



دانشکده مهندسی
کامپیوتر و فناوری اطلاعات



دانشگاه صنعتی امیرکبیر
(پلی تکنیک تهران)

دانشگاه صنعتی امیرکبیر
دانشکده مهندسی کامپیوتر و فناوری اطلاعات

سیستم‌های توزیعی پروژه سوم

پرهام الوانی

۶ دی ۱۳۹۶

در فاز اول دو تابع `map` و `reduce` با استفاده از تابع `FieldsFunc` در زبان گو که برای قطعه قطعه کردن رشته‌ها استفاده می‌شود پیاده‌سازی شدند. همانطور که در صورت پروژه نیز شرح داده شده بود قرار بود رشته‌ی ورودی با هر کاراکتر غیر از حروف قطعه قطعه شود و به این منظور از تابع `IsLetter` استفاده شد.

در فاز دوم یک سیستم توزیع شده برای انجام `MapReduce` در قالب یک کتابخانه پیاده‌سازی می‌گردد.

یک بخش از کد این کتابخانه در قالب `master` وظیفه دارد `register` شدن `worker`ها را سازمان‌دهی کند و کارهایی که در سیستم موجود است را به آن‌ها تخصیص دهد. بعد از اتصال هر `worker` یک `go routine` برای آن اجرا می‌گردد. این `go routine` در ابتدا تلاش می‌کند از کانال تسک‌های `map` تسک استخراج کرده و آن را به `worker` خود تخصیص دهد. در پایان هر تسک نتیجه‌ی آن در کانال `mapDone` منتشر می‌گردد. بعد از اینکه `master` مطمئن شد نتیجه‌ی تمام `nMap` را جمع‌آوری کرده است با بستن کانال `reduceStart` به تمام `go routine`هایی که `worker`ها را مدیریت می‌کنند اطلاع می‌دهد، می‌توانند وارد تسک‌های `reduce` شوند.

در ادامه همان روند تسک‌های `map` برای تسک‌های `reduce` طی شده و نتایج در کانال `reduceDone` منتشر می‌گردد.

با پایان یافتن تسک‌های `reduce` برنامه به صورت کامل خاتمه می‌یابد. با توجه به روندی که شرح داده شد اضافه شدن `worker`ها در میانه‌ی کار امکان‌پذیر بوده و در صورت `timeout` شدن هر `worker` تسک آن دوباره به کانال تسک‌ها اضافه شده و یک `worker` دیگر آن را برعهده خواهد گرفت.