

Application of Visual Analytics For Thermal State Management in Large Data Centers

M. C. Hao¹, R. K. Sharma¹, U. Dayal¹, C. Patel¹, R. Vennelakanti¹, D. A. Keim²

¹ Hewlett Packard Laboratories, Palo Alto, CA

² Department of Computer and Information Science, University of Konstanz, Germany

Abstract

Today's large data centers are the computational hubs of the next generation of IT services. With the advent of dynamic smart cooling and rack level sensing, the need for visual data exploration is growing. If administrators know the rack level thermal state changes and catch problems in real-time, energy consumption can be greatly reduced. In this paper, we apply a cell-based spatio-temporal overall view with high-resolution time series to simultaneously analyze complex thermal state changes over time across hundreds of racks. We employ cell-based visualization techniques for trouble shooting and abnormal state detection. These techniques are based on the detection of sensor temperature relations and events to help identify the root causes of problems. In order to optimize the data center cooling system performance, we derive new non-overlapped scatter plots to visualize the correlations between the temperatures and chiller utilization. All these techniques have been used successfully to monitor various time-critical thermal states in real-world large scale production data centers and to derive cooling policies. We are starting to embed these visualization techniques into a handheld device to add mobile monitoring capability.

CR Categories and Subject Descriptors: I.3.3 [Computer Graphics]: Picture/Image Generation – Display Algorithms; H.5.0 [Information Systems]: Information Interfaces and Presentation – General

1. Introduction

Cooling is the major operational cost in a data center. Excess heat adversely affects uptime and equipment performance and leads to a premature end of life for equipment. Thermal state management is crucial for a cost-effective and sustainable operation of large scale data centers to support the growth in computing demand [SH08, BSP06].

1.1 Background

Figure 1 displays a typical data center air-conditioning environment with under-floor air distribution. Computer Room Air Conditioning (CRAC) units cool the exhaust hot air from the computer racks. Cooling infrastructure comprises chillers, cooling towers, computer room air conditioning units and primary/secondary pumps. Data center's IT infrastructure includes servers, network and storage devices housed in standard racks.

There are usually five inlet sensors (blue) used for measuring incoming air temperature to the server racks as well as five outlet sensors (red) for measuring the out-going air temperatures from a server rack. The sensors are arranged in descending order with sensor T1 closest to the vent tile within a rack. In a normal situation, the sensor air temperatures have an ascending order ($T1 < T2 < \dots < T5$).

Energy consumption in data center cooling depends on the distribution of cool air to the racks and the extraction of heat from the hot exhaust air. A refrigerated or chilled water cooling coil in the CRAC unit extracts the heat from the air and cools it to a range of 10 °C to 18 °C.

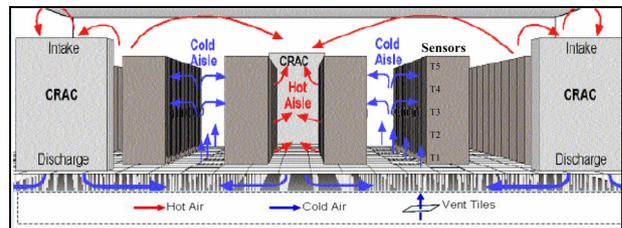


Figure 1: A typical raised-floor data center

The sensors are arranged in ascending order with sensor T1 closest to the vent tile within a rack. The sensor air temperatures should be always kept in an ascending sequence $T1 < T2 < T3$.

The air movers in the CRAC units pressurize the plenum with cool air which enters the data center through vented tiles located on the raised floor close to the inlet of the racks. Typically the racks are laid out in rows separated by hot and cold aisles. This separation is done for thermal efficiency considerations. Air inlets for all racks face cold aisles while hot air is expelled to hot aisles. A number of other equipment layout configurations and non-raised floor infrastructures also exists. Chillers provide chilled water to the data center room that houses the server racks and other IT equipment.

Most data centers have four critical problem thermal states: (1) high sensor temperatures (e.g., temperatures exceeding 35°C), (2) out-of-bound temperature-based metrics such as Supply Heat Index (SHI) [SBP02], (3) out-of-sequence sensor temperatures within a rack (e.g., $T1 > T2$ in Figure 1),

and (4) abnormal chiller utilization. Detection of hot spots is necessary to prevent temperature related failures while detection of abnormal air flows and metrics helps to identify energy inefficiencies. There were no tools available for administrators to reveal interactions between sensors in real-time, which are especially needed after a reorganization of racks, computers, chillers, or CRAC units. Detecting these thermal problems usually takes more than one day per 100 racks. For a large size data center with thousands of racks, the cost can easily reach more than 100 man hours. Since most thermal problems are sensitive to physical location, an onsite deployable tool is of great advantage to administrators. Onsite problem identification and escalation prevention is much easier with a mobile platform.

Currently the application of visualization techniques to the data center infrastructure is rudimentary and limited to the representation of geometry and modeling/design issues. Also, visualization techniques have not been applied to real-time thermal management data from sensors. Current approaches and guidelines adopted to address energy efficient thermal management include numerical modeling and parametric studies. Expedient visual analytics and evaluation of immense amounts of data are needed.

1.2 Our contribution

Our goal was to enable data center administrators to identify and quickly correct abnormal thermal states to meet sustainability targets such as energy efficiency, emissions, and cooling optimization. We have developed an interactive management system that allows administrators to detect and mitigate time-critical thermal events. This system provides a spatio-temporal overview of operations data with visual analytics capabilities and queries. In addition, we are integrating these visual analytics techniques to HP iPAQ devices to provide mobile capabilities for onsite evaluation.

In order to optimize the data center cooling system performance, we derive new non-overlapped scatter plot techniques to visualize the correlation between the temperatures and chiller utilization. We also cluster the temperatures with similar chiller utilization for data center managers to identify different patterns of cooling operations.

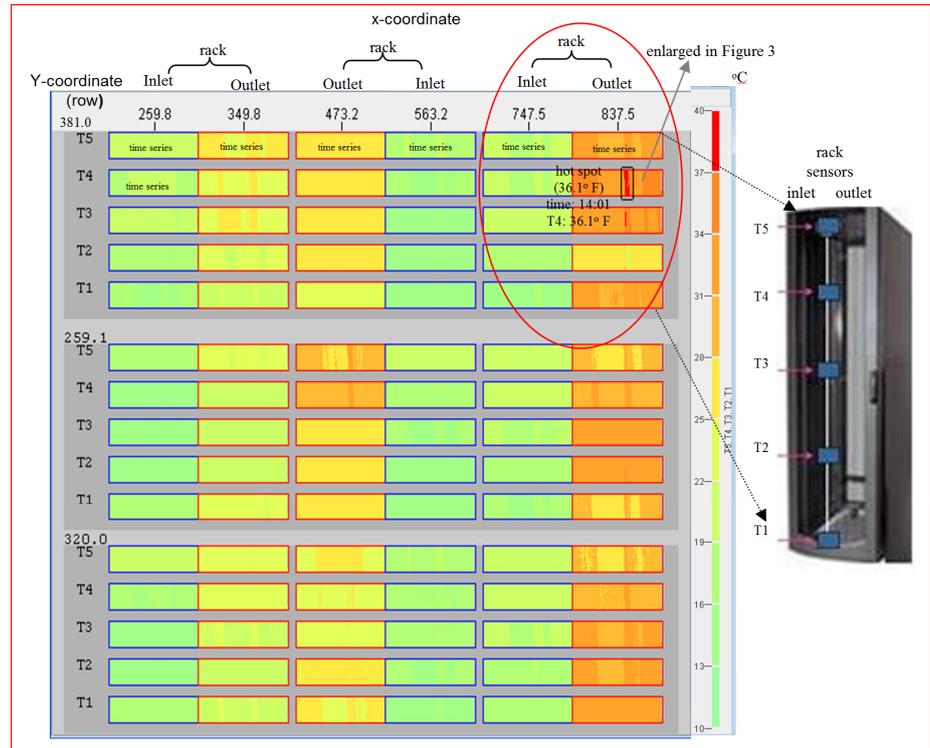


Figure 2: A spatio-temporal overall view

(x-axis: rack aisles; y-axis: rack rows; rectangle: temperature time series; color: temperature)
Shows an overview of 9 racks residing at location (x, y). Each rack has 5 inlet sensors (T1-T5) with blue border and 5 outlet sensors (T1-T5) with red border. Each sensor has a temperature time-series.

2. Related work

Early interactive spatio-temporal displays were designed by Andrienko et al. [And95]. Andrienko successfully developed a series of interactive spatio-temporal visualization techniques for displaying values of numeric attributes referring to different moments in time and locations in space. Their work focuses on user-controlled data transformation aggregation and dynamic linking of data displays. Andrienko's recent work [And08] uses novel ways of aggregating movement data for visual analysis, providing both traffic-oriented and trajectory-oriented views of the data. Andrienko employs temporal histograms and clusters to show results of movement positions and various trips. The framework combines visualization with database techniques for handling large datasets which is similar to our current work on spatio-temporal thermal data. Both Andrienko and our visualization techniques enable human cognition and reasoning to direct and control further analysis.

Compieta et al. [CMBF07] proposed a data-mining system to explore large spatio-temporal data sets. They have developed 3D visualization techniques to display the mining outcome, such as displaying association rules and variable distributions. Chen et al. [CXGH08] used advanced filters, aggregations, and pre-fetching techniques to visualize more than a billion records with commodity hardware.

To visualize multi-attribute data sets, Seo [SS05] developed the rank-by-feature framework. The low-dimensional projections such as histograms and scatter plots are useful for understanding the multi-dimensional data sets.

Van Wijk et al [WS99] developed a calendar-based visualization of large time series data combined with a clustering technique for users to quickly identify changes over time. Munzner’s BinX [MUN04] can dynamically visualize time series across aggregation levels. Moore et al.’s Knowledge Planes [MCFR04] are designed to monitor and manage data computing infrastructures and events in data centers.

In contrast to the above mentioned techniques, we focus on the exploration of sensor temperature distributions and thermal state changes in real-time without aggregation. Also, we emphasize on how to visualize the relationships among multiple attributes (e.g., racks, inlet and outlet sensors, temperatures, chiller utilization, etc). For dealing with hundreds of sensor time series simultaneously, we add high-resolution and automatic analysis to our previous multi-resolution time series [HDKS07], so that data center administrators can simultaneously analyze high-volume temperature changes.

3. High resolution temperature visualization

In this paper, we introduce different cell-based visual analytics techniques with visual queries and anomaly markers to detect various thermal states. Data center administrators can visualize and interact with these techniques and by selecting certain racks or sensors they are able to incorporate their domain knowledge in real-time.

A spatio-temporal overall view

Figure 2 illustrates a sample mapping of nine racks of multi-dimensional thermal data to a 2-D space. We provide both an overall view for the data center thermal performance and interactive “data exploration” views for troubleshooting. Each rectangle contains a temperature time series for one sensor. Hot spots are marked automatically for thermal correlation analysis and root-cause queries.

High resolution time series

To detect abnormal operating conditions in a large thermal data set without overplotting, we use a high-resolution color cell-based time series to plot actual values over time. In Figure 3, cells of a time series are arranged from bottom to top and left to right according to time. The size of the cells is automatically scaled down as more cells are displayed. The color of a cell is the temperature value of a sensor for each time interval. Specific colors (i.e., red, orange, etc.) represent more severe conditions and alert the administrator to take immediate action. For example, sensor T4 has a persistent high temperature (over 35°C) starting at time 14:01 in the marked area. Users can easily mouse over a cell and display the time and temperature value of that cell. Also, users can query on a marked area for what happens in a certain time period.

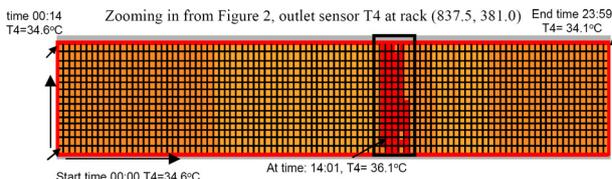


Figure 3: A high-resolution time series with a marked anomaly (96 X 15 cells; at 1-minute intervals; color: temperature)

4. Abnormal state detection

We apply two new visual analytics techniques, Marker [HDKSM09] and Correlation Visual Query [HDKMS07], to visualize the thermal state problems. The Marker, as shown in Figure 3 (sensor T4) and Figure 5 (e.g., CB0619.T2), automatically draws a rectangle around the abnormal data points and their surrounding areas when the points exceed a certain user specified threshold. These markers guide the user to find dangerous conditions and identify problems quickly. Administrators can choose different color maps to represent temperatures as needed. A rainbow color map is commonly used in most data centers with blue for low, orange for intermediate, and red for hot temperatures. To distinguish these three different states, in Figures 5 and 6 a non-uniform color map is used for displaying the uneven high density dataset [HD06].

4.1. Visualization of hot spots

After locating the anomalies, the operator can inspect the marked areas and perform a visual query. The query invokes correlation mining techniques to compute the correlation between the attributes of the marked data with all other attributes in the data set. As a result of the visual correlation query the two sensor temperatures with the highest positive or negative correlations are shown. The operator can learn from the query results, which other servers have a highly correlated workload and refine the workload accordingly. Figure 4 shows the top two sensors T1 and T2 which are more related to T4 than T3 and T5 (not shown). The following observations can be made:

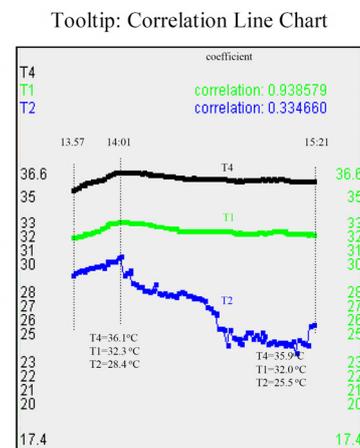


Figure 4: Correlation within a rack T4 is highly correlated to T1 with a correlation coefficient > 93%. T4 is less correlated with T2. T2 has a drop in temperature which may be caused by a reduced server workload or an air flow change. The correlation coefficients of T3 and T5 with T4 (not shown) are lower than 0.3.

1. T4 (black line) has a persistent high temperature above the threshold ($>35^{\circ}\text{C}$) over a long period of time from 13:57 to 15:21.
2. T1's (green line) temperatures are highly correlated with T4. Both T1 and T4 temperatures change at the same pace. There is a similar air flow in T1 and T4.
3. T2's temperature (blue line) drops after 14:01. The administrator knows that this may be caused by either a change in air flow or workload in the servers around T2.
4. All other temperatures have a lower correlation and are therefore not relevant for further inspection.

Note that the correlation analysis compares the selected temperature with the temperatures of all racks and shows the two sensor temperatures with the highest positive or negative correlations and therefore helps administrators to also diagnose the thermal events across racks.

4.2. Visualization of metrics

Metrics are used to describe the physics behind the process responsible for the thermal states. In Figure 3, we show the sensor temperature changes over a time period to observe each sensor temperature patterns, trends, and outlets. In Figure 5, we visualize an hourly Supply Heat Index (SHI) metric. SHI is a measure of recirculation of the exhaust air stream at the racket inlet to capture the mixing of hot and cold air near the racks [SBP02]. High SHI can lead to high energy consumption in data centers. Each colored cell represents a metric value at a 1-minute interval. Visualizing the evolution of metrics using color is better than simply plotting the actual values over time to avoid overplotting while visualizing large data volumes. Figure 5 shows many SHI values (red) which are beyond the normal bound of 0.85 and are marked on rack CB0619 sensors T2 to T5. With this information, administrators can quickly identify sensors with high supply heat index values indicating regions of high recirculation and then make the needed adjustment to the racks.

4.3. Out-of-sequence temperature detection

Within a rack, the sensor air temperature should be always kept in an ascending sequence (e.g., $T1 < T2 < T3 \dots$) from bottom to top. Out-of-sequence air temperatures are usually caused by servers not being stacked properly or by gaps between servers leading to the mixing of hot and cold air. To identify these problems and to improve energy efficiency, we use a pair-wise color cell comparison matrix. Each cell represents a two dimensional vector (location, time). A maximum delta value (e.g., $T1 + \Delta^{\circ}\text{C} < T2$) is used to automatically determine racks with out-of-sequence sensors and show them in the visualization (Figure 6). The delta

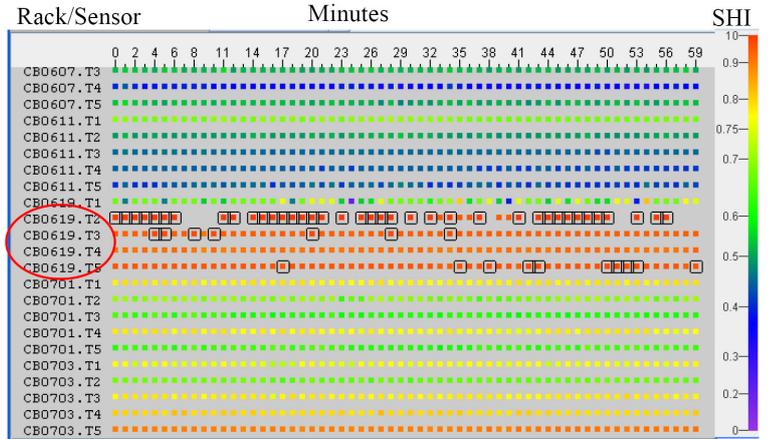


Figure 5: A Supply-Heat-Index (SHI) hourly metrics map
Each cell represents a 1-minute interval, the color of a cell is the SHI value.
Rack CB0619, Sensors T2, T3, T5 with SHI >0.85 are marked.
(x-axis: minutes (0-59) only intermittent values are shown; y-axis: rack)

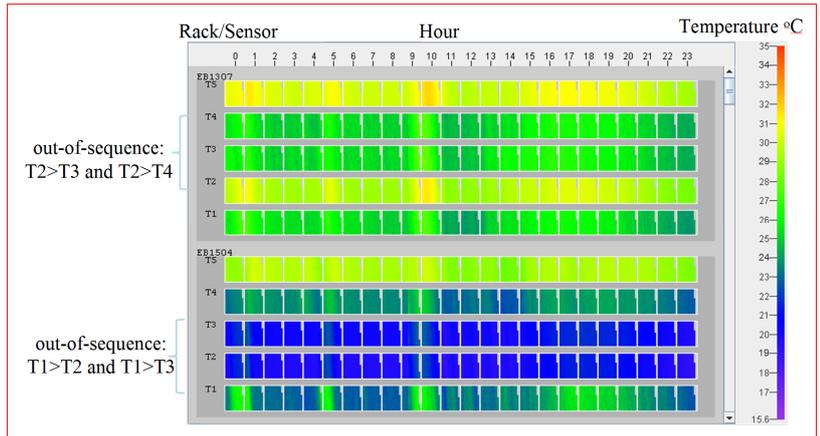


Figure 6: Visual time series for finding out of sequence sensors
(x-axis: hours 0 to 23, y-axis: rack/sensor T1 to T5;
color: temperature with a non-uniform color scale to show dense data)

value can be dynamically adjusted at run time. In Figure 6, the sensors in rack EB1307 and EB1504 have a number of out-of-sequences events, such as the temperature of rack EB1307 T2 (yellow) is higher than T3 and T4 (green).

Using this method, administrators are able to identify and correct the air flow of hundreds racks in minutes with energy savings of up to 20%. Most importantly, users can see temperature out-of-sequence distributions. Prior operation had to copy the temperature values, paste them into a spreadsheet sheet, and then highlight the abnormal sensors. This process usually took several hours and is prone to error.

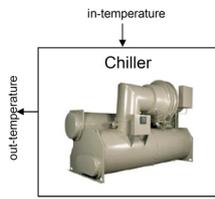


Figure 7: A Chiller
The chiller utilization efficiency is a function of in-temperature T_{in} , out-temperature T_{out} , and utilization %U.

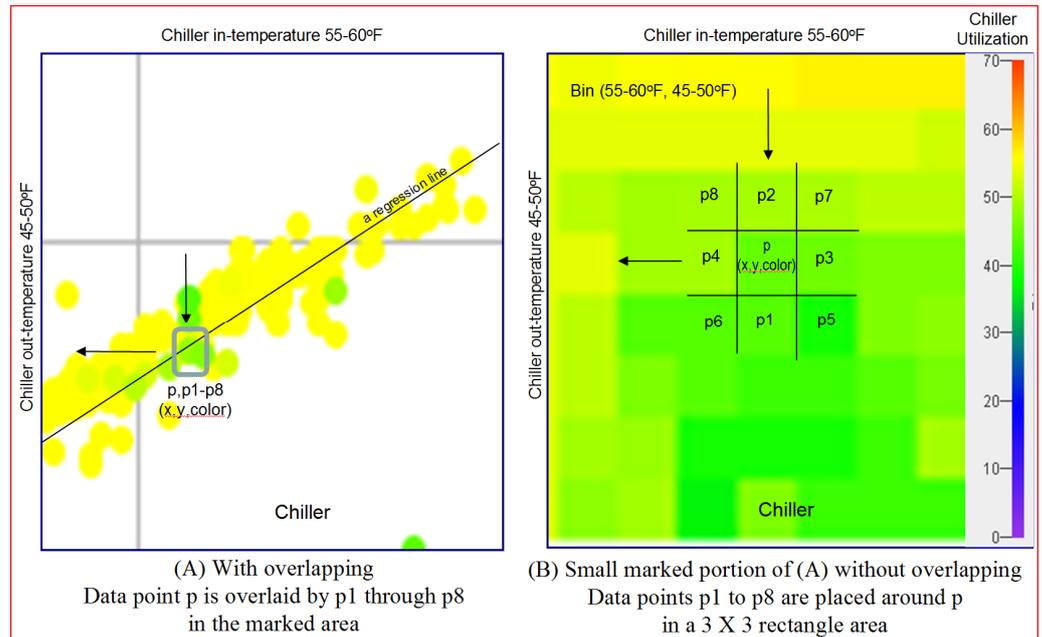


Figure 8: Chiller utilization scatter plots (with and without overlapping)
Chiller data points p1 to p8 have the same location as data point p (T_{in} , T_{out} , %U) with different values of %U (e.g., 40-50%, from green to light green) in the rubber-band area.

5. Chiller utilization scatter plots

Chillers (Figure 7) extract heat from the warm water and provide cold water to the CRAC units in the data center. The extracted heat is ejected into the environment through cooling towers or heat exchangers. Evaluation of the efficiency of the chillers and the estimation of operational uptime before the next maintenance downtime are crucial tasks. Visual monitoring of the utilization of chillers can greatly reduce operating expenses and downtime.

Examples of a data center manager's typical concerns regarding chillers are:

- Q1. How to optimize cooling system performance?
- Q2. How is the chiller operating (e.g., non-operational, low, high, or normal)?
- Q3. What are the hourly chiller usage patterns?

To answer the above questions, we start by employing the well-known scatter plots to visualize two attributes, chiller in-temperature and out-temperature. There are a number of interesting visualization techniques introduced to solve the overlap problem in scatter plots, such as HexBin scatter plots [HexBinR] determining the brightness value of each HexBin cell depending on the number of data points in the cell, and Jitter scatter plots [JitterR] which are used in R and SAS to display jittered values against the covariate.

To visualize chiller in-temperature and out-temperatures with chiller utilization, we map a third attribute to color which

represents chiller utilization as shown in Figure 8A. In the colored scatter plot, a data point is defined with three-attributes (x-axis, y-axis, color) where the x-axis is the chiller in-temperature, the y-axis is the chiller out-temperature, and color is the chiller utilization. Administrators can easily see the high correlation between in-temperatures and out-temperatures. All data points lie close to the regression line, such as data point p (58°F, 47°F, 40%), which indicates that when the chiller in-temperature has a medium-high value, then the chiller's corresponding out-temperature will also be in medium-high with a medium chiller utilization (green).

However, there are several data points such as p and p1 to p8 that are located at the same x- and y- coordinates in a scatter plot, thus causing overlap. Figure 8A shows that the data point p and p1-p8 have the same x-axis and y-axis location. The data point p is overlaid by the subsequent data point p1 to p8. Overlap causes two problems in visualizing data distributions and patterns: (1) the number of overlaid data points is unknown and (2) the chiller utilization value of overlaid data points is not visible.

In order to answer the first question (Q1) which is important for optimizing cooling system performance, we need to understand the correlation of the temperatures with the chiller utilization. The overlaid data points mask the true distribution. To solve this problem, we introduce a non-overlapped scatter plot technique as shown in Figure 8B. Similar to the colored scatter plot (Figure 8A), the x-axis represents the in-temperature and the y-axis represents the out-temperature.

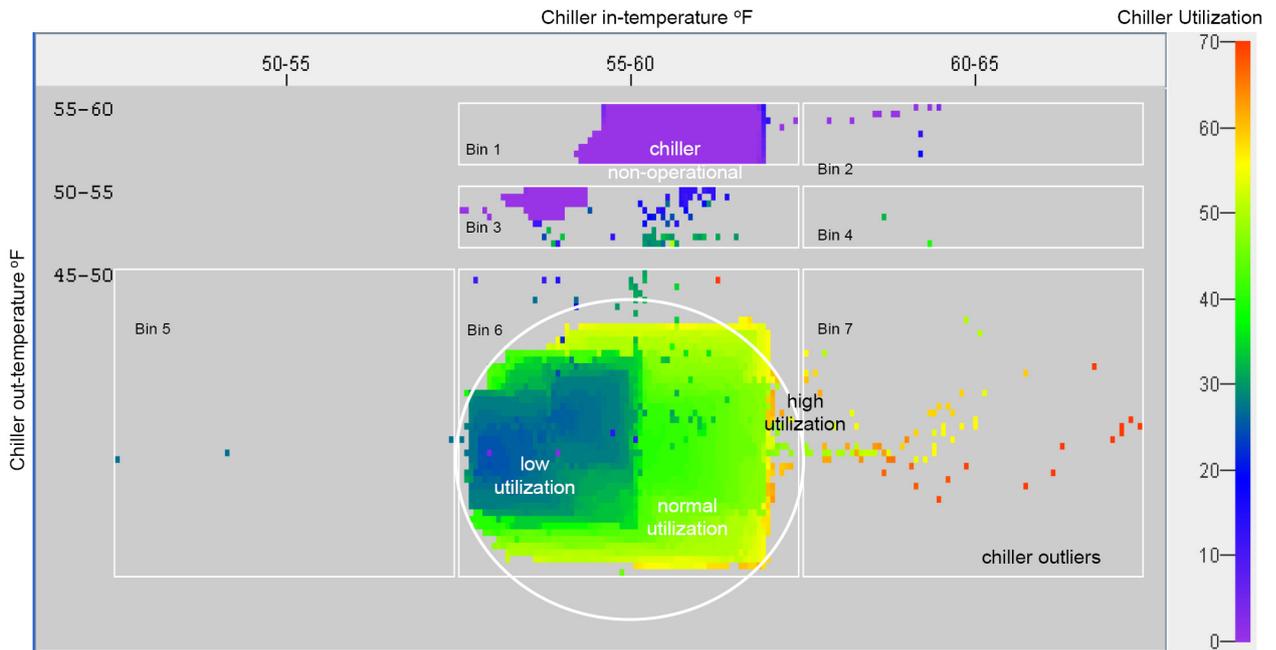


Figure 9: Non-overlapped chiller utilization scatter plot

Each data point $p(x, y, \text{color})$ represents a measurement with three attributes
 x : in-temperature, y : out-temperature, and color: chiller utilization.

Utilization is a measure of power consumed for a given cooling load.

Bin 6 has three utilization clusters: low (blue), normal (green), and high (yellow/orange).

The color of a data point represents the value of a third attribute, e.g., chiller utilization (%U), from low (blue), medium (green to yellow), to high (orange to red).

A data point $p(x, y, \text{color})$ is now represented by a cell, which is placed inside the corresponding temperature bin. A bin is defined as the number of data points having the (x, y) coordinates within a certain x and y value ranges. The binning of the x - and y -axes and the size of a bin are determined automatically according to the number of the chiller's in- and out- temperatures. For example, a data point p with $(x, y, \text{color}) = (58^\circ\text{F}, 47^\circ\text{F}, 40\%)$ is positioned within the bin $(55\text{--}60^\circ\text{F}, 45\text{--}50^\circ\text{F}, \text{utilization } 40\%)$.

The placement algorithm for our non-overlapping chiller utilization scatter plots uses the size of the bin (width X height) and the empty spaces around the already occupied data points to compute the best location for the data point that would be overlaid in a regular scatter plot. Data points P_1 to P_{10} with the same (x, y) coordinates will be placed in a neighborhood around the already plotted data point P according to the similarity of their colors. These points are sorted by the chiller utilization and are placed close together in the order shown in Figure 8B to form clusters.

The following steps describe the data point placement algorithm (Figure 8B):

1. Place the first data item P in the location (x, y) .
2. Place the 8 data items (P_1 – P_8) around P in a rectangular shape in the direction of top, bottom, left, right, and diagonals to form the rectangular shape. The placement sequence can be altered as long as we place the 8 data points as close to p as possible.
3. Place the remaining data points (e.g., P_9 , P_{10} , etc.) around the rectangle based on following location selection criteria: The neighborhood of the data points should have similar values (colors).
4. Enlarge the bin size as the number of data points with the same coordinates grows.

The resulting visualization reveals interesting distributions and patterns, (i.e., the blue, yellow, and green clusters formed by the low, high, and normal utilization respectively in Figure 9) which otherwise would have been lost.

Figure 9 illustrates one of the time-critical applications of weekday cooling system performance. Each data point (x, y, color) contains the in- and out- temperatures and chiller utilization at 1-minute intervals, and other related information (i.e., chiller id, rack id, etc.) The bins on the x -axis have

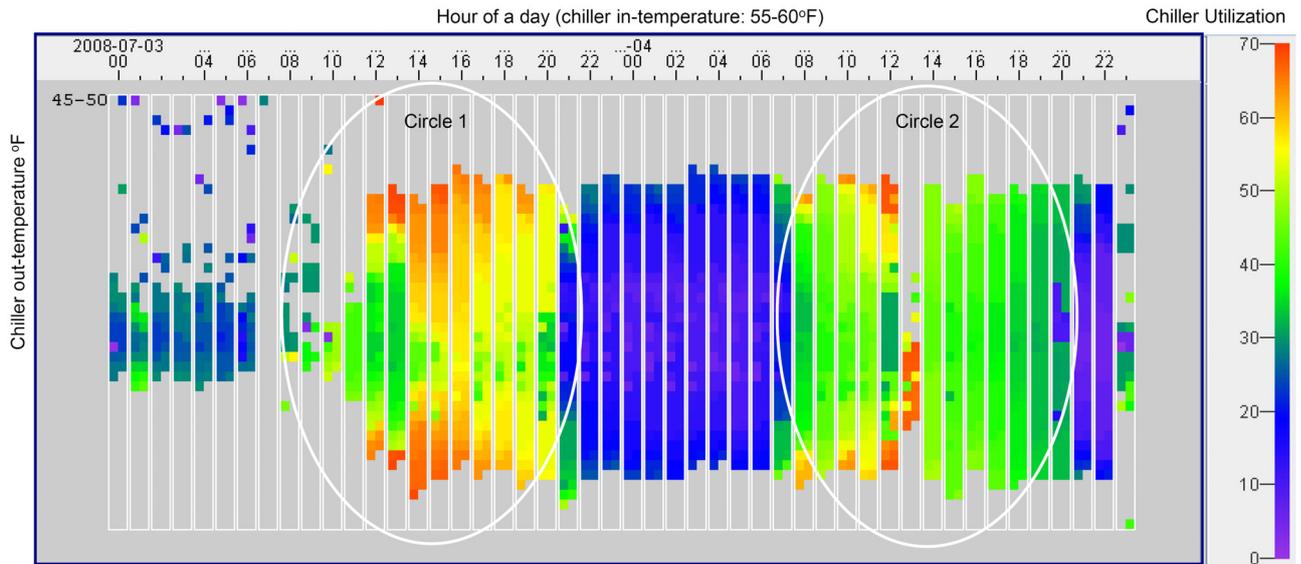


Figure 10: Chiller hourly utilization (drilldown from Figure 9, Bin 6)
(x-axis: hour of a day, y-axis: chiller out-temperature, color: chiller utilization)
Chiller has high capacity of utilization during hours 8 to 20 on 7/3 and 7/4 in the white circles 1 and 2.
Most colors are green, yellow, and red to indicate that chiller has high utilization during the day and early evening.

equal width. The bins on the y-axis have different heights according to the total number of data points within the bins.

In this non-overlapped chiller utilization scatter plot, administrators can mouse over an interesting data point and drilldown to the record level information, for example, to find which chiller is not operational or has a high utilization. Further, the administrator can easily observe the following facts to answer the second question (Q2) on the chiller operations:

- There is a high correlation between the temperatures and chiller utilization as shown in Bin 6 and Bin 7. The chiller utilization increases from blue, to green, yellow, and red when the in-temperature increases from 50°F to 65°F with the corresponding out-temperature increasing from 45°F to 50°F.
- The data points in Bin 6 show high variation in utilization of the chiller. The large blue, green, yellow, and orange areas show that there are clusters with low, medium, and high utilization. The high utilization clusters indicate where large volumes of water are needed to provide cooling.
- A few scattered outliers (red) in Bin 7 use the high 70% utilization to bring the in-temperatures of 60°C to 65°C to out-temperatures of 45°C to 50°C. Bin 7 requires more attention from administrators to reduce the heat generated by overloaded servers.
- The purple clusters in Bin 1 and Bin 3 indicate that the chiller for those racks and servers is not operational.

- Bin 4 and Bin 5 have the least number of data points (green and blue). Bin 6 (blue, green, yellow, and orange area) has the most number of data points and has the most interesting chilling process for data center managers to optimize its cooling performance.

Hourly chiller usage pattern can be obtained by drilling down into Bin 6 and changing the x-axis to represent time (24 hours). The y-axis is used for chiller out-temperature and color shows chiller utilization. This provides the answer to the third question (Q3). The resulting visualization shown in Figure 10 confirms that chiller utilization is higher during the day and early evening as shown in Circle 1 and Circle 2 (more red, orange, yellow, and green) than early morning and late evening (mostly blue). This is due to high server heat caused by running high computational-workloads on servers. Interestingly, circle 1 and circle 2 indicate different operational patterns. Administrators can drill down for further investigation.

6. Evaluation

Our new visual analytics techniques have been successfully employed on a daily basis in seven data centers of different sizes, ranging from 3,000 to 14,000 sq. ft. with hundreds of racks. Several billion records from data centers have been analyzed. Our data center managers and administrators have provided positive feedback: *“Using the visual analytics overview with automated anomaly markers and queries, we are now able to analyze the data easily within seconds instead of hours it would take without these visual techniques and we are able to rectify thermal state problems in real-time”*. The following are some of their reported results:

- A 20% energy savings by using the new out-of-sequence sensor temperature visual detection techniques.

Without the out-of-sequence temperature detection, administrators had to copy the temperature values manually from the interfaces, paste them onto a spreadsheet, and then highlight the abnormal sensors. This was a tedious and error-prone process.

- A 10-15% power savings by using non-overlapping scatter plots to optimize chiller efficiency, thus enabling administrators to observe chiller utilization patterns and their correlation of in- and out-temperatures. With this information, data center managers were able to better manage their cooling strategies: (1) chiller real time operations, (2) server workload balance, and (3) chiller maintenance.

Previously, administrators had to calculate the differences between the in-temperatures and the out-temperatures to find the correlations with the chiller utilization. Regular scatter plots only provided two coordinates (x, y), resulting in limited understanding of the relationships between utilization and actual temperatures.

7. Conclusions

The application of the new visual analytics techniques for thermal state detection has met the data center requirements of running their operations, enabling quick insight into the root-cause factors and allowing the administrators to take corrective measures.

We have started to embed the visual analytics capabilities into a mobility-enabled visualization platform, Data Center Mobile Studio, which is hosted on HP iPAQ 210. This empowers onsite administrators to visualize current thermal state information. The key technical work is centered on the ability to perform progressive visual analytics on data either from a local database cache or a web service.

Our future work will focus on combining historical data analysis with trends to predict anomalies before they happen so that preemptive measures can be taken, possibly before problems occur.

Acknowledgements

The authors wish to thank Meichun Hsu for her encouragement, and thank Manish Marwah and Cullen E. Bash for providing comments and suggestions.

References

[And95] G. Andrienko and N. Andrienko, Visual Exploration of the Spatial Distribution of Temporal Behaviors, Proceedings of the Ninth International Conference on Information Visualization (IV'05).

[And08] G. Andrienko and N. Andrienko, Spatio-temporal Aggregation for Visual Analysis of Movements. Proceedings of the IEEE Symposium on VAST 2008, OH.

[BPS06] C. Bash, C. Patel, R. Sharma, "Dynamic Thermal Management of Air Cooled Data Centers", IEEE Itherm06.

[CMBF07] C. Compieta, S. D. Martino, M. Bertolotto, F. Ferrucci, T. Kechadi, Exploratory Spatio-Temporal Data Mining and Visualization, Journal of Visual Languages and Computing. 2007.

[CXGH08] S. Chan, L. Xiao, J. Gerth, P. Hanrahan. Maintaining Interactivity While Exploring Massive Time Series. Proceedings of the IEEE Symposium on VAST 2008, OH.

[HD06] M. Hao, U. Dayal, D., Methods for visualizing graphical data sets having a non-uniform graphical density for display. US Patent 7,046,247, May, 2006.

[HDKMS07] M. Hao, U. Dayal, D. A. Keim, D. Morent, J. Schneidewind: Intelligent Visual Analytics Queries. IEEE Symposium on Visual Analytics Science and Technology, CA, 2007.

[HDKS07] M. Hao, U. Dayal, D. A. Keim, T. Schreck. Multi-Resolution Techniques for Visual Exploration of Large Time-Series Data. Proceedings: IEEE VGTC Symposium on Visualization, EuroVis 2007.

[HDKSM09] M. Hao, U. Dayal, D. A. Keim, Sharma R. K., Mehta A. Visual Analytics of Anomaly Detection in Large Data Streams. Proceedings: Visual Data Analysis Conference, CA. 2009.

[HexBinR] HexBin Scatter Plot released by R System. http://rss.acs.unt.edu/Rdoc/library/hexbin/doc/hexagon_binning.pdf.

[JitterR] Jittered scatters are used by the R and SAS systems.

[MCFR04] J. Moore, J. Chase, K. Farkas, P. Ranganathan. A Sense of Place: Toward a Location-aware Information Plane for Data Centers; HPL-2004-27.

[MUN04] T. Munzner. BinX: Dynamic Exploration of Time Series Datasets across Aggregation Levels, Proc. IEEE Symposium on Information Visualization, p. 215, 2004.

[SH08] R. Sharma et al., "Energy Flow in the Information Technology Stack", ACM Compute 2008.

[SBP02] R. Sharma, C. Bash, C. Patel, "Dimensionless parameters for evaluation of thermal design and performance of large-scale data centers", 8th ASME/AIAA Joint Thermophysics and Heat Transfer Conf., St. Louis, MO, June 2002.

[SS05] J. Seo, B. Shneiderman, A rank-by-feature framework for interactive exploration of multidimensional data Information Visualization 2005, Palgrave Macmillan Ltd.

[WS99] J. J. Van Wijk, E. R. Van Selow. Cluster and calendar based visualization of time series data. Proceedings: IEEE Symposium on Information Visualization, NJ, 1999.