

# Optical Links Monitoring Setup

S. Billingsley, Southern Methodist University  
September 29th, 2022

## Setup

The optical data links for the ATLAS Liquid Argon Calorimeter are required to have a high efficiency and stable performance over a period of several years. To verify their stability, a test setup has been constructed consisting of 64 optical links located in Fondren Science Building 22a at Southern Methodist University. This setup is continuously monitored for the parameters crucial to the proper operations of the links. They need to be monitored to ensure the components have not been compromised in some way because these components were produced in the same batch as those installed in the ATLAS LArC. Should any of these components start to show a significant decrease in performance a supervisor can be alerted and forensic analysis can be performed to ensure the components in the ATLAS detector are not compromised as well.

To monitor the links, they are connected to a DC power supply that powers VCSELs used as a light sources that send pulse signals through the optical fibers to the receivers. The links are all connected to an optical spectrum analyzer (OSA). The OSA cycles through each of the links and analyzes the light output for the quality of the transmission by measuring power, wavelength spectra, and power at the peak of the spectrum. In the system, there are 64 channels that need to be monitored. To do this, a laptop controls a switch that changes which channel is analyzed through the OSA. Once the data is collected by the OSA, it is stored locally on a storage device and passed to SMU Box where the waveform can be analyzed and plotted by other programs.

## Data Structure and Storage

The data is stored directly from the OSA (raw data) on SMU Box in a library called "optical monitor raw data" and backed up on the local machine for redundancy. The raw data is stored as a text file with two columns: wavelength and optical power. The data is a histogram so the wavelength column describes the bins of wavelength in nanometers and the optical power column describes how much power in Watts is in the bin.

The raw data is organized in folders in SMU Box. The highest level is the folder "Waveform\_data1", then there are folders for each of the 64 channels, in each of these folders are folders for the date, and lastly the individual files. The file names carry information. An example of a file name is "waveform\_data\_channel1\_2022-09-12\_23-17-47.txt." This explains the channel it holds information for (in this case channel 1), the date and time, all split with an underscore.

This raw data is analyzed and uploaded to a Google sheet. It has a separate worksheet for each channel and another for each parameter where all the channels are concatenated so they can be compared. The structure of the individual channels is the daily average of 'mu', 'sigma', 'peak\_wavelength', 'peak\_power', 'total\_power', and 'total\_dBm'. The data is indexed by its date in the first column. Since there are about 100 timestamps per day for each channel the daily

average is taken for each value and that's the value is stored. This prevents any latency in uploading data, prevents hitting file size limitations, and the parameters of each channel do not fluctuate enough to require every timestamp to be analyzed.

## Software

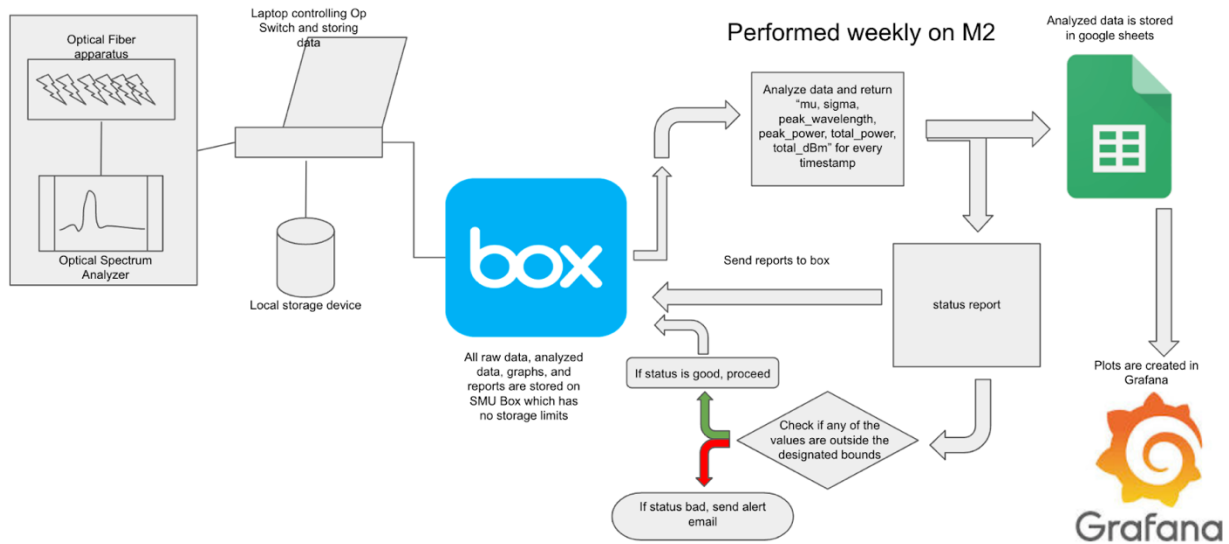


Diagram1. Flowchart detailing how the software works (explanation below)

## Analysis

The analyzing software allows the raw data to be quickly understood by the user to monitor the condition of the components. Since the raw data is a snapshot of the spectrum, the statistics of the waveform are reported. The script reads in all the titles of the existing snapshots and can recognize when the snapshot was taken from the title of the file. This saves time because the whole file doesn't need to be read in. This is crucial because the data is coming from Box and the only way to read the files is through FTPS which has a high latency per call. After reading the name of the file, the script reads the last date from the existing analyzed data in google sheets. It then drops all the snapshots that have already been analyzed. It then analyzes all the snapshots, organizes them into a dataframe, and takes the daily average of all the values. The analyzed data is then saved onto a google spreadsheet.

The script is connected to google sheets through an API. For privacy and so anyone in the group can connect, a new google account was made. To be authorized to edit files through the API, a service account had to be created in the Google console service and the Google Sheets API and Google Drive API needed to be authorized by the account holder. Once everything is authorized, the credentials for the service account are saved on the machine where the script is running and is pointed to using the pygsheets library. Using this library any spreadsheet can be accessed and any sheet within it can be returned as a dataframe.

To analyze data, if the file is older than or equal to the last analyzed day, it is read through FTPS from Box into a dataframe. X and Y values were extracted from the dataframe by setting the wavelength bins to be the X and the height of the bins (optical power) to be the Y. From here, the average wavelength, the standard deviation of the spectrum, brightest wavelength,

power at the peak of the spectrum, and total power in units of Watts and decibel-meters. The analyzed values are returned as a list for the corresponding timestamp for the file that the data came from. From here the data is resampled as the daily average and saved to the spreadsheet for that channel.

## Alerts

Using the same API for Box and Google Sheets, the analyzed data can be monitored to check if it is outside a value that is considered acceptable. On a daily basis, the alerting software will pull the analyzed data from Google Sheets and write a report for each value for each channel. These reports are uploaded to the “Reports” folder in Box. To determine the value of the report, a 7-day moving average is calculated for each value for each channel. After doing this, it creates an upper and lower limit from this average based on some parameters decided by the user. If the current value is outside these limits, when the report is written, it is saved to the “Bad Reports” folder as well as the “Reports” folder. On the report, the value that was triggered will be indicated. To receive a notification, the user must modify their Box settings to receive notifications when there is an upload to the “Bad Reports” folder.

## Automation

Since the raw data is constantly updated, it is necessary to analyze the data frequently. To do this without relying on someone to run the analyzing and alerting scripts and to maintain a consistent schedule, the scripts are run with CRON. CRON is managed on ManeFrame, SMU’s supercomputer. A simple shell script was created to run both python scripts. The scripts are scheduled to run every 8 hours to reduce the time it takes for the program to run. The longer the time in between runs of the analyzing script, the longer it takes to run. This is because the amount of data to analyze per channel increases. Since Box uses FTPS it takes a long time to read files from the website to the script.

## Plots

Since the goal of the solution is automatic monitoring, an online solution for plotting was used. The software chosen was Grafana. Grafana is a website that allows multiple users to log in and monitor dashboards. The dashboards can have a number of different panels to monitor different types of data. In our case, just plots are generated. The different plots that are created are all with respect to time and are the average wavelength, standard deviation, power at the peak, the wavelength at the peak, and total power in Watts and dBm of the waveform for all channels. The comparison of all these values can be seen in Figure 1. Then there are 64 additional plots, one for each channel. This is depicted in Figure 2. Each of these additional plots has all of the previously mentioned values that can be selected to be viewed separately. All of the plots are interactive as well, so the time range that is viewed can be modified as well as the size of the plot.

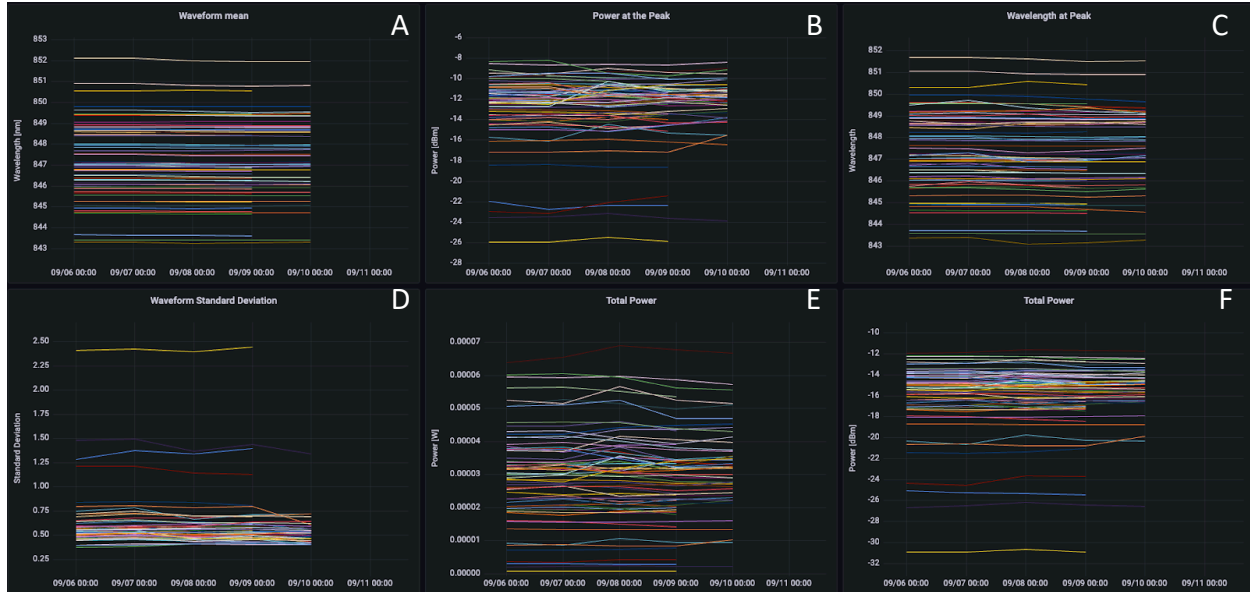


Figure 1. Plot A is the waveform mean with wavelength in units of nanometers on the vertical axis, plot B is the power at the peak with power in units of decibel-meters on the vertical axis, plot C is the wavelength at the peak with wavelength in units of nanometers on the vertical axis, plot D is the standard deviation of the waveform, plot E is the total power with power in units of Watts on the vertical axis, and plot F is total power with power in units of decibel-meters on the vertical axis.

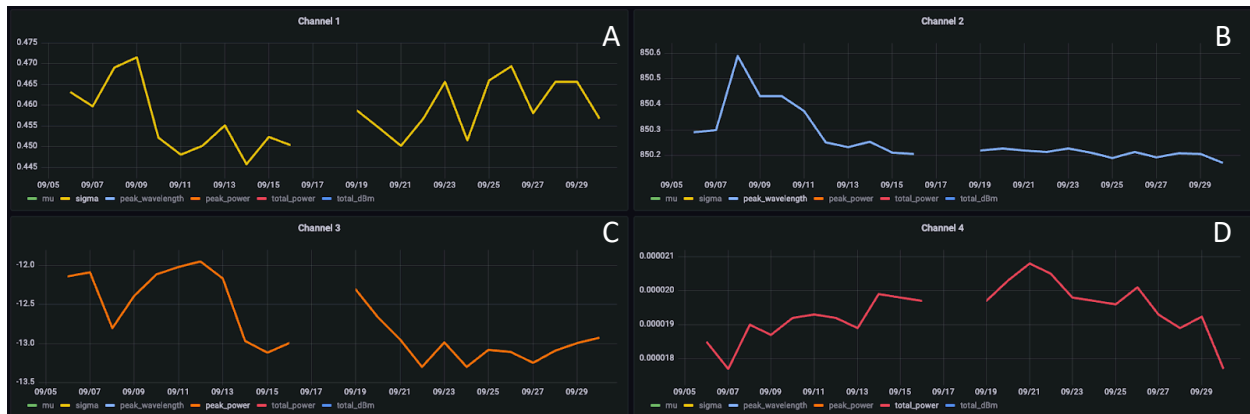


Figure 2. Shows an example of different parameters of 4 different channels all with time on the horizontal axis. Plot A shows channel 1's standard deviation of the waveform, plot B shows channel 2's peak wavelength with wavelength in units of nanometers on the vertical axis, plot C shows channel 3's power at the peak with power in units of decibel-meters on the vertical axis, and plot D shows channel 4's total power with power in units of decibel-meters on the vertical axis.

## Access

To modify or access any part of the software you must notify Sully Billingsley at [sbillingsley@smu.edu](mailto:sbillingsley@smu.edu). The software for analysis, alerts, and automation is stored and runs in the group directory `/work/group/spectrum_monitor` on SMU's maneframe. To have access to these packages you must be added to this group directory by contacting Sully, who will then contact OIT. For access to the plotting software, you can contact Sully and ask for the credentials to the Grafana website. To receive notifications from the alerting software contact Sully to be added to the SMU Box directory titled "optical monitor raw data". Once added your

notification settings must be modified to allow for notifications for uploads from the “Bad Reports” folder.

## Conclusion

A reliable and robust solution has been achieved to monitor the optical fiber components. As much of the process as possible is located on the cloud to allow for access to monitoring from anywhere in the world and to maintain consistency. The analyzing software is run on SMU’s supercomputer, all data is stored on either SMU Box or Google Sheets, and plotting is done on a cloud-driven website, Grafana. The solutions in SMU Box and Grafana now allow anyone with proper access to receive notification or look at the available plots with no prior knowledge of the system or needing to run software.