Reengineering Cloud Data Centers

Josef Spillner
Zurich University of Applied Sciences

Alan Sill
Texas Tech University

Data centers should be among the most automated, robust environments on the planet. Instead, they often remain stuck in the past as agglomerations of outdated technologies. It is time to take this problem seriously and directly address the needs of data centers for modern automation and control through analytics-driven methods.

Computation now extends widely and deeply into all areas of society, business, and educational fields. This trend is only increasing with time. In light of the deep dependence on advanced computation in all areas of life, it is necessary to re-examine the role of the data centers that support these functions, taking into account considerations of security, speed and latency of access, dynamic provisioning of resources, transportability, protection, and provenance of data, as well as other engineering factors.

The goal of this special issue is to present original articles on new methods that extend the state of the art and that highlight best practices to achieve highly capable, robust, scalable data centers that are well-suited to serving the infrastructure needs of large-scale clouds in an increasingly distributed world. From mega- to micro-data centers and edge nodes, diverse architectural choices emerge, each with its own new challenges. The data centers themselves, of course, exist only within the context of complete data processing ecosystems, and so their designs have to be responsive to the needs of their particular workflows, user communities, and software and hardware providers.

After a rigorous selection in which the vast majority of submissions were not possible to include for space reasons, but all were evaluated carefully, we are delighted to present in this special issue three compact, very well thought out and well-presented articles that address important themes concisely and with, we believe, deep insight.

These articles each take on an important aspect of the use of data centers. In the article "Accelerator Virtualization in Fog Computing: Moving from the Cloud to the Edge," authors Blesson Varghese, Carlos Reaño, and Federico Silla consider the complete relationship between cloud and edge computing, addressing ways in which technology is needed to accelerate functional operations at each stage of processing. This is a very comprehensive article that takes on a number of topics that relate to the effective use of clouds in fog computing, as well as the performance and communication factors that must be considered.

26

The article "Boosting Energy Efficiency and Quality of Service through Orchestration Tools," by Alessandro Carrega, Giancarlo Portomauro, Matteo Repetto, and Giorgio Robino, proposes a novel approach to the related problems of improving energy efficiency and performance through application of new software orchestration paradigms to provide and update context information automatically for elastic services. They present not just a concept for such quality-of-service (QoS) improvement, but also an experimental framework that can be evaluated and implemented by others interested in this problem.

Finally, the article "PRTuner: Proactive-Reactive Re-Replication Tuning in HDFS-based Cloud Data Center," by Thanda Shwe and Masayoshi Aritsugi, gives a specific, sophisticated, and highly replicable example of application of modern analytics methods to important problems in managing storage in large-scale data center environments. Their analysis places this work into context with other similar systems so that its conclusions can be applied to other settings.

ACKNOWLEDGMENTS

ABOUT THE AUTHORS

Josef Spillner is a senior lecturer (associate professor) and head of the Service Prototyping Lab at Zurich University of Applied Sciences in Switzerland. His research interests include cloud-native applications, service tooling, and cloud accounting and billing. With his team, he works on challenging topics such as microservices, function-as-a-service, and time series analysis. Before founding the lab, he founded the Open Source Service Platform Research Initiative to promote reusable software for scientific work. His work approach continues to promote international exchange, most recently at Itaipu Technology Park and Indian Institute of Technology, and modernization of research and publication approaches. He published a doctoral dissertation about meta-quality of services and a habilitation treatise about stealth computing in multi-cloud environments. Contact him at josef.spillner@zhaw.ch.

Alan Sill is the managing director of the High-Performance Computing Center at Texas Tech University, where he is also an adjunct professor of physics. He is also a visiting professor of distributed computing at the University of Derby. He co-directs the multi-university US National Science Foundation Industry/University Cooperative Research Center for Cloud and Autonomic Computing (CAC). Sill has a PhD in particle physics from American University and has an extensive track record of work in particle and nuclear physics and scientific computing, including cloud and grid computing. He serves as president of the Open Grid Forum, is an active member of many computing standards working groups and roadmap committees, and has served as principal organizer, program committee chair, or general chair for a number of large-scale conferences and workshops. Contact him at Alan.Sill@ttu.edu.