



## ActiveFence Fullstack Candidate Home Assignment

Thanks for taking the time to complete this assignment. The **Challenge** is to create a Spider Service (aka crawler).

### The system should have the following components:

1. Front-end
  - a. A simple form for entering scrape requests which will include at least the basic params for a scrape job:
    - i. `startUrl` - the url to start scraping from
    - ii. `maxDepth` - the maximum depth to crawl down to from the start url
    - iii. `maxTotalPages` - the max number of pages for the entire scrape job
  - b. A display of the actual jobs with as much relevant information as you think would be helpful.
2. Back-end (Node.js. ES6 syntax required)
  - a. An api service for the front end to communicate with. It should implement either REST over http or a well structured api over websockets.
  - b. A crawler service that handles the actual crawling, with the following specs:
    - i. Use the BFS algorithm. BFS here means that all items of depth N finish processing before any items from depth N+1 start.
    - ii. Stop crawling a job when it reaches `maxDepth` or `maxPages`, whichever comes first.
    - iii. For each page, save at least:
      1. `title` - The document.title of the page
      2. `depth` - Current depth being scraped
      3. `url` - The URL that was scraped
      4. `links` - All hrefs in the anchor tags in the page
3. Include a README.md with at least the following information:
  - a. How to build and run the project.
  - b. Solution overview
  - c. What features would you have added to the spec? What would you have done better given more time?

### Notes & Hints:

- **Focus on Back-end functionality. Front-end should be lean and practical.**
- Build for horizontal scale. Pure in-memory solutions will not be accepted.
- Good level of documentation in code is required.

### Bonus points:

- Unit tests welcome.
- Make the UI update in real time.

**Deliverable:** Create a private repo with your solution on bitbucket.org and share it with [activefence.rnd@gmail.com](mailto:activefence.rnd@gmail.com).

*Feel free to reach out to us with any questions.*

**Good Luck!**