# I/O Performance Analysis Framework on Measurement Data from Scientific Clusters

Michelle Koo[1], Wucherl Yoo (Advisor)[2], Alex Sim (Advisor)[2]
[1]University of California Berkeley, [2]Lawrence Berkeley National Laboratory
michellekoo@berkeley.edu, {wyoo,asim}@lbl.gov

## INTRODUCTION

This project was motivated by the observations that I/O performance analyses can be conducted on scientific applications from monitored performance measurement data from scientific clusters. I/O performance behavior of the Palomar Transient Factory (PTF) application was observed by analyzing performance data collected on NERSC Edison. PTF is a wide-field automated survey that records images of transient objects in the sky. These images are sent to NERSC and LBNL for processing through the near real-time image subtraction data analysis pipeline, where time executions of each step in the pipeline were recorded. To analyze performance behavior of this PTF data analysis pipeline, an interactive I/O analysis framework was developed.

## RESULTS

Processing 1.6TB of system logs, Python and Apache Spark were used to sort measurement data from the PTF database into different Resilient Distributed Datasets (RDDs). iPython and Python's Matplotlib were used to create visualization tools to aid in determining hidden I/O bottlenecks. PTF was executed on compute nodes each with two 12-core CPUs: Intel Xeon E5-2670 and 64GB memory.

As shown in Figure 1, the PTF analysis pipeline consists of 38 checkpoints, and each bar represents the amount of time the PTF pipeline took on average each day. It is evident that the top three checkpoints with longest execution times are checkpoints 25, 31, and 36. Average daily percentage calculations taken over a span of 64 days reveal: checkpoint 25 takes 11.16%, checkpoint 31 takes 14.79%, and checkpoint 36 takes 23.72%.

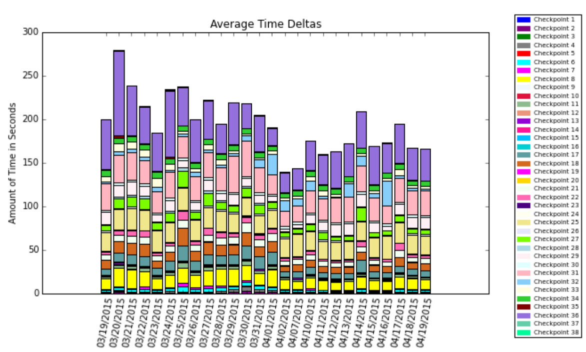


**Figure 1. Amount of time each operation takes**

From these findings, checkpoint 36 was found to have the longest execution time, and was further investigated as a bottleneck where performance should be improved. Checkpoint 36 involves the Transients in the Local Universe (TILU) query - a geometric query that correlates a table of known galaxies elliptical shapes and orientations with incoming candidates. Figure 2 shows the box plot of all timestamps of checkpoint 36 for each day.

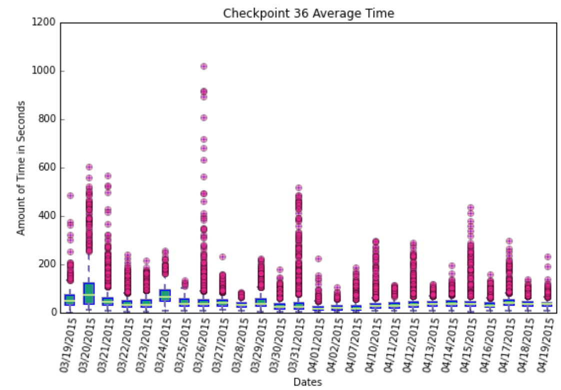


**Figure 2. Checkpoint 36 timestamps**

From Figure 2, notable dates containing many outliers with long execution times were further analyzed in Figure 3—showing a scatter plot of all timestamps of checkpoints 25, 31, and 36. The build up of execution times shown in Figure 3 after 12:20 P.M. reveals a possible system anomaly.

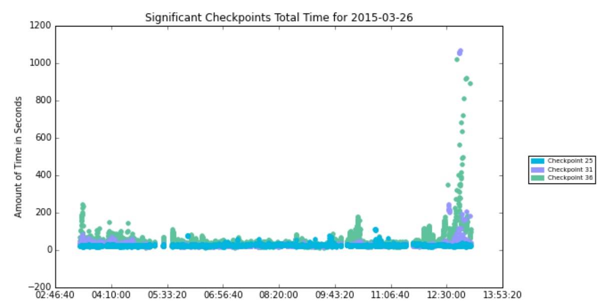


**Figure 3. All timestamps of checkpoints 25, 31, and 36**

To compare all three checkpoints with shared PTF field names with the longest execution times, timestamps associated with PTF field names and checkpoints 25, 31, and 36 were averaged, summed, and graphed in Figure 4.

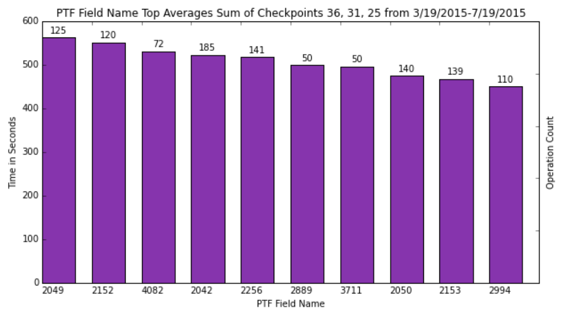


**Figure 4. Top PTF field names**

From Figure 4, PTF field name 2049 was calculated to execute the longest, and was displayed in detail in Figure 5.

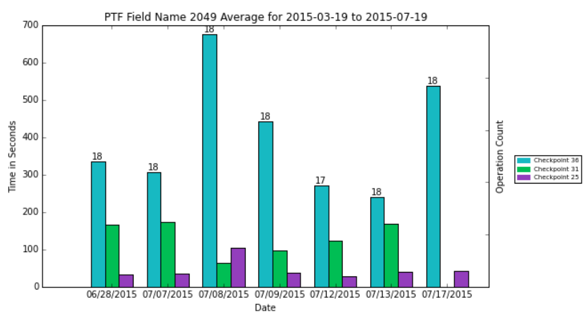


**Figure 5. PTF field name 2049**

These results led to the development of an interactive I/O analysis framework, pictured in Figure 6, which generates these figures by allowing users to enter start and end dates to query timestamps. The user has the option to choose between different types of graphs, checkpoints, and file types. When executed, a complete graph will be saved in the project directory and displayed below the user's queries. These experiments ran on a cluster with several hundred machines with two 8-core CPUs: Intel Xeon E5-2670 and 64GB memory. This tool allows users to interactively explore performance data for maximal understanding of their applications.

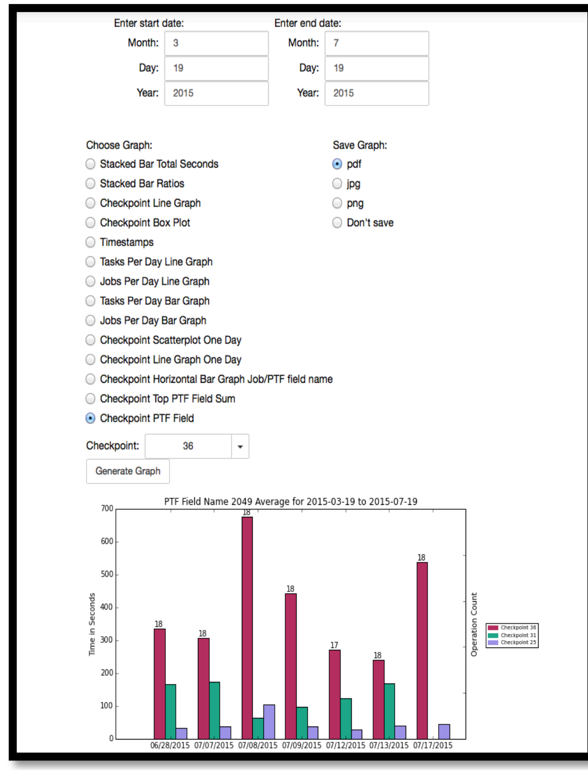


**Figure 6. Screenshot of the interactive visualization tool**

# CONCLUSION

Visualization tools were developed to analyze I/O measurement data from the PTF application on NERSC Edison to identify I/O performance bottlenecks. Results revealed checkpoint 36 as a bottleneck utilizing a quarter of the total execution time. Scientists inform that checkpoint 36 initially accesses "local" galaxies, which result in a high hit rate, but as PTF field counts increase, execution time slows, possibly causing this bottleneck. This analysis has led to two main observations: exposure of a possible system anomaly illustrated by increasing execution times in Figure 3, and that PTF field name 2049 had the longest average execution time. More research into the definitive reasons behind these observations will lead to improvement on the overall performance and efficiency of the PTF analysis pipeline.

This analysis contributed to the development of the I/O analysis framework tool, which allows users greater autonomy to extensively analyze I/O performance characteristics. It allows queries of different factors to research all possible sources of and uncover concealed I/O bottlenecks. The development of this tool will continue to advance research of I/O performance behavior characteristics and will be applied to other scientific applications in the future.

# REFERENCES


Nugent, Peter and Cao, Yi and Kasliwal, Mansi (2015) *The Palomar Transient Factory.* In: Visualization and Data Analysis. Proceedings of SPIE. No.9397.