Monitoring the Health of Computer Networks with Visualization

VAST 2012 Mini Challenge 1 Award: "Efficient Use of Visualization"

A. Kachkaev* I. Dillingham† R. Beecham‡ S. Goodwin§ N. Ahmed¶ A. Slingsby‖

giCentre, School of Informatics
City University London

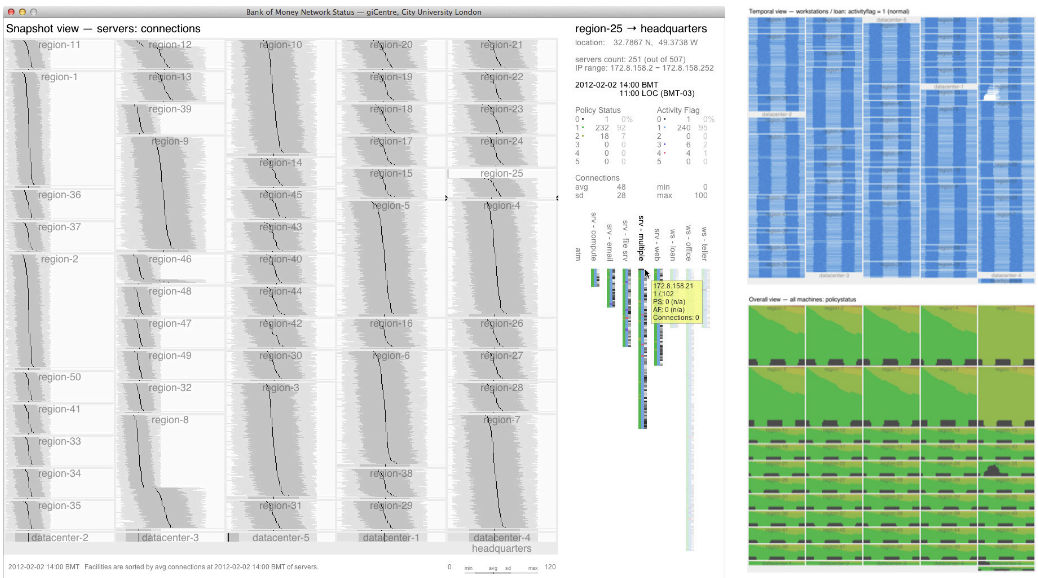

Figure 1: Left: snapshot view and details-on-demand panel showing the health of all 4,095 facilities. Regions, which contain facilities, are arranged geographically (with the exception of datacentres). Facilities are sorted by mean number of connections. Top right: excerpt showing temporal view, one category. Bottom right: excerpt showing temporal view, all categories. Regions are arranged alphabetically.

1 Introduction

The complex computer networks of large organisations contain many machines of many types, used in many geographic locations. Although system administrators should monitor the health of each machine, they need to do so within the context of the whole computer network. Our visualization presents the health of a fictitious financial institution's computer network at a snapshot in time and over a time range, and preserves the important aspects of each facility's administrative and geographic context. Using the "Bank of Money" VAST Challenge dataset, our visualization allowed us to correctly identify several areas of concern, as well as hypothesise about their causes.

*e-mail: alexander.kachkaev.1@city.ac.uk
†e-mail: iain.dillingham.1@city.ac.uk
‡e-mail: roger.beecham.1@city.ac.uk
§e-mail: sarah.goodwin.1@city.ac.uk
¶e-mail: nabihakk@gmail.com
‖e-mail: a.slingsby@city.ac.uk

2 Space efficient, context-preserving layout

An important requirement that emerged from our initial use of Google Earth[1] and Mondrian[2] was for a space efficient layout that preserved as much information as possible about a facility's geographic location, as well as its position in the institution's administrative hierarchy. HiDE[3] offers a spatially ordered treemap layout [4] that meets this requirement; although it became clear that the software would not scale to handle large amounts of data, HiDE helped us to explore the design space quickly [2].

The application features three overviews (Figure 1); the user can apply the spatially ordered treemap layout in each case and can pan and zoom to an area of concern.[4] The first view shows one category of one attribute at a snapshot in time; the second, one category of an attribute and the third, all categories of an attribute, over a time range. In the second view, the time range can be specified in either universal or local time (Figure 2); the former allows the user to explore cyclical patterns that are affecting machines in different timezones.

[1] http://www.google.com/earth/
[2] http://stats.math.uni-augsburg.de/mondrian/
[3] http://gicentre.org/hide/
[4] Several additional layouts, as well as sorting and filtering options, are available in the application.

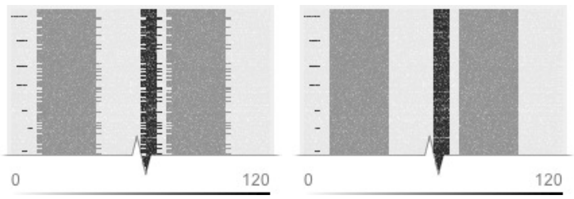

Figure 2: A range specified in universal and local time.

3 Simultaneous overview plus details-on-demand

The overviews are complemented by a details-on-demand panel, which shows the situation within a facility (Figure 3). This panel is updated by hovering over, or using the arrow keys to traverse, an overview. The proportion of machines in each attribute category, as well as summary statistics that describe the number of connections to the machines in the facility, are provided above a vertically stacked bar chart of the number of machines of each type (e.g. web server, ATM). This chart, which can be panned and zoomed, uses colour to represent all categories of all attributes for the selected facility.

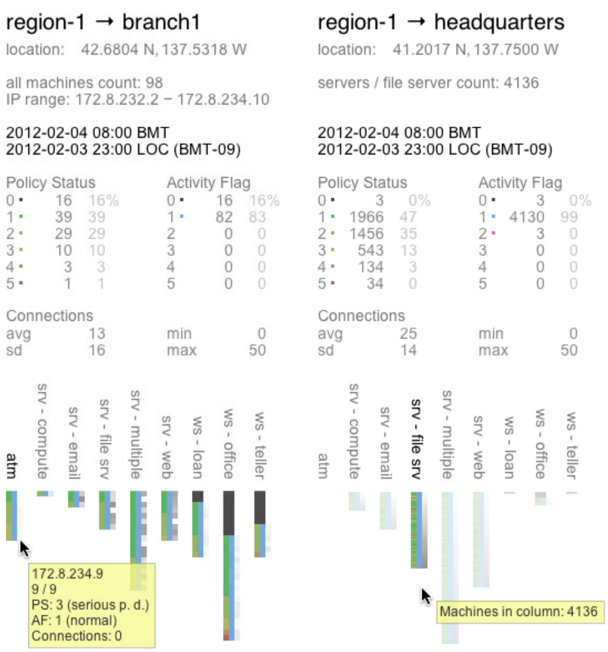

Figure 3: The details-on-demand panel for two facilities. Sorting, filtering and a tool tip allow the user to monitor the health of each machine within the context of the whole computer network.

4 User interface

Keyboard-based interaction is used throughout the application to improve efficiency; to enhance cognition, interaction is with the data [1], rather than with UI elements. This ensures that most of the screen is devoted to the status of the computer network, maximising the data-ink ratio [3].

5 Scalability

Two scalability issues are relevant: the first relates to the amount and type of data that the application can handle; the second to the scalability of the representations themselves.

The application is built on a 15GB PostgreSQL database and has a 2GB memory footprint; it can run on commodity hardware and could scale to handle larger amounts of data. Additional attributes could also be represented: although this would require updating the source code, the application is written in Java, using open source software libraries (Processing[5] and the giCentre Utilities[6]).

The overviews are based on a pixel representing a facility, which means that the entire computer network (containing nearly a million machines) can be represented in 1,280 by 1,024 pixels. Although additional facilities may necessitate additional displays, given the global coverage of the institution we feel that such a relatively small expense is reasonable.

6 Colour

An important requirement was the consistent use of colour across the application, where each hue should have a fixed brightness level to allow colour scaling in the single category temporal view (Figure 4). Two colour schemes were developed to meet this requirement: one for nominal and one for ordinal attributes (Figure 5).

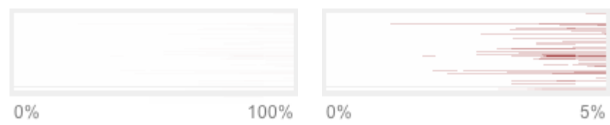

Figure 4: Colour scaling in the single category temporal view highlights hidden trends.

Figure 5: Colour schemes for nominal and ordinal attributes.

7 Conclusion

Our visualization allowed us to correctly identify several areas of concern in a fictitious financial institution's computer network. For the interested reader, a video[7] and submission report[8] (including links to download the application and inspect its source code) are available on the web.

Acknowledgements

The authors wish to thank Graham Dove, Jason Dykes and Ali Ramathan for their design and analysis suggestions.

References

[1] J. W. Crampton. Interactivity types in geographic visualization. *Cartography and Geographic Information Science*, 29(2):85–98, 2002.

[2] A. Slingsby, J. Dykes, and J. Wood. Configuring hierarchical layouts to address research questions. *IEEE Transactions on Visualization and Computer Graphics*, 15(6):977–984, 2009.

[3] E. R. Tufte. *The Visual Display of Quantitative Information*. Graphics Press, Cheshire, CT, USA, 2nd edition, 2001.

[4] J. Wood and J. Dykes. Spatially ordered treemaps. *IEEE Transactions on Visualization and Computer Graphics*, 14(6):1348–1355, 2008.

[5] http://processing.org/

[6] http://www.gicentre.org/utils/

[7] http://vimeo.com/47441537

[8] http://gicentre.org/vast2012/