

# Laboratory Assignment 1.3 - Summarizing your Findings / Lab Setup

- Due Sep 29 by 11:59pm
- Points 100
- Submitting a text entry box or a file upload
- Available Sep 23 at 12am - Sep 29 at 11:59pm

This assignment was locked Sep 29 at 11:59pm.

## Laboratory Assignment 1.3 - Recap of your Experience

*You may complete this as groups; however, each person will need to submit their work.*

### Prerequisites:

- Functional Hadoop / Jupyter / Dask / pyspark etc. cluster. The procedure I provided for Docker will work best with this lab.
- Knowledge of the system, architecture, layout.
- Jupyter Notebooks knowledge.
- Some Python, experience with Python package includes.

### Purpose:

The purpose of this lab is to demonstrate the use of your cluster. At this time, you likely have a "3 node system" composed of a master, worker, and edge node. While this isn't distributed per se, we can test our cluster functionality before expanding in a later lab.

### Procedure:

1. Procedure is referenced in parts 1.1. and 1.2. You will need these completed for this write-up. Once you have a working cluster, complete the deliverables below.

### Deliverables:

For this part, please submit the following through Canvas:

1. An overview of the architecture used. I.e. did you use Docker or another approach? If you used another approach, please briefly explain how you implemented this lab and any results you would like to share.
2. For those who went with the Docker implementation, please tell me more about the experience, any hiccups, and your successes.
3. In 1.1 you built the system, in 1.2, you began testing. I'd like to see one or two more test cases. You may use the test cases in ZyBooks, or develop your own. The word search is a great starting point.
4. Minimum 2-4 page write up in Docx (Word) or PDF format explaining the points above. Adding diagrams / additional technical documentation helps me understand the systems and how they

"work" on different architectures. Be sure to include:

1. Any supporting documentation, diagrams, notes.
  2. Sources you used to implement this lab and any documentation you used to troubleshoot.  
Sources should be complete and not "google / ChatGPT". Include links or professional citations please.
  3. Brief section about any challenges / your successes.
  4. Include your findings from 1.2's testing. I.e. what worked, what didn't? And what were the results?
  5. Conclusion statement and next steps to expand the cluster (feel free to provide a theoretical as we will implement in lab 2 - expanding your cluster).
5. This lab assignment will be submitted through Canvas and is due Sunday, September 29th @ 11:59pm. I may extend if there is reason to.
6. You will receive 100 pts for the implementation, 100pts for testing, and 100 pts for this part of the lab exercise. Your submission should be of high quality.
7. Use of AI is allowed; however, your voice / statements are important for the learning components of this lab. Please use AI as a guide and NOT to formulate your submissions. These should be in your own words via paraphrasing.