Future Generation Computer Systems ▮ (▮▮▮▮) ▮▮▮–▮▮▮

Contents lists available at ScienceDirect

Future Generation Computer Systems

journal homepage: www.elsevier.com/locate/fgcs

Loom: Complex large-scale visual insight for large hybrid IT infrastructure management

James Brook [a], Felix Cuadrado [b], Eric Deliot [a], Julio Guijarro [a], Rycharde Hawkes [a], Marco Lotz [a], Romaric Pascal [a], Suksant Sae-Lor [a], Luis M. Vaquero [c,*], Joan Varvenne [a], Lawrence Wilcock [a]

[a] Hewlett Packard Enterprise, UK
[b] School of Electronic Engineering and Computer Science of Queen Mary University of London, UK
[c] Department of Computer Science. University of Bristol, UK

HIGHLIGHTS

- Interactive insight extraction techniques need to be extended for manageability software.
- An extended multiplex graph powers the "Thread" visualisation model.
- The extended multiplex graph delivers interactive response times (few seconds).
- Loom excels at aggregating data and on relational queries to these aggregations.
- It enables visual interactive spotting of relationships between aggregations of managed entities.

ARTICLE INFO

Article history:
Received 5 October 2016
Received in revised form 26 May 2017
Accepted 8 August 2017
Available online xxxx

Keywords:
Management
Cloud
Scale
Visualisation
Complexity
Extreme scale visual analytics
Visual analytics

ABSTRACT

Interactive visual exploration techniques (IVET) such as those advocated by Shneiderman and extreme scale visual analytics have successfully increased our understanding of a variety of domains that produce huge amounts of complex data. In spite of their complexity, IT infrastructures have not benefited from the application of IVET techniques.

Loom is inspired in IVET techniques and builds on them to tame increasing complexity in IT infrastructure management systems guaranteeing interactive response times and integrating key elements for IT management: Relationships between managed entities coming from different IT management subsystems, alerts and actions (or reconfigurations) of the IT setup. The Loom system builds on two main pillars: (1) a multiplex graph spanning data from different ITIMs; and (2) a novel visualisation arrangement: the Loom "Thread" visualisation model.

We have tested this in a number of real-world applications, showing that Loom can handle million of entities without losing information, with minimum context switching, and offering better performance than other relational/graph-based systems. This ensures interactive response times (few seconds as 90th percentile). The value of the "Thread" visualisation model is shown in a qualitative analysis of users' experiences with Loom.

1. Introduction

Information Technology (IT) infrastructures produce a myriad of messages and alerts that a human operator receives in order to make an assessment and execute some action that takes the IT infrastructure management (ITIM) system to a normal operating mode.

Many IT issues in an infrastructure are never analysed and resolved. This results in apparently benign early symptoms of failure (e.g. low priority alerts) being ignored until they have caused major problems. For example, an unusually slow drive tends to be ignored until the database running on it becomes unusable and transactions fail.

Most companies deal with this problem by continuously refining a set of thresholds on well-known metrics (Service Level Agreements, SLAs). This strategy has proven to be extremely efficient and effective when all services rely on a single ITIM in relatively modern IT setups [1].

* Corresponding author.
E-mail address: luis.vaquero@hpe.com (L.M. Vaquero).

http://dx.doi.org/10.1016/j.future.2017.08.013
0167-739X/

Unfortunately, there are several factors that set most IT setups apart from this ideal scenario:

- The IT infrastructure is the reflection of the history of a corporation and it tends to include many ITIMs that have been added over decades
- The extreme abundance of open source ITIMs released by all sorts of companies in an attempt to reduce costs, make recruitment easier, commoditise competing solutions, and/or share some knowledge back with the community
- High heterogeneity of the software stack running on the IT infrastructure (tools for accounting, relationships with suppliers, provisioning, provenance, human resources, and many other) makes management more difficult
- The continued effort towards setting programmable infrastructure that was started by virtualisation/cloud technologies and is continued by "containerisation", software defined network and network function virtualisation
- Miniaturisation of the user interfaces taking us from a world of "mobile first" to a situation closer to a "mobile only" where the ability to touch to manage the infrastructure is increasingly appealing

All these factors foster:

- The appearance of hidden interdependences (switching a host off with HPE iLO will kill the Kubernetes containers running on it but these two systems are not necessarily connected)
- The presence of partially overlapped functions so that an IT operator can exert the same effect via a different ITIM. One could, for instance, shutdown a container by directly using Docker or by invoking an appropriate Kubernetes command.

Hiring more administrators and increased automation (like testing, building and deploying [2]) have been the responses to cope the increased complexity of modern IT setups. In spite of these strategies, considerable human effort is still required to oversee and authorise semi-automated decisions, and full automation is often times only possible after careful human analysis of management data [3] .

Unfortunately, we have reached a point where increasing the number of administrators is simply not enough. Current projections estimate there will be 80 billion connected devices by 2020.[1] The number of servers a single sysadmin can manage range from 100 to 20000 depending on the complexity of the IT setup. This maps to 80-16000 times the number of U.S. majors in computer science in 2012.[2]

Poorly managed infrastructures result in security problems (e.g. unpatched devices attacked on well-known vulnerabilities), operational issues (e.g. slow times to deploy or upgrade a running service) and, eventually, in poor user satisfaction and profit losses.

Visual analytics (VA) is the "science of analytical reasoning facilitated by interactive visual interfaces" [4]. Extreme-Scale Visual Analytics (ESVA) focuses on large scale complex and possible multi-sourced data [5]. ESVA establishes a close interplay between (1) a fast (often in-memory) analytics backend capable of merging data from several sources to create insightful summaries and (2) fine interactive visualisations that promote exploration of summarised data and quick extraction of details of interest. Thus, ESVA enables humans to focus their full perceptual and cognitive abilities on the analytical process and extract actionable insight from complex data [5,6]. Similar guidelines to deliver interactive exploratory tolls are also advocated by Shneiderman in his famous mantra [7]: "Overview first, zoom and filter, then details-on-demand". We refer to these as interactive visual exploration techniques (IVET).

While these techniques have been widely applied across disciplines [5,8], IT infrastructure management has remained indifferent to their advancements except for a few position statement papers that claim that a new breed of systems (known as "Data Visualisation Management Systems") is needed to increase collaboration between visualisation and data infrastructures [9,10].

IT management requires to build on interactive visual exploration techniques to cope with information overload. In addition, there are some peculiarities of IT management that have not been explored by current IVET solutions:

- Alerts also need to be summarised and visualised together with the elements of the IT infrastructure they are associated with
- Spotting sources of problems and filtering noise in IT management makes it essential to highlight relationships across managed entities in different ITIMs. For instance,"discovering a failure in network link A caused errors in application X via server 1"
- IT management is not just about extracting insight from management data: the ability to exert "Actions" to change the state of the many ITIMs in the infrastructure in a coordinated way becomes critical to reduce cognitive overload
- As many IT operators can be "on the go" and access the management interfaces from displays with different form factors, the ability to touch on representations of the managed entities becomes increasingly interesting

Our system, Loom,[3] is the first system that:

- (1) Enables interactive touch-enabled exploration. This is possible thanks to a new graph model that includes managed IT nodes and aggregations of interest, which result from queries by the user. Loom can manage data from several ITIMs, creating multiplex aggregated graphs.
- (2) A new interactivity model where elements in the frontend are mapped to entities in the backend but only an abstract representation and a summary of the managed IT infrastructure elements are sent to the frontend.
- (3) A set of internal optimisations that make Loom capable of coping with large heterogeneous scale IT installations in interactive response times (less than 30 s).

The rest of this paper is structured as follows: Section 2 presents how Loom builds on visual interactivity and exploration techniques for large full stack IT setups containing several ITIMs. After presenting the main functionalities and the fine orchestration required between frontend and backend, we present Loom design in Section 3. The advantages of Loom are exemplified with two use cases in Section 4. We evaluate Loom in terms of performance and user experience and present the results in Section 5. Loom is compared with state of the art systems in Section 6. The main conclusions are summarised in Section 7.

2. The loom visual insight extraction model for large-scale IT management

Corporate IT users have been characterised in enterprise usability studies [11]. Loom caters for two of them. Advanced users

[1] http://www.forbes.com/sites/gilpress/2014/08/22/internet-of-things-by-the-numbers-market-estimates-and-forecasts/.

[2] http://archive2.cra.org/uploads/documents/resources/taulbee/CS_Degree_and_Enrollment_Trends_2010-11.pdf.

[3] The name comes from the metaphor that summarising information can help users "weave" their own "tapestries" rendering a custom overview of the ITIMs in their IT setup.

Please cite this article in press as: J. Brook, et al., Loom: Complex large-scale visual insight for large hybrid IT infrastructure management, Future Generation Computer Systems (2017), http://dx.doi.org/10.1016/j.future.2017.08.013.