

Single Crawler Overview

- Send unique URLs to workers
- Keep pages if they are unique
- Terminate crawling when done

Program Termination Conditions

- No URLs left
 - Numgoproc =0
 - Len ulist=0
- Max URLs reached
- Timeout
- No new unique pages are being found

Worker Overview

- Make http request
- Parse request
- Send back URLs
- Send back created page

Worker Termination Conditions

- Error requesting
- Error parsing
- Taking too long