BlueSight: Automated Discovery Service for Cloud Migration of Enterprises

Dannver Wu[1]([✉]), Jinho Hwang[2], Maja Vukovic[2], and Nikos Anerousis[2]

[1] University of California, Berkeley, CA, USA
dannver@berkeley.edu
[2] IBM T.J. Watson Research Center, New York, USA
{jinho,maja,nikos}@us.ibm.com

Abstract. Migrating legacy enterprise infrastructures to the cloud is highly desirable due to greater versatility, lower management costs, as well as improved scalability. However, the large scale of these systems makes transforming the current architecture a long and difficult process that involves weeks or even months of manual collection and analysis of data. BlueSight serves to expedite and simplify this process by collecting the data through an agentless process and analyzing the collected data to determine which and how applications should migrate.

Keywords: Cloud computing · Migration analytics · Workload discovery

1 Introduction

The benefits of cloud platforms, which include reduced management costs and greater versatility, create a need to migrate existing applications running on older systems. However, this migration process typically involves a lengthy and costly process that requires human-to-human interactions and user input. In this process, engineers must decide which applications to migrate and if the application should be directly moved as a single entity or be broken down and refactored into independent micro-services.

On the large scale of enterprise systems, ranging from hundreds to thousands of servers, it is difficult to understand and analyze the legacy applications to construct a plan for migration. Currently, an experienced team of migration engineers is required to decide the best method of migration for each application. Different applications require different changes to move to a cloud platform (e.g. re-hosted, re-factored, etc.).

In this paper we describe an automated discovery service for the migration and transformation of enterprise systems to the cloud. Discovering the architecture is particulary challenging due to the complexity of the application dependencies. The service automatically collects the necessary data and generates a detailed report to be used by migration engineers to determine a plan for the transformation process. In this report, we provide detailed information about

the infrastructure, as well as classify the applications based on their suitability for transformation to the cloud, and propose a migration pattern for the application. Additionally, a clear visual interface provides users the ability to clearly understand their infrastructure by filtering, clustering, ranking, and viewing the dependency and geographical information.

2   System Overview

BlueSight consists of two main steps, collection and analysis, in order to create a comprehensive system report for migration teams. Applications have many significant dependencies on a multitude of underlying components, such as the OS, various system and security configurations, etc. The collection process is executed as an agentless method, either automatically through remote access or manually uploaded by the user, providing detailed data about the complex dependencies within the architecture, as well as various server metrics. Once this data is collected, the entire system is displayed through a clear visual graph with filtering, ranking, and clustering functionalities to allow the user to understand and prioritize specific applications for migration.
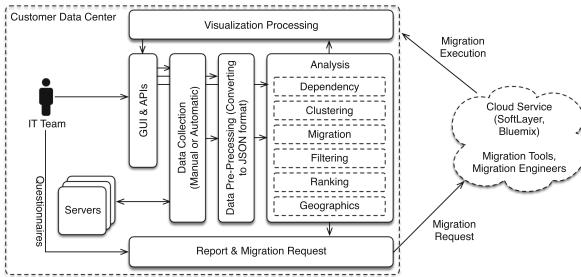
Fig. 1. BlueSight architecture.

Figure 1 illustrates the BlueSight architecture and workflows throughout the migration analytics process. The IT team in the customer data center can use BlueSight through a GUI or through APIs. When the IT team registers migration candidate servers and triggers "Data Collection", BlueSight accesses the servers automatically through ssh (linux) or samba (Windows) connections and collects information [6,7]. Then, the raw data is parsed and converted into the JSON format. The analysis engine takes over the data and runs through the different analytics: dependency, clustering, migration, filtering, ranking, and geographics [1,3–5]. The visualization processing engine generates graph data for the GUI interface. When the IT team decides to migrate any applications or servers, BlueSight will trigger the migration and generate a report of specific information about the applications or servers will be sent to cloud migration engineers who will make a migration plan [8]. Migration orchestration tools, such as the IBM cloud migration orchestrator, can also be integrated into the process [2].

3   Functions/Features

3.1   Collection

BlueSight takes an agentless discovery approach, using automated scripts to collect data from the servers [1]. Since enterprise customers are very sensitive to security, it is extremely hard to obtain machine credentials without compromising privacy. Thus, BlueSight is deployed into the customer's data center and isolated from the outside (even from IBM). The collection process can be executed automatically through remote access, or alternatively the user can run the scripts themselves and upload the resulting archived files to BlueSight. This process collects data about system properties, CPU/memory/disk/network usages, network statistics, as well as specific dependencies, services, processes, and applications of each server.

3.2   Analysis

Once the data is collected, the user is presented with a visualization of all of the servers, OS, middlewares, instances, applications, and databases in the system. There are several ways to get more detailed information. Users can view a summary of the entire infrastructure enumerating the unique applications, total discovered applications, servers, as well as overall averages of system metrics and a list of the unique applications. Additionally, clicking any node in the visual graph will display specific information about the server, including the applications/middlewares that are being run on that server and the hardware resources and usage statistics.

Users are also provided functionalities to manipulate the graph. In the top right corner are options to show or hide components based on type. Filters allow users to display servers that fit user-defined thresholds of CPU usage, disk usage, network usage, etc. A search-box and an application list on the side allow users to search for and display specific applications along with the connected components (e.g. OS, server).

Being able to group servers by a variety of metrics helps users to determine the significance of migrating each application. BlueSight conducts a migration analysis of the architecture, classifying servers into different migration patterns, suggesting whether each could be retained in the current state, retired, re-hosted, re-platformed, re-factored, or re-architectured. Each pattern is a good indication of where the servers migrate into. BlueSight also allows users to cluster the servers in the graph through a variety of statistical algorithms and view each group of servers visually.

4   Case Study

In our case study we will demonstrate the service using the data of an anonymous customer that consists of 121 servers running 631 instances of 37 unique
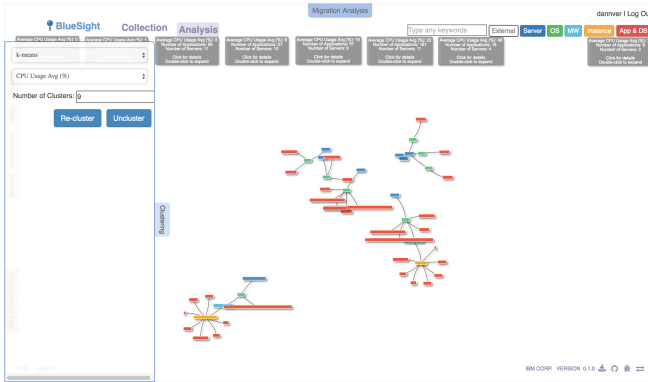
Fig. 2. Clustering in BlueSight.

applications. The collection script collects data and analyzes the application dependencies. In the GUI view, as seen in Fig. 2, the user is presented with the entire architecture as a graph with the various components as nodes, and the relationships and dependencies represented as edges. There are options to filter the graph by types of nodes. On the left, there are menus to further filter the graph through various system metrics, rank the applications using these same metrics, as well as generate clusters. Figure 2 is the result of clustering the graph by average CPU usage using k-means, to form 9 groups. The clustered nodes are displayed on the top as larger gray nodes, and each one can be expanded to show the contents. Clicking the cluster node shows additional details concerning the contents of the selected cluster.

One of the most exciting parts of this interface is in the menu on the top of the screen, which shows the results of the automatic migration analysis. This shows the user which migration pattern each application fits, which allows the customer to quickly and easily decide which applications to focus on first during the transformation. A more in-depth video of the service is available at http://bit.ly/BlueSightDemo.

5   Conclusion

In this paper, we described an automated discovery service for the migration of enterprise systems to cloud platforms. BlueSight solves the issue of being able to understand large enterprise infrastructures and assists both the user and migration teams in deciding what applications to transform, as well as how to do the transformation.

References

1. Hwang, J.: Computing resource transformation, consolidation and decomposition in hybrid clouds. In: 2015 11th International Conference on Network and Service Management (CNSM), pp. 144–152, November 2015

2. Hwang, J., Huang, Y.W., Vukovic, M., Anerousis, N.: Enterprise-scale cloud migration orchestrator. In: 2015 IFIP/IEEE International Symposium on Integrated Network Management (IM), pp. 1002–1007, May 2015
3. Hwang, J., Liu, G., Zeng, S., Wu, F.Y., Wood, T.: Topology discovery and service classification for distributed-aware clouds. In: 2014 IEEE International Conference on Cloud Engineering (IC2E), pp. 385–390, March 2014
4. Hwang, J., Huang, Y.W., Vukovic, M., Jermyn, J.: Cloud transformation analytics services: a case study of cloud fitness validation for server migration. In: Proceedings of the 2015 IEEE International Conference on Services Computing, SCC 2015, pp. 387–394. (2015). http://dx.doi.org/10.1109/SCC.2015.60
5. Hwang, J., Vukovic, M., Anerousis, N.: FitScale: scalability of legacy applications through migration to cloud. In: Sheng, Q.Z., Stroulia, E., Tata, S., Bhiri, S. (eds.) ICSOC 2016. LNCS, vol. 9936, pp. 123–139. Springer, Cham (2016). doi:10.1007/978-3-319-46295-0_8
6. Jermyn, J., Hwang, J., Bai, K., Vukovic, M., Anerousis, N., Stolfo, S.: Improving readiness for enterprise migration to the cloud. In: Proceedings of the Middleware Industry Track, Industry papers, NY, USA, pp. 5:1–5:7 (2014). http://doi.acm.org/10.1145/2676727.2676732
7. Nidd, M., Bai, K., Hwang, J., Vukovic, M., Tacci, M.: Automated business application discovery. In: 2015 IFIP/IEEE International Symposium on Integrated Network Management (IM), pp. 794–797, May 2015
8. Vukovic, M., Hwang, J.: Cloud migration using automated planning. In: NOMS 2016–2016 IEEE/IFIP Network Operations and Management Symposium, pp. 96–103, April 2016