شکل بالا می‌توانید نمونه ای از صفحه های پیاده سازی شده در این پروژه را مشاهده کنید. پس ثبت نام و ورود به سایت شما با صفحه داشبورد مواجه می‌شوید. در این صفحه می‌توانید به ادامه ویرایش پروژه قبلی خود بپردازید یا توسط قالب های از پیش تعیین شده پروژه جدیدی شروع کنید. اضافه کردن اکثر زبان های برنامه نویسی ممکن است ولی در حال حاضر از Python، Node.js، C++ و پشتیبانی می‌شود. اضافه کردن زبان جدید به سادگی امکان پذیر است که در ادامه در بخش Frontline به آن می‌پردازیم.



شکل ۳-۵: لیست پروژه ها

[این صفحه خالی است]

Abstract

YARCEE(Yet Another Remote Code Execution Engine) is a code running service that relies on Firecracker to spawn microVMs and execute the code over HTTP.

Keyword: Remote code execution - MicroVM - VMM



Babol Noshirvani University of Technology
Faculty (or School) of ...

A Thesis

*Submitted in Partial Fulfillment of the Requirement for the Degree of Master of Science in
Computer*

Remote code execution engine

by:

Mohammad ali Ali panah

Supervisor:

Dr.Gholami

August, 2023