Group Project Report

CMPT 353 Summer 2020

Orion Hsu ohsu@sfu.ca

Wendy Huang wha57@sfu.ca

Vincent Law vclaw@sfu.ca

August 13, 2020

1 Introduction

1.1 Problem

Linus Tech Tips did a video about an experiment where they took measurements of air temperatures at the inlet and outlet of radiators in an enclosed wind tunnel. LTT's experiment attempts to determine the effectiveness of stacking multiple radiators as a cooling solution for a water cooling loop with a CPU and GPU. LTT presented the data in the video exactly as they gathered it, without applying the techniques found in CMPT 353 such as LOESS or Kalman smoothing, or without using statistical properties to determine the significance of their claims. In this project we seek to:

- Clean and smooth the data provided
- Learn about the relationship between stacking radiators on radiator exhaust temperature
- See how many radiators is most optimal for thermal performance

2 Methodology

2.1 Data Acquisition

2.1.1 Setup

Seven 240MM crossflow radiators from Alphacool were used to acquire raw data, with low profile fittings to stack those radiators together and 3D printed end caps for holding the radiators. Wind tunnel was snapped together from clear acrylic letting the air flow go between radiators. Two thermistor-type temperature probes were attached to each radiator to measure the temperature difference of water entering and exiting the radiator.

Collin from Linus Tech Tips created a data logging system which samples data from the sensors every 250 milliseconds (four times per sensor) with the help of Arduino code. The team then converts the resistance of the thermistor senator embedded in each of the radiators into Celsius degrees using the Steinhart-Hart Equation. Afterwards, they utilized trimpods to calibrate each sensor to a known temperature using a water bath. Once all the sensors are readings within 0.5^{o} of each other, the data is pushed out via a serial connection to a computer running PuTTy which outputs all the data into a text file for later processing.

In terms of heat source, the team paired a 32-core Threadripper 3970x CPU with Nvidia Titan V, which emits 500 watts of heat into the system. The team also used a D5 pump with no-name reservoir for the system's water source.

2.1.2 Experiment

The datasets from Linus Tech Tips includes 6 different valid tests and 2 invalid ones. Among the valid tests, only 3 of them were used for the experiment done in the YouTube video.

In the following paragraphs, we will be discussing the experiment method and findings of each test.

Test 1

Each radiator was paired with a Noctua NF-F12 fan and used the fully assembled stacked radiators down the windtunnel, with air flowing through the radiators from sensor 0 to sensor 7. For the GPU stress test, FurMark, an intensive OpenGL benchmark was used to fully utilize the graphic card. Cold water with the same starting temperature was used for the experiment.

The result for test 1 indicates that most of the cooling work is done by the first radiator, with the sequential radiator performing less and less work; the reason being the first radiator's contact with the coldest air and the hottest water coming directly from the system. The point of negligible returns was reached at the sixth radiator in the stack. The radiators got progressively hotter in sequence within the first two minutes, and stayed the hottest when FurMark was closed, with the latter radiators reheating the water coming from the first radiator.

Test 2

The second test had the exact same setup as the first one, with the only difference in the reversed waterflow direction in the second test (see Figure 1). The result from test 2 indicates a minor improvement as all of the temperature of the radiators increase more slowly while the stack got hotter.

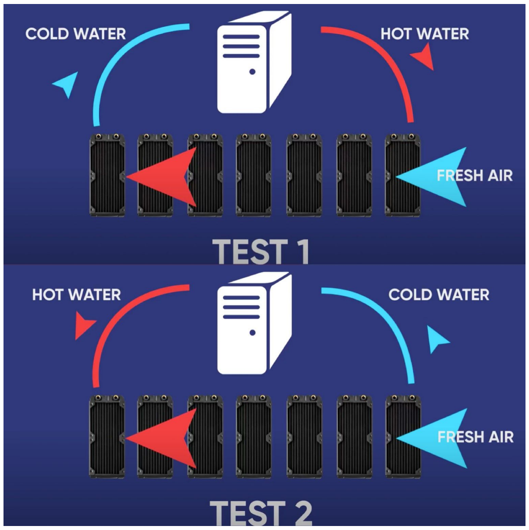

Figure 1. Water flow directions for test 1 and test 2 (Linus Tech Tips, 2020, 07:20–07:40).

Test 3

Different from the prior two, the third test had a more sensible setup, including one bank of fans and two radiators. Since there was less volume in the loop, the temperature rose much quicker. The Linus Tech Tips team ran the same test with a single radiator instead of two, and both hit a peak temperature of 38.5 degree Celsius. The result from the third test indicates that having a second radiator does not have much impact on cooling.

2.2 Data Cleaning

The initial data provided by Colin from Linus Tech Tips was in a Microsoft Excel spreadsheet with all the tests put together. To make each test more friendly to load into Pandas Dataframes, each of the tests were copied into individual csv files. This allows for each test to be loaded into a separate Dataframe without having to split them apart using Pandas operations. Column names in the tests were renamed to be more consistent with each other by taking the version that was the most concise. These first steps were done directly to the csv files, outside of Dataframe manipulation in Python. Now the data is ready to be loaded into Pandas Dataframes.

The data itself is very complete with only a total of 11-25 missing values for each of the sensors out of around 4900 rows for each of the tests. Removing NA values in the Dataframe was enough to get usable data. Duplicate columns were assigned a ".X" for each duplicate of the same column. A cursory look at each of the Dataframes shows no extreme outliers in the low temperature range or in the high temperature range when looking at the min and the max values.

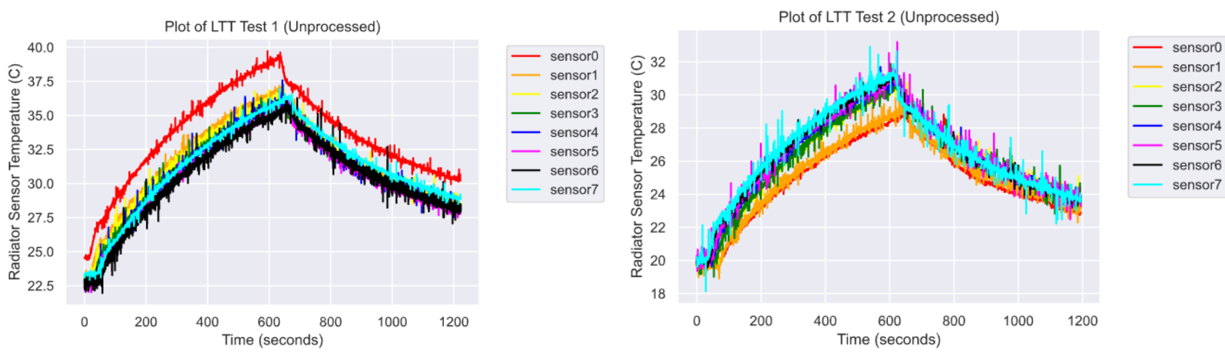

Figure 2. A cursory glance at two tests.

Figure 2 shows that while there are no major outliers in the tests, there is a fair amount of noise in the data. With this amount of noise, it is difficult to get a clear picture of what is happening in each of the tests.

2.3 Data Smoothing

Each of the tests contains around 1200 seconds of data, with 0.25 seconds in between each individual observation in each sensor. There is a high amount of observations which allows for LOESS smoothing to be viable. In this report we will explore both Kalman smoothing and LOESS smoothing and compare which provides a clearer picture and which may be closer to the actual temperature values.

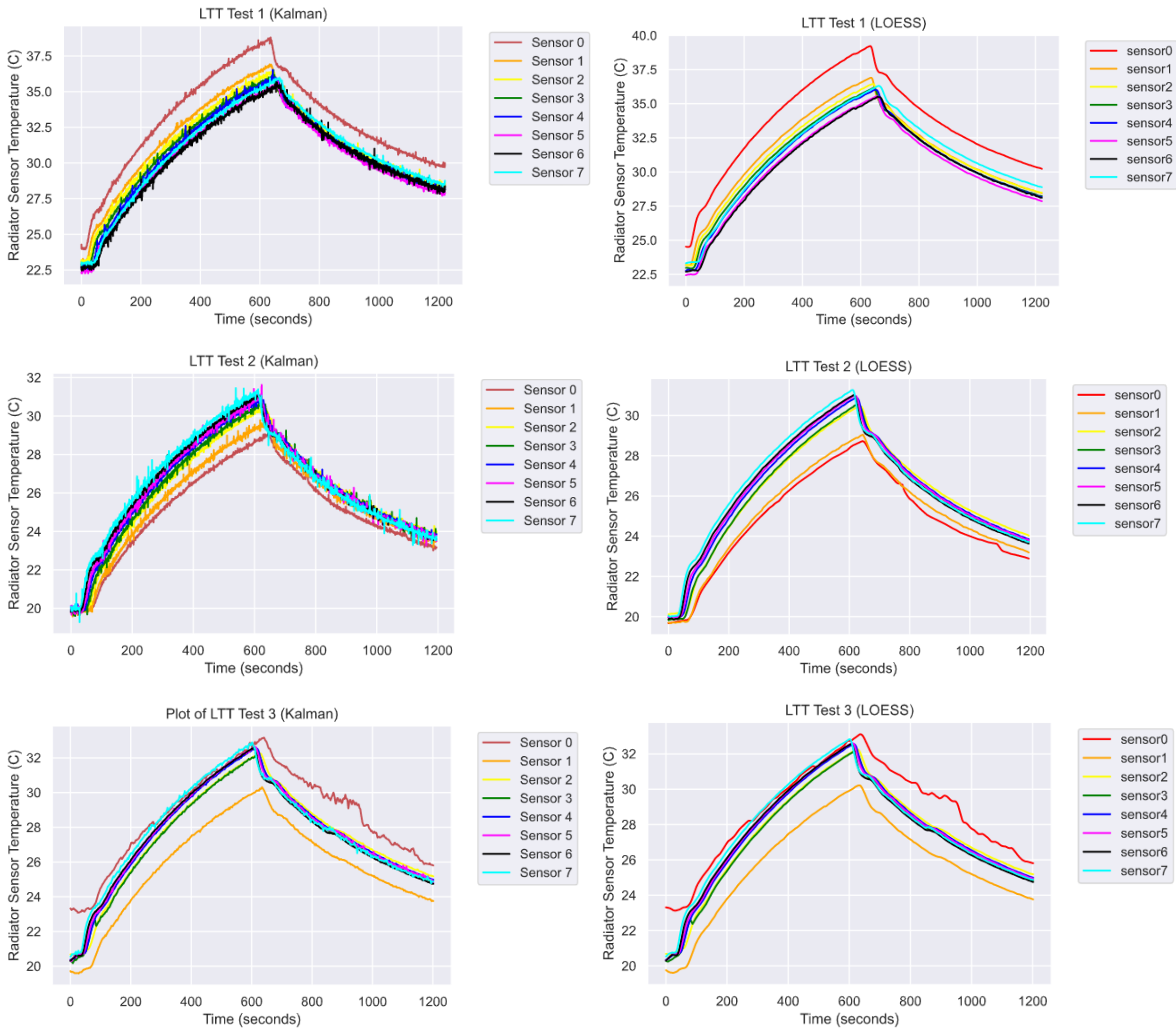

Figure 3. Comparing two types of smoothing.

In this comparison, the LOESS smoothing seems to provide a clearer picture of the temperature graph for tests 1 and 2, but Kalman performs nearly as well in test 3. This may be due to Kalman smoothing requiring much more precise tweaking of its parameters to be effective, while LOESS smoothing is simple to set up and provides a smoother graph due to the high number of observations. LOESS smoothing appears to create the results closest to the actual temperature, therefore statistical analysis will be done using the LOESS smoothed Dataframes.

The smoothed data shows clearly the effects of heat soak on the radiators further down, especially during the cooling down phase. Sensor 7 in test 1 LOESS (top right of figure 3) shows the effect most clearly, when the temperature of the radiator supposed to be the coolest goes above other radiators, showing that the hot air from further back is heating it up.

3 Results

3.1 Intentions of the Original Tests

To understand the conclusions that can be made from the data provided, we must understand the parameters and purpose of the original tests done. Each of the tests in the original Excel file has a set of different criteria designed to test out different hypotheses regarding the configuration of the radiators.

	Test 1	Test 2	Test 3	Test 6	Test 7	Test 8
Purpose	Test 8 radiators together.	Test 8 radiators together.	Unused by LTT.	Unused by LTT.	Test a single radiator.	Test 2 radiators together.
Config.	Hottest radiator gets air first.	Coldest radiator gets air first.	Hottest radiator gets air first (?)	Single radiator, three sensors.	Single radiator, two sensors.	Two radiators, three sensors.

Figure 4. Summary of each test corresponding to each csv file.

3.2 Difference in Temperature

In this section we will explore what information can be determined from the trends in the data by examining the difference in temperature in tests 1 and 2. The temperature delta provides a view of effectiveness from radiator to radiator as an alternative to simply viewing the temperature over time graph. If the final temperature minus the initial temperature is lower, then the temperature would have been lowered. This will allow us to find the point where adding more radiators to the setup would not be of any benefit. To find this point we will go sensor by sensor.

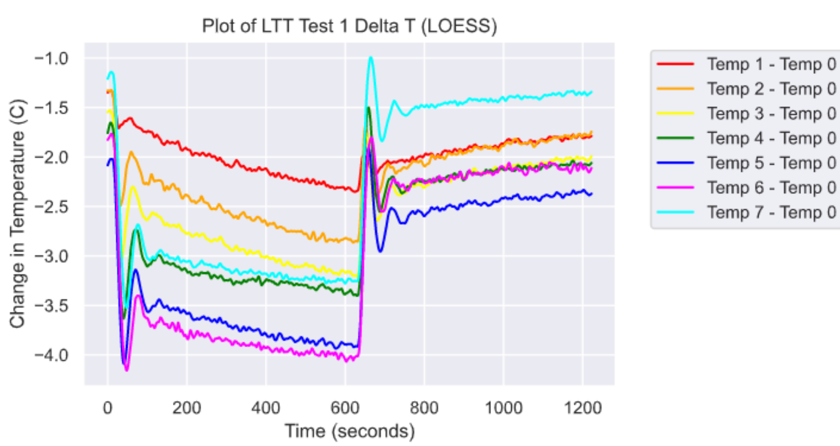

Figure 5. Change in Temperatures (Sensor 0)

Figure 5 shows that every radiator added after the first one provides a benefit, so looking at the delta T figures for radiators further up in the stack is needed.

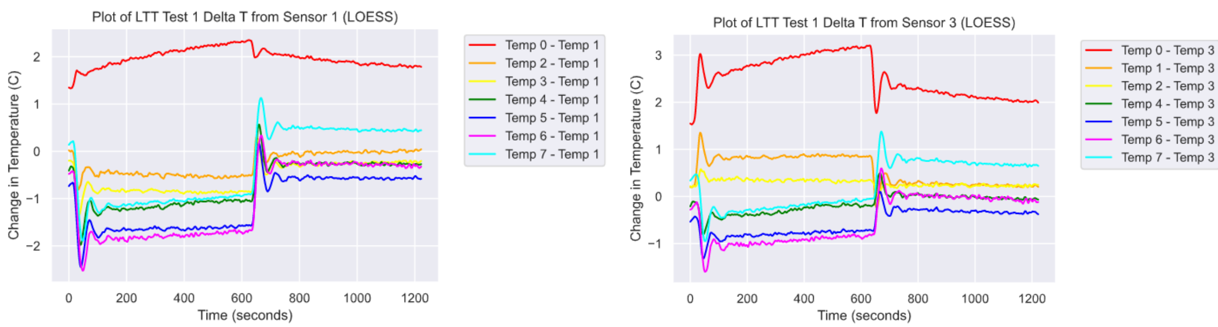

Figure 6. Change in Temperatures (Sensor 1, Left, Sensor 3, Right).

The difference between temperatures for temperatures beyond sensor 3 do not provide a strict improvement in temperature. For test 1 where the hottest radiator gets air first and flows towards the coldest, three radiators seems to be the point where the benefits are present throughout the test, while ignoring the efficiency penalty of diminishing returns. A similar comparison can be done for test 2.

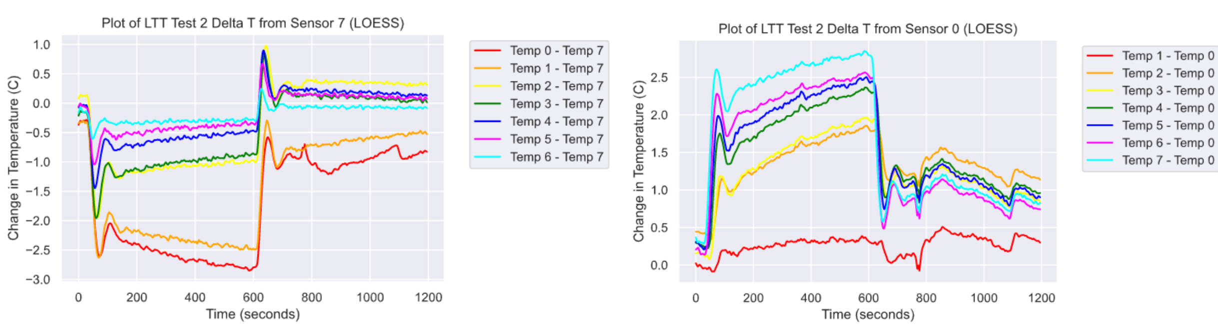

Figure 7. Changes in Temperatures (Sensor 7, Left, Sensor 0, Right).

A glance at the data shows that there is an improvement albeit diminishing when the coldest radiator gets air first. Figure 7 shows the two extremes from the first radiator in the stack to the last radiator in the stack. Having the coldest radiator receive air first lessens the effect of heat soak, although it is still present notably in the cooldown phase after 600 seconds where sensor 2 reports the highest temperature.

3.3 Statistical Tests

To find out whether the differences in the mean temperatures are significant from radiator to radiator, ANOVA was proposed to be used. However, ANOVA has assumptions that must be true in order for the results to be valid. One of the assumptions is each of the groups have equal variance. To show equal variance, Levene's test was used with the null hypothesis that the variances are the same. A Levene's test for test 1 and test 2 show that $p < 0.05$ for each, therefore we reject the null hypothesis, making ANOVA unsuitable for the data we have. An alternative to

ANOVA noted by [Scipy's documentation](#) is the Kruskal-Wallis test. It is similar to ANOVA, however it tests whether the medians differ rather than the means, which reduces the power of the test. The Kruskal-Wallis test does not have the assumptions that ANOVA makes, which makes this test suitable for our data. The p-value for the Kruskal-Wallis test was < 0.05, so we followed up with a post hoc analysis using Tukey's HSD.

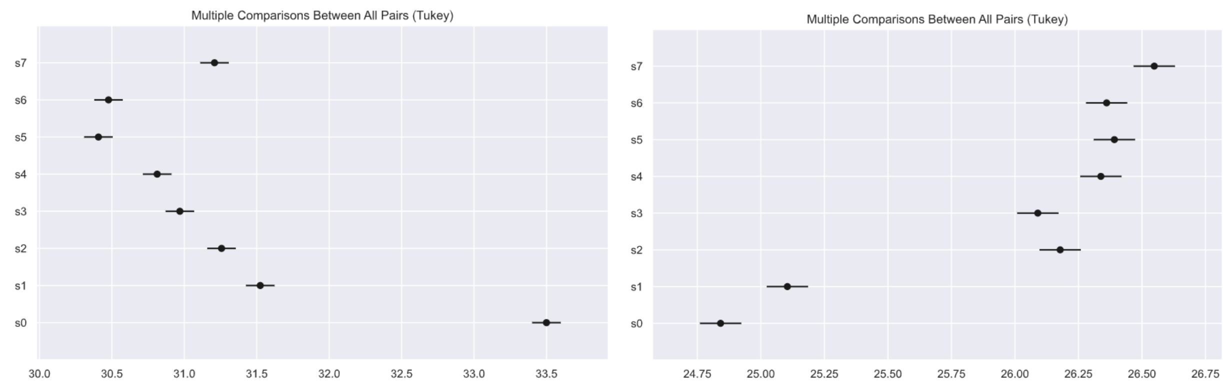

Figure 8. Pairwise Tukey for Test 1 (left) and Test 2 (right).

Figure 8 shows the post hoc test results for the LTT tests 1 and 2. For test 1, we can determine that there are different means notably in sequential sensor comparisons, for example sensor 0 to sensor 1, but ending from sensor 3 to sensor 4 since it is unknown whether the means are different. For test 2, the coldest radiator gets air first. That configuration has the potential to heat up air slightly, reducing the efficiency of each radiator after the first, but allowing for larger stacks and avoiding the heat soak effect. On the Tukey comparison, it is shown that beyond sensor 0 and sensor 1, the means afterwards are more tightly grouped together, making sequential comparisons difficult, however sensor 2 to sensor 7's means differ.

3.4 Conclusion

To conclude, we can determine that for a wind tunnel setup with stacked radiators and a similar hardware setup, if the radiators are arranged such that the hottest radiator gets air first, then 3 radiators is the most effective with minimal heat soak, while if the coldest radiator gets air first, then 2 radiators has the most impact, but it scales minimally with minimal heat soak 3 radiators onwards.

4 Discussion & Limitations

In retrospect, some analysis of the dual and single radiator tests would have established a baseline measurement for the analysis of many radiators. A test can be conducted between the three temperature values of the two radiator tests to make a conclusion based on statistical tests, even if the Dataframe graph and the Youtube video make the connection clear that two radiators have the most positive effect.

Regarding the quality of the data, it is worth noting there is no guarantee that the input wattage, stated to be 500 watts is consistent, therefore making the amount of heat going in not constant as well. This is caused by dynamic boost behaviour in modern CPUs and GPUs such as the Threadripper and Titan V used in the video, which will lower their clock speeds and power expenditure according to the temperature sensors on the chip and not the water or air temperature. Another factor to consider is firmware level power throttling from programs like Furmark. The 500 watt figure is presumed to be from the TDP figures of the Titan V (250W) and Threadripper 3970x (280W). However, TDP is not apples to apples from manufacturer to manufacturer. The air temperature measured is from an inconsistent source and may cause noise in the data or cause the curve to be flatter than it is and lower the maximum temperature. This likely will not harm the results in any significant manner, but a rerun of the test would be needed to confirm that claim.

A conclusion from the Linus Tech Tips video indicated possible flaws in their tests which causes the perceived benefits for stacking radiators. The mistake comes from where the system is not able to reach the equilibrium point, which makes the water temperature shown cooler temporarily. However, the water would eventually heat up again as the time goes on. The mistake comes from the increase in thermal mass within the radiators themselves as well as their water capacity.

5 Project Accomplishment Statement

Orion's Accomplishments

- Calculated statistical test figures using Kruskal-Wallis test and Tukey HSD to determine the effectiveness of radiator stacking in a wind tunnel system.
- Implemented visualization and analysis of statistical results by working with group members.

Wendy's Accomplishments

- Experimented with different methods of calculating running average (RA columns) used in the original datasets to determine a formula for the test 3 data.
- Comprehended the method of data acquisition and findings described in Linus Tech Tips's video and summarized them for the written report

Vincent's Accomplishments

- Coordinated with two other classmates to conduct an experiment to determine the effectiveness of stacking multiple radiators as a cooling solution for a water cooling loop with a CPU and GPU.
- Facilitated group discussions to achieve project's objectives.

6 References

Linus Tech Tips. (2020, July 6). *Is this even going to work…* [Video file]. Retrieved from https://www.youtube.com/watch?v=vauAJl29xlw