# lab1

## worker

### `map`任务：

| field(args)申请 | 类型 | 解释 |
| --- | --- | --- |
| | | |

| field(reply) | 类型 | 解释 |
| --- | --- | --- |
| Filename | string | 对应input split文件名 |
| ID | string | 该worker是几号map worker |
| R | int | 有多少个reduce worker |
| Task | string | 任务类型MAP/REDUCE |
| Message | string | RPC信息 |

| field(args)完成 | 类型 | 解释 |
| --- | --- | --- |
| ID | string | 该worker是几号map worker |
| Task | string | 任务类型MAP/REDUCE |

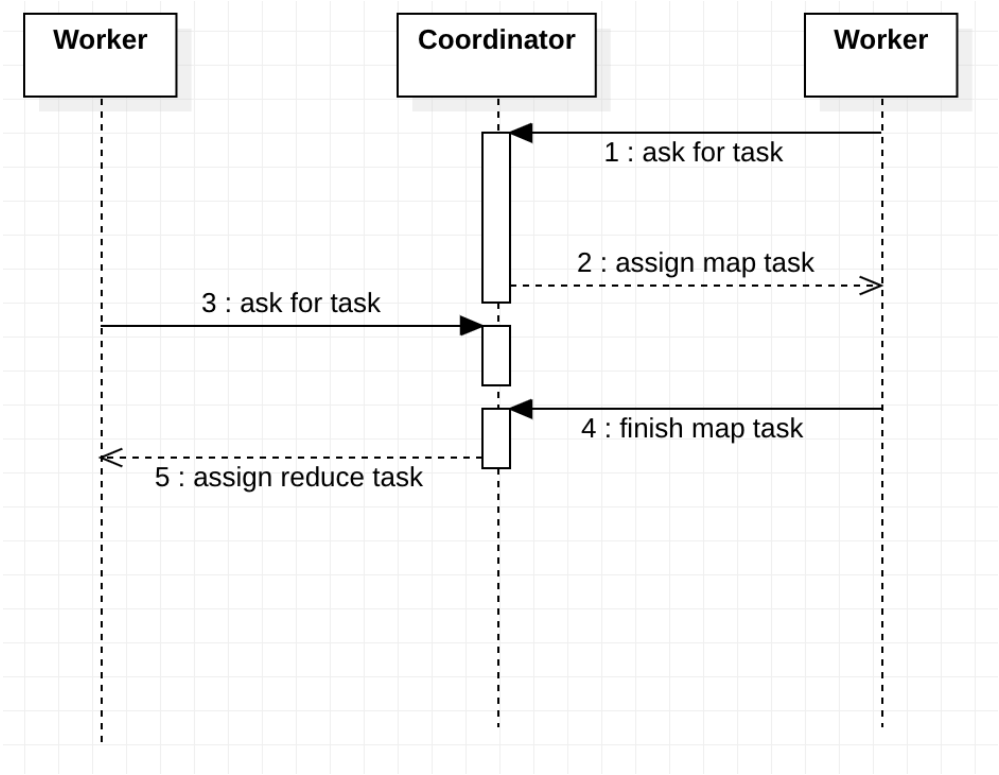
### `reduce`任务：

| field(args) | 类型 | 解释 |
| --- | --- | --- |

| field(reply) | 类型 | 解释 |
| --- | --- | --- |
| Filenames | []string | 中间文件 |
| Task | string | 任务类型MAP/REDUCE |
| Message | string | RPC信息 |
| ID | string | 该worker是几号reduce worker |

| field(args)完成 | 类型 | 解释 |
| --- | --- | --- |
| ID | string | 该worker是几号reduce worker |
| Task | string | 任务类型MAP/REDUCE |

- 时序图示意




- 获取文件名和map函数，进行map函数处理，用json存储中间keyvalue结果，返回给Coordinator存储位置和中间文件的文件名
- rpc定义了请求的参数和返回值

## Coordinator

Coordinator要维护一个map任务列表，一个reduce任务列表。每个任务有状态，workerid，文件名

对于并发请求的处理

用过用同步锁的话，每次查表前要lock，查完表unlock

查表的场景：getTask要查表找个任务，finishTask要确定完成的worker就是之前分配的worker，对于超时的任务要重置状态。

map未分配任务表：

| Field | 解释 |
| --- | --- |
| 任务名(filename) | input file |
| | |

map正在进行任务表

| Field | 解释 |
| --- | --- |
| 任务名(filename) | input file |
| 执行者(worker id) | 交给哪个worker做了 |

map已完成任务表

| Field | 解释 |
| --- | --- |
| 任务名(filename) | input file |
| 执行者(worker id) | 交给哪个worker做了 |

reduce未分配任务表

| Field | 解释 |
| --- | --- |
| 任务名(filenames) | Intermediate files（按reducer分配 |
| 执行者(worker id) | 要交给哪个worker做 |

reduce正在进行任务表

| Field | 解释 |
| --- | --- |
| 任务名(filenames) | Intermediate files（按reducer分配 |
| 执行者(worker id) | 交给哪个worker做了 |

reduce已完成任务表

| Field | 解释 |
| --- | --- |
| 任务名(filenames) | Intermediate files（按reducer分配 |
| 执行者(worker id) | 交给哪个worker做了 |

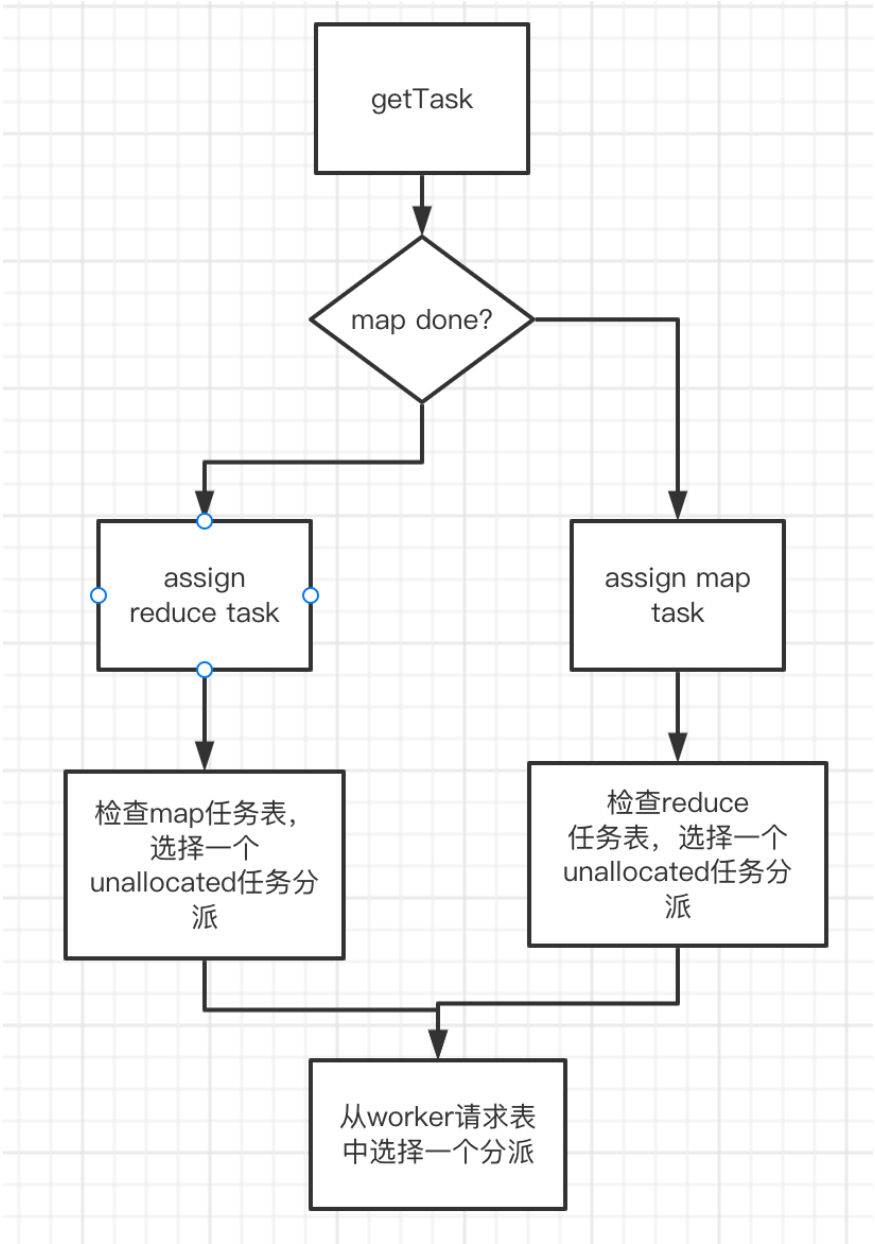
- 任务状态

| 编号 | 状态 | 解释 |
| --- | --- | --- |
| 0 | unallocated | 任务未分配 |
| 1 | processing | 任务已分配 |
| 2 | finish | 任务已完成 |

`getTask`：

| field(args)申请 | 类型 | 解释 |
| --- | --- | --- |

| 任务 | 类型 | 解释 |
| --- | --- | --- |
| 查看任务表 | string | 该worker是几号map worker |
| Task | string | 任务类型MAP/REDUCE |




`Finish`：

| field(args)完成 | 类型 | 解释 |
| --- | --- | --- |
| ID | string | 该worker是几号map worker |
| Task | string | 任务类型MAP/REDUCE |

- 对于sequencial而言

```
$ go run -race mrsequential.go wc.so pg*.txt
```

- 对于并发调度

```
# main/mrcoordinator.go
$ go run -race mrcoordinator.go pg-*.txt
# main/mrworker.go
$ go run -race mrworker.go wc.so
```