Visual Evaluation of Cloud Infrastructure Performance Predictions

Holger Stitz*
Johannes Kepler
University Linz

Samuel Gratzl†
Johannes Kepler
University Linz

Harald Rogner
Johannes Kepler
University Linz

Marc Streit‡
Johannes Kepler
University Linz

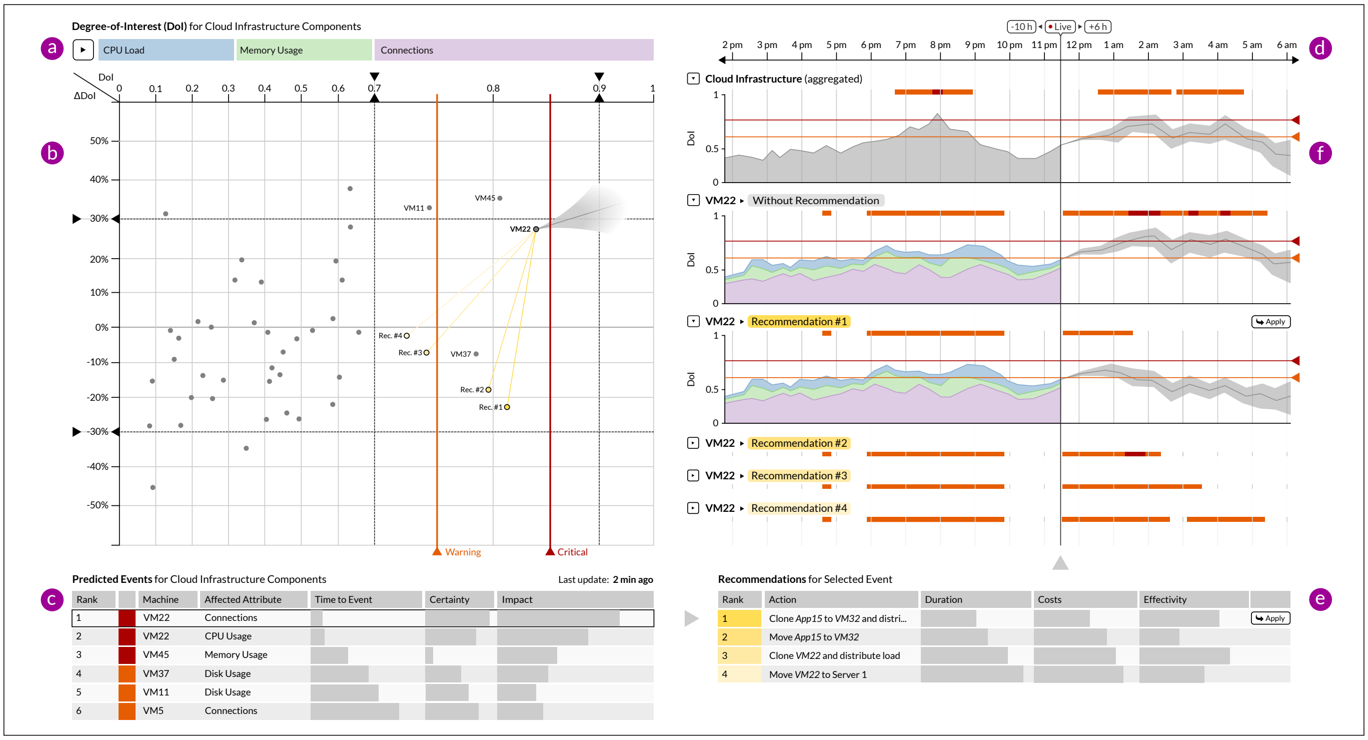

Figure 1: *Design sketch of the proposed visualization concept. (a) The administrator can configure the degree-of-interest (DoI) that determines how critical the status of a component currently is. (b) The ThermalPlot shows the current as well as predicted positions of components in the DoI space. (c) Ranked list of predicted performance bottleneck events. (d) Selected time span with past and future. (e) Ranked list of possible countermeasures for avoiding the predicted event. (f) Detail view visualizing the past and predicted items' performance using peak bars (collapsed state) and streamgraphs (expanded state).*

Abstract

The workload of a cloud infrastructure is continuously changing. Administrators need to find the balance between quality of service and cost efficiency. Proactively avoiding performance bottlenecks is a challenging endeavor due to the number and heterogeneity of the network components, the relationships among them, and the associated attributes. While pattern detection and simulation methods can be used to predict the performance of the network, efficient tools for exploring and evaluating those predictions are missing. In this poster we present first results of a design study for visually evaluating cloud infrastructure performance predictions. We combine established visualization techniques for exploring large item collections with custom techniques for investigating predicted time-series data, allowing the administrator to effectively monitor, evaluate, and optimize cloud infrastructure.

Keywords: Temporal data, prediction, simulation, cloud.

*e-mail: holger.stitz@jku.at
†e-mail: samuel.gratzl@jku.at
‡e-mail: marc.streit@jku.at

1 Introduction

Cloud infrastructure is the backbone of our digital world and consists of a large number of heterogeneous components, including servers, virtual machines, and services. Administrating such infrastructure is challenging due to two contradicting goals: (1) the infrastructure must be highly available and responsive and (2) it should operate as efficient as possible in terms of costs and energy consumption. Finding a trade-off between these conflicting goals requires effective tools for monitoring and optimizing such infrastructures. Potential performance bottlenecks are usually addressed only after they occur. However, in order to increase the reliability and efficiency of cloud infrastructures, the goal is to prevent future bottlenecks before they happen.

In this work we present a visualization concept that enables administrators to browse recorded historical data, monitor live performance data, and explore predicted performance data. The ultimate goal of our visualization concept is to let administrators visually investigate predicted performance bottlenecks (events) and compare alternative recommendations for preventing these potential problems.

2 Related Work

The increasing availability of real-time multi-attribute datasets has triggered recent research in prediction-based modeling and visualization. Cloud monitoring tools have been surveyed by Fatema *et al.* [1]. In alignment with their proposed taxonomy, our work focuses on the two cloud operational areas: *capacity planning* and *fault management*. Commercial solutions, such as *IBM SmartCloud Monitoring*[1], *HPE Cloud Optimizer*[2], or *Idera Uptime Infrastructure Monitor*[3] separate the monitoring views from the views that are designed to optimize particular machines. In contrast to existing tools, our concept combines capabilities to monitor live data with the means to evaluate predicted performance data as well as possible ways to address potential bottlenecks and critical states into a single view.

3 Visualization Concept

Our visualization concept is targeted at administrators who should be able to monitor cloud infrastructures and optimize them by visually evaluating a list of predicted events and recommendations. Figure 1 shows a design sketch that combines our earlier work on *CloudGazer* [4], *ThermalPlot* [3], and *LineUp* [2] with a new prediction view that is tailored to the specific domain problem. The concept consists of a series of linked views that are described in the following.

***ThermalPlot* View** The *ThermalPlot* visualization provides an overview of the cloud infrastructure by positioning the components in a space that maps the criticality of components to the x-axis and the positive and negative change of the criticality to the y-axis (see Fig. 1(b) and [3]). Consequently, the higher the DoI value, the more critical a component is, and the further on the right it will appear in the *ThermalPlot* space. How critical the status of a component is, is calculated by a configurable DoI function (see Fig. 1(a)) that is computed as a weighted sum of multiple performance attributes, such as CPU load, RAM usage, and the number of currently opened connections. Which performance attributes are taken into account by the DoI computation depends on the use case. Additionally, we show two vertical threshold lines that discretize the criticality of components into two states: 'warning' and 'critical'. The administrator can freely configure the thresholds of the states by changing the vertical position of the lines via drag and drop.

Predicted Events Ranking In the lower part of the visualization we show an interactive ranking containing a list of predicted critical events (see Fig. 1(c) and [2]). Events are the result of a prediction model that uses the historical and live data collection to discover potential bottlenecks. Each event consists of a cloud infrastructure component and a specific attribute that is expected to be critical, together with the estimated time span until the event occurs (the shorter the closer), the certainty of the prediction (the longer the more likely), and a value that quantifies the impact the event has on the perspective or infrastructure (the longer the bigger). An example for an event would be a rapid increase in the number of connections that a component must handle, which also results in a higher CPU load. Selected events are visualized in the *ThermalPlot* space with the predicted position and a funnel that encodes the certainty (see Fig. 1(b)).

Timeline The timeline (see Fig. 1(d)) shows the selected time range (e.g., 10 hours into the past and 6 hours into the future) for the DoI computation (past to live) and prediction (future). Adjusting the selected time range triggers a re-computation of all DoI values and events.

Recommendations Ranking Selecting an event from the event ranking (see Fig. 1(c)) brings up a separate ranking with recommendations that are supposed to prevent the event from happening (see Fig. 1(e)). Each recommendation consists of a description for the action and a multi bar chart representing the predicted duration that is needed to apply this recommendation (the shorter the faster), the cost estimation (e.g., bandwidth; the longer the more expensive), and an effectiveness estimation to reduce the event's impact (the longer the more effective). Recommendations are also shown (with their respective color coding) in the *ThermalPlot* space (see Fig. 1(b)) and in the detail view.

Detail View The detail view (see Fig. 1(f)) shows a selected component from the *ThermalPlot* or the event ranking view, together with all recommendations to potentially solve the issue. Each item can be represented as *peak bar* (collapsed state) or as a detailed streamgraph (expanded state). The peak bar visualizes DoI values that are above a certain threshold in the corresponding threshold color. The predicted performance changes that are expected for the different recommendations are also visualized using peak bars. When expanding the representation, we reveal the more detailed DoI streamgraph that visualizes the full time-series data within the selected time span. Available data in the past is visualized as a stacked DoI streamgraph that represents how much each attribute contributes to the aggregated DoI value over time. For predicted future performance values we show the predicted maximum, expected, and minimum DoI value. This encoding helps the administrator to judge how uncertain predictions are and what the predicted worst, average, and best case is.

4 Discussion

The stacked DoI streamgraph in the detail view visualizes the contribution of multiple attributes over time. In our concept, we only show the aggregated uncertainty for the overall combined DoI value. However, streamgraph representations cannot encode the individual uncertainty for each attribute. As a next step, we plan to evaluate different visual encodings that address this limitation.

Furthermore, our current concept does not visualize relationships between components in the infrastructure. However, this knowledge can be critical for the administrator to better judge side-effect that might be introduced by recommended infrastructure changes. In future work we will explore ways of how to include this topological information into our concept.

5 Conclusion

In this poster we presented a concept to visualize predicted data for cloud infrastructures. In our visualization concept administrators can investigate predicted events based on historical data and also evaluate different proposed changes to prevent these events from happening. As a next step, we will gather feedback from domain experts to further improve our concept.

Acknowledgements

This work was supported by the Austrian Research Promotion Agency (FFG) via the PIPES-VS-DAMES project (840232).

References

[1] K. Fatema, V. C. Emeakaroha, P. D. Healy, J. P. Morrison, and T. Lynn. A survey of Cloud monitoring tools: Taxonomy, capabilities and objectives. *Journal of Parallel and Distributed Computing*, 74(10):2918–2933, 2014.

[2] S. Gratzl, A. Lex, N. Gehlenborg, H. Pfister, and M. Streit. LineUp: Visual Analysis of Multi-Attribute Rankings. *IEEE Transactions on Visualization and Computer Graphics (InfoVis '13)*, 19(12):2277–2286, 2013.

[3] H. Stitz, S. Gratzl, W. Aigner, and M. Streit. ThermalPlot: Visualizing Multi-Attribute Time-Series Data Using a Thermal Metaphor. *IEEE Transactions on Visualization and Computer Graphics*, 2016.

[4] H. Stitz, S. Gratzl, M. Krieger, and M. Streit. CloudGazer: A Divide-and-Conquer Approach for Monitoring and Optimizing Cloud-Based Networks. In *Proceedings of the IEEE Pacific Visualization Symposium (PacificVis '15)*, pages 175–182. IEEE, 2015.

[1] `https://www.ibm.com/`
[2] `https://www.hpe.com/`
[3] `https://www.idera.com/`