SLA : Smart Log Analytics

POSTER

Kemal A. Delic
Enterprise Services
Hewlett-Packard Co.
Grenoble, France

Jeff A. Riley
Enterprise Group
Hewlett-Packard Co.
Melbourne, Australia

***Abstract*—Enterprise Applications represent the major legacy workloads in the contemporary, corporate Data Centers. They will stay for several years in the future and will represent the principal headache of the always diminishing support operation staff. In this poster we present an innovative approach aiming to automate routine operations chores, support analyst's investigation and enable timely business insights into overall application landscape status and trends.**

Keywords—enterprise applications, workloads, logs, analytics, machine learning, closed-loop control

I. ENTERPRISE APPLICATIONS ARCHITECTURE

Legacy Enterprise Applications (EA) are typically constructed as a stack of several servers (e.g. web, application database), providing dependable, daily services to several thousands of users and called three-tier applications. They are usually monitored and managed by layers, and typically by separate organizations and operators, which creates important inefficiencies. We see this as the key problem, which could be better addressed by considering the entire EA stack holistically (Fig. 1).

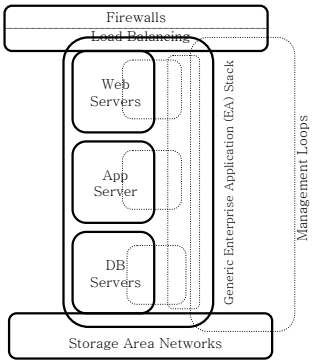

Fig. 1 Typical EA Stack (adapted from [1])

All EA components have logs, traces events and cases captured either from hardware and/or software by various types of monitoring agents. Three levels of closed loops are:

- an outer loop controlling entire EA,
- a middle loop controlling all servers together, and
- an inner-loop controlling only specific servers.

It should be obvious that the bi-directional log data and overall information flow will provide efficient monitoring and management of the EA domain [1]. We will outline next the overall architecture of the smart log system.

II. MONITORING AND MANAGEMENT ARCHITECTURE

Our fundamental [2] idea is embodied in a naturally layered architecture, so that the data and information flow from the infrastructure and applications into a log analytics engine which has an intricate, layered organization and drives and feeds several presentation and interaction layers (Fig. 2). This is a highly distributed, real-time architecture in which strategic business decisions are made by humans, while low-level chores are executed without human intervention.

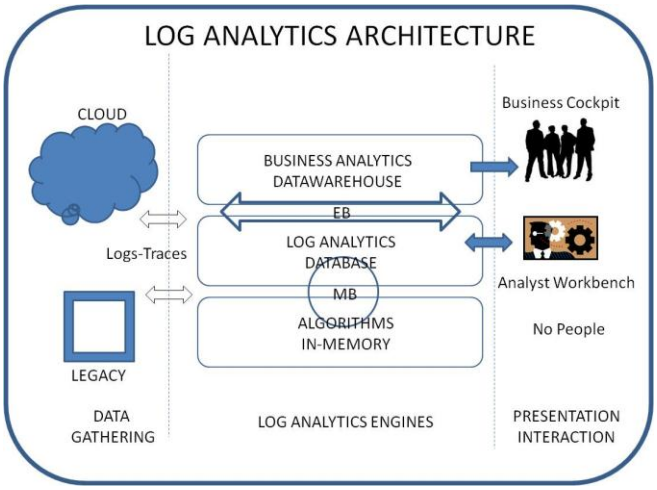

Fig. 2 Layered Log Analytics Architecture

It is important to notice that the log analytics engine is stratified into fundamental, low-level algorithms working on in-memory data sets feeding database layers above which, in turn, drive the complex set of tools for the analyst's workbenches. They are interconnected via a message broker (MB) enabling asynchronous, real-time capabilities of the system. The innovative here is making this architecture distinctively better from known systems.

Above those two layers, we have the data warehouse which delivers functionality of the business cockpit – enabling insights into the state of all corporate applications, and driving

prediction engines that enable management of an entire business application landscape. The Enterprise Bus (EB) interconnects all layers, providing timely communication. The Cockpit is different from the usual dashboards as it represents a closed management loop in which timely, strategic decisions can be made, while routine decisions are automated and executed without human operators in the loop. This specific deployment of Message Broker and Enterprise Bus is the key to Cockpit functionality. Furthermore, one should observe that enterprise servers and devices are creating continues streams of GB of data, which then amounts to TB of data to be processed into analytics and which then belongs by the scale into domain of Big Data Analytics (BDA).

In the following section, we dive slightly deeper into the database layer analytics, and discuss the technologies deployed for log and trace analytics management.

III. ZOOM-IN: DATABASE SERVER MONITORING

Currently, most problems with databases are resolved either via embedded scripts following if-then-else logic, or with the engagement of database analysts to manually and visually interrogate very large database and application logs, trying to understand the cause of the problem and develop an appropriate remedy and problem solving.

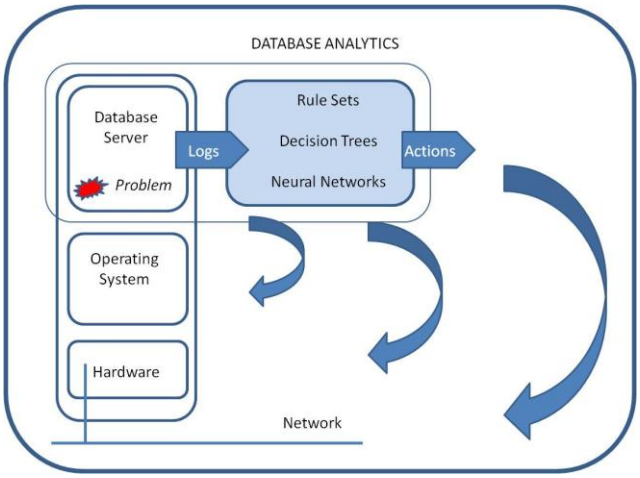

Fig. 3 Conceptual Database Log Analytics Design

We suggest another approach, in which (at least) three types of specialist knowledge will be captured and utilized. A **Fuzzy Rule Base System** will capture designer and expert knowledge, devised from hypothetical and real case studies, as well as knowledge automatically discovered from the knowledge bases of the underlying EA systems. We think ooperational knowledge is best represented via **Decision Trees** capturing knowledge gained during real systems operations. The decision trees will be constructed in-part automatically by mining knowledge from operational knowledge bases, and fine-tuned by injecting expert knowledge. Finally, we suggest that **Neural Networks** are the best technology to recognize problem signals and patterns in huge log and trace collections. The neural network will be trained automatically with knowledge mined from operational knowledge bases and logs. The well established Black Board AI paradigm is used as the architectural choice, enabling cooperation/collaboration of fundamentally very different technology systems. This fits well three types of problem classes here: heuristic, deterministic and probabilistic. Objective then is to cover entire problem space, as completely as possible. In this part of the system, we implement an innovative step which makes corroborative knowledge management system, by far, more effective and efficient. We are currently experimenting with this technological and architectural environment. We also realize that the behavior of web, application and database servers have very different working circumstances that need to be carefully explored and simulations developed [3].

IV. FUTURE TECHNOLOGY DEVELOPMENTS

The system we have presented is called Smart Log Analytics (SLA) – as we aim to provide an autonomic, effective and efficient system from which we expect to develop an appropriate set of technologies changing the cost-effectiveness equation for the specific target of Enterprise Applications. But we are also looking to the possibility of applying the same approach to cloud-resident applications. For those types of applications, the volume of logs will be much larger and different from legacy enterprise applications, but our hope is that experiences gained with SLA will enable some benefits for further research in the area of cloud-resident applications.

We also envisage the creation of a huge repository of logs for training and simulation purposes, and expect that some of the smart log analytics will necessarily require cloud computing resources [4] . The direction of this research will be more about exhaustive exploration of enterprise workloads and automatic judgment against the infrastructure and resources necessary to execute them. This will in turn enable semi-automation of application migration projects, which are largely done today thanks to human intelligence and practical experiences.

We are currently driven by business priorities and a focus on legacy application landscapes, while in the future we see this whole area developing as the deployment of a variety of Artificial Intelligence technologies on/above Big (log) Data collections. We also hope that this line of research [5] may help improve the performance of IT systems and offer improvement hints to EA designers and architects.

Acknowledgment

We are experimenting with several types of applications with sufficiently large log traces, trying to establish a test environment which will provide a basis for deployment in a production environment. We also acknowledge that we report only about our work at very high level and without providing any in-depth details, as we report from on-going industry experimentation. One should also observe that we are using only our own references as the trace of our own industrial research while indicating work of others in large collection of work of others [5].

References

[1] K. Delic and M. Walker, "Architecting Enterprise Grids: Possible Inflection Points", IADIS International Conference on Applied Computing, Salamanca, 18-20 February 2007, pp. 113-121.

[2] K.A.Delic, W.M. Green, U. Dayal, "Active Enterprise Analytics", HP OVUA 2006 Workhop, Sophia Antipolis/Nice, June 2006.

[3] K. Delic, J. Riley, C. Bartolini and A.Salihbegovic., "Knowledge-Based Self-Management of Apache Web Servere".Proceedings of the 21st International Symposium on Information, Communication and Automation Technologies (ICAT'07). Sarajevo, Bosnia and Hercegovena, October 2007

[4] K. Delic and J. Riley, "Enterprise Knowledge Clouds: Architecture and Technologies" In Borko Furht and Armando Escalante, editors, Handbook of Cloud Computing, chapter 10, pages 239-254. Springer, 2010. ISBN 978-1-4419-6523-3.

[5] Comprehensive Literature Collection and Survey : System Trace and Log Management – 1980-2015 – Private Collection, 2015