Clever Title Goes Here

Shaun Brandt, Hema Kumar, Michael Novak, Mandar Patil, Yu Yang
Department of Computer Science
Portland State University
Portland, Oregon, United States

Abstract—As the need for high performance computing has increased, large data centers containing many individual servers (hundreds, or even thousands in some cases) have been constructed. The continuing increase in data center size has led to a greater need for sophisticated cooling solutions. However, providing efficient cooling for large clusters of servers at relatively low cost relies heavily on understanding the overall effectiveness of the cooling system. We present a set of custom tools and extensions to an existing tool that provide useful analysis of data center cooling performance. Additionally, we use data from an existing server cluster to demonstrate the effectiveness of these tools.

Keywords; power, cooling, integrated analysis, clustering, Perftrack

I. INTRODUCTION

As the use of Internet-connected devices has grown and the usage of the Internet by connected users has become more sophisticated, the infrastructure required to support these users has grown increasingly larger and more complex. Companies like Google [6], Apple [7] and Amazon [8] have deployed a number of large computer clusters, called data centers, throughout the world. While data centers provide a powerful solution to the problem of scaling computing resources to meet the needs of users, such environments introduce a number of unique difficulties. Among these is the ever-increasing power and cooling requirements that are an inevitable consequence of placing so many servers together in a single location. While many data centers have been constructed in locations close to cheap, easily available sources of power, lower electricity costs are not the only concern. Keeping hundreds, or even thousands, of servers cooled effectively requires increasingly innovative solutions. Poor management of cooling systems can lead to undesirable cost overhead, and can limit the extent to which a data center can be expanded. Therefore, the ability for data center operators to easily and effectively monitor cooling and performance is essential.

In this paper, we discuss methods that we have developed to provide solutions for metric collection and analysis of cooling data in a data center. The first involves custom software that generates graphs and tables based on metrics that cover key subsets of collected data, while the second modifies an existing data analysis tool, called Perftrack [1,2], to add support for data clustering. While these two tools currently work somewhat independently, we believe that future work could tightly couple the two approaches, making a complete analysis tool set that's fully integrated into Perftrack.

The remainder of the paper is organized as follows: Section 2 provides an introduction to the environment where the data used for analysis has been collected, as well as information about the data itself. Section 3 discusses the first approach used for analysis, based on custom scripts, and covers the metrics presented in the resulting tables and charts. Section 4 discusses modifications to the Perftrack tool that add support for clustering of cooling data using off-the-shelf clustering software. Section 5 presents results of the two approaches and discusses how the methods outlined in Sections 3 and 4 could be combined to produce a single Perftrack-based tool for data analysis. Section 6 discusses opportunities for future work, and Section 7 concludes. An appendix containing a generated sample report is also provided.

II. THE COMPUTING ENVIRONMENT AND DATA

To effectively analyze the cooling and power usage of a data center, a large collection of power and temperature sensors is needed to collect relevant data. For the analysis presented in this paper, we used data collected from PNNL's Energy Smart Data Center. This is the same data used by some previous analysis work [1,2].

The server cluster at the Energy Smart Data Center consists of multiple groups of server racks. In our analysis, we concentrate on eight particular racks, labeled A1 through A4, and B1 through B4. Of these 8 racks, one (A2) is devoted solely to network equipment, and therefore is not directly addressed in this paper. Of the 7 remaining racks, two (A3 and A4) use air cooling, while the remainder use a spray cooling method that introduces a non-conductive liquid (Fluorinert, in this case) directly onto the components being cooled [1]. These spray-cooled racks include thermal management units (TMUs) that contain pumps and a heat exchanger, allowing excess heat from the server racks to be removed. A set of chiller units provides a method of extracting heat from the data center as a whole. Chiller units consume a large amount of energy and are also subject to physical wear over time [4], so monitoring the effectiveness of these units is especially important.

The collected data spans two separate, short time periods: one from April 25, 2011 and one from May 5, 2011. Two different levels of workload, referred to as 'low density' and 'high density', are included. The low density workload was designed to spread a set of processes evenly among multiple racks, with relatively few processes running per server, while the high density workload runs a similar set of processes using only the servers from a single rack.

While the data covers a relatively short time span compared to data used by related works [3,4], we were still able to extract useful information. Unfortunately, some of the sensor readings, especially those related to server component temperatures, were invalid. This limited the set of metrics that we were able to analyze. This limitation is discussed briefly in Section 6. As a result, we concentrate on a few key metrics involving power usage and cooling efficiency of the center's chillers and TMUs.

III. Analysis Using Custom Tools

A. Parsing and assembly of raw data into graphs and tables

To process the data in an effective manner and to allow the process to be automated, a series of scripts was written in Haskell, though earlier prototypes include components written in Perl, C and shell script. The current version of the tool performs all data parsing using source files in CSV format, either generated by Perftrack or obtained from another source, creating output suitable for passing to Gnuplot. Graphs and tables are then generated, including cluster graphs (described further in section 3.2). The resulting image files can then be embedded into the final analysis report.

B. Clustering

The use of data mining techniques to analyze the effectiveness of cooling techniques in data centers has been explored in previous work [3,4]. However, these data mining approaches rely on the availability of data spanning a large amount of time. Additionally, the approaches addressed in those papers use a technique called motif analysis, which divides the data into larger time slices, allowing clustering based on these slices rather than individual data points. Unfortunately, insufficient data is available in our data sets to use motif-based clustering, so we use individual data points as the items to be clustered.

For both low-density and high-density loads, we use the following metrics for clustering:

- Cooling load. This is a measure of the amount of heat generated by a data center. Since almost all power consumed by installed equipment is dissipated as heat, this number is approximately equivalent to that equipment's power consumption. The standard unit of measure for cooling load is the kilowatt.
- Coefficient of Performance (COP). The COP of a chiller unit measures how efficiently the unit provides cooling, and is defined as the ratio between the amount of cooling provided and the amount of power consumed to provide that cooling. Each chiller has a separate COP value.
- Temperature deltas. The chillers and condenser units all monitor inlet and outlet water temperatures. With this information, we can calculate water temperature deltas (the difference in temperature between the inlets and outlets).

An open source tool called ELKI [5] is used to perform the actual clustering analysis. A list of data points generated by the Haskell scripts is provided to ELKI, which returns a set of clusters, each containing a subset of the data points. Each point is also listed with a time slice; each row of the output, corresponding to a single data point, lists the trial and time slice corresponding to that point.

The graphs generated in Section 3.1 are then modified in a novel way. We have created a script to sort the cluster data by time, resulting in a sorted list of time slices and the cluster to which the time slice belongs. The original graph images are then further modified by Gnuplot, highlighting each time slice using a color allocated for each individual cluster. The resulting graphs, which we call time-correlated cluster graphs, show which cluster ELKI assigned to each particular time slice, giving a quick overview of what points are considered similar. We believe that this display technique could be extended to the work done in [4], allowing data center managers to quickly identify individual motifs within the original time-based data.

C. Automatic report generation

As a final step, all of the charts, tables, and time-correlated cluster graphs are assembled together into a single report using LaTeX, and the resulting document is rendered in PDF format. In addition, information about the metrics used and data collected is placed at the beginning of the report. The resulting document is an essentially complete, easily published paper that provides information about the environment, results, and cluster analysis, with no human interaction required. This script can even be run automatically, allowing reports to be assembled and generated on an ongoing basis (as a cron job, for example).

IV. Modifications to Perftrack to Support Clustering

V. Results

VI. Future Work

The analysis performed in this paper was done after-the-fact, based on data generated approximately one year ago. This data was provided in a relatively simple format (CSV), but the Energy Smart Data Center automatically collects this data into a performance monitoring database called FRED [1]. An optimal implementation of the data analysis features described in this paper would be based on a direct connection to the data placed into FRED, and the analysis results could then be generated in real time.

In lieu of a direct connection to FRED, the work done in this paper could be improved upon by the availability of additional data in the formats we have already used; however, the amount of data required to make analysis fully effective would be measured in days, if not longer. The clustering results we generated are somewhat mixed due to the short time span of the data, and we believe that further experimentation with long data sets would generate more useful results. Also, the data that was available had many data points that were either invalid or repeated, presumably due to malfunctioning sensors or incorrect collection and storage of sensor results. This missing data includes all of the temperature metrics for the server hardware components, including CPU, memory and ASIC temperatures. If these values were present, we could extend the work we have done to measure efficiency of the TMUs and chillers relative to the overall effectiveness of the cooling provided.

Additionally, we would like to extend the time-correlated cluster graphing techniques introduced in Section 3 to work with the motif-based analysis performed by Patnaik et al. [4], but unfortunately the results of their work are no longer available, so the analysis tools written by their group would need to be rewritten.

Finally, the ability to access and modify Perftrack was relatively limited. The code is based on older libraries that are not easily available on our test systems, so the ability to add features to the Perftrack code base was only available late in the project. Given suitable time, we could extend the tool to integrate all of the features from Section 3. The resulting tool would be easier to use than the custom scripts that we generated, and the database connection would ensure that the source data used for analysis and clustering is always up to date.

VII. Conclusion

References

[1] R. L. Knapp, K. L. Karavanic, S. Krishnamoorthy, and A. Marquez, “Power- and cooling-aware parallel performance diagnosis,” Proceedings of the IASTED International Conference on Parallel and Distributed Computing and Systems (PDCS 2011), December 2011, pp. 435-442.

[2] R. L. Knapp, K. L. Karavanic, and A. Marquez, “Integrating power and cooling data into parallel performance analysis,” Proceedings of the 2010 International Conference on Parallel Processing Workshops (ICPPW 2010), pp. 489-496.

[3] D. Patnaik, M. Marwah, R. K. Sharma, and N. Ramakrishnan, “Data mining for modeling chiller systems in data centers,” Advances in Intelligent Data Analysis IX, 2010, pp. 125-136.

[4] D. Patnaik, M. Marwah, R. Sharma and N. Ramakrishnan, “Sustainable operation and management of data center chillers using temporal data mining,” Proceedings of the 15th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD '09), July 2009, pp. 1305-1314.