TREO
Technology, Research, Education, Opinion

Multi-Level Root Cause Analysis in large scale Hybrid Cloud IT-Services

Sebastian Wind, Sebastian.Wind@fau.de; Andreas Both, Andreas.Both@datev.de

Enterprise IT-Services often consist of a huge number of software and hardware components, which have grown historically, driven by constantly changing technology. Hence, a mix of legacy and state-of-the-art components is present in any well-established company. Analyzing such systems in the cause of a problem is cumbersome, but possible. Well researched approaches like Root Cause Analysis in Public Clouds [1] do not tackle aspects like running multiple platforms. Other methodologies such as RCSF [2] are only using application level monitoring. Hence, root cause analysis of real-world IT environments are extremely difficult or incomplete. Today, this problem is addressed by highly experienced personal, who are enabled to maintain the systems, to analyze recognized problems, and finally to uncover the root cause of such problems. The possible retirement of personal, the investment of manual labor as well as the requirement for higher velocity in the whole recovery process demand an enhanced root cause analysis process.
The root cause of an incident is very hard to locate because of the more and more complex dependencies inside the services. Existing methods are generating a flood of events that cannot be handled easily by the human operators. It is very important to do the troubleshooting in time, because every minute of downtime can be costly. Most of the existing methods are relying on extensive monitoring and are observing the relationship by the timestamp of the alarms. Those solutions are designed for specific platforms or just work on application level.

In contrast to previous approach, we will focus on establishing a generalized, holistic approach while also establishing features, which lead to a high automation of the analytics steps. The challenge in determining the root cause is to find the connection between anomalies across the service components. That connection must be singled out through different isolation levels (application level, platform level, infrastructure level). Hence, we plan to automate the process of recognizing the correlation between the anomalies, in order to identify a pattern and display the most likely root causes across the isolation levels with high precision.

We propose a methodology that is based on standard IT-management processes (ITIL) to assist in effectively identifying root causes. By using standard processes, we can leverage the structures, that are already in place and help to isolate root causes more quickly as well as more efficiently than (semi-)manual approaches. On the implementation level two requirements provide the main challenges: (1.) the approach has to be as highly unobtrusive as large investments on legacy component are not justified as well as (2.) a strong generalization is required to be enabled to cover efficiently the following innovation cycles of services without knowing upcoming features.

References

[1] Weng, Jianping & Wang, Jessie & Yang, Jiahai & Yang, Yang. (2017). Root cause analysis of anomalies of multitier services in public clouds. 1-6. 10.1109/IWQoS.2017.7969155.
[2] Wang, Kui & Fung, Carol & Ding, Chao & Pei, Polo & Huang, Shaohan & Luan, Zhongzhi & Qian, Depei. (2015). A methodology for root-cause analysis in component based systems. 243-248. 10.1109/IWQoS.2015.7404741.