# Editorial for FGCS Special issue on "Time-critical Applications on Software-defined Infrastructures"

Zhao, Z.; Taylor, I.; Prodan, R.

**DOI**
10.1016/j.future.2020.07.056
**Publication date**
2020
**Document Version**
Author accepted manuscript
**Published in**
Future Generation Computer Systems


**Citation for published version (APA):**
Zhao, Z., Taylor, I., & Prodan, R. (2020). Editorial for FGCS Special issue on "Time-critical Applications on Software-defined Infrastructures". *Future Generation Computer Systems*, *112*, 1170-1171. https://doi.org/10.1016/j.future.2020.07.056

# Editorial for FGCS Special issue on "Time-critical applications on software-defined infrastructures"

**Dr. Zhiming Zhao**
**Multiscale Networked Systems**
**University of Amsterdam**
**Email: z.zhao@uva.nl**

**Dr. Ian Taylor**
**Computer Research Center**
**Notre Dame University,**
**Cardiff University**
**Email: ian.j.taylor@gmail.com**

**Prof. Radu Prodan**
**Institute of Information Technology**
**University of Klagenfurt**
**Email: radu.prodan@aau.at**

# Introduction

Performance requirements in many applications can often be modelled as constraints related to time, for example, the span of data processing for disaster early warning [1], latency in live event broadcasting [2], and jitter during audio/video conferences [3]. These time constraints are often treated either in an "as fast as possible" manner, such as sensitive latencies in high-performance computing or communication tasks, or in a "timeliness" way where tasks have to be finished within a given window in real-time systems, as classified in [4]. To meet the required time constraints, one has to carefully analyse time constraints, engineer and integrate system components, and optimise the scheduling for computing and communication tasks. The development of a time-critical application is thus time-consuming and costly.

During the past decades, the infrastructure technologies of computing, storage and networking have made tremendous progress. Besides the capacity and performance of physical devices, the virtualisation technologies offer effective resource management and isolation at different levels, such as Java Virtual Machines at the application level, Dockers at the operating system level, and Virtual Machines at the whole system level. Moreover, the network embedding [5] and software-defined networking [6] provide network-level virtualisation and control that enable a new paradigm of infrastructure, where infrastructure resources can be virtualised, isolated, and dynamically customised based on application needs.

The software-defined infrastructures, including Cloud, Fog, Edge, software-defined networking and network function virtualisation, emerge nowadays as new environments for distributed applications with time-critical application requirements, but also face challenges in effectively utilising the advanced infrastructure features in system engineering and dynamic control. This special issue on *"time-critical applications and software-defined infrastructures"* focuses on practical aspects of the design, development, customisation and performance-oriented operation of such applications for Clouds and other distributed environments.

# Content of the issue

This special issue publishes ten papers after a rigorous review process. The papers cover different topics related to time-critical applications and software-defined infrastructures, including infrastructure optimisation, task scheduling, network control, and software engineering.

An effective software framework for developing and operating time-critical in Cloud can reduce the development cost and improve the application delivery. The SWITCH workbench is an example developed in the EU Horizon 2020 SWITCH project that covers different phases across the application lifecycle, including virtual infrastructure planning and provisioning, application deployment, and runtime control. Štefanič et al. presented the workbench architecture and a novel cooperative programming and control model for developing and operating time-critical applications in Cloud environments [7]. In the same project, Knight et al. explore the utilisation of graphics processing units in Cloud applications, integrated into the SWITCH workbench to support time-critical applications [8].

Taal et al. profile the behaviour of partial critical path algorithms for planning customised virtual infrastructures for time-critical applications [9]. The partial critical path algorithm initially proposed in [10] for scheduling the execution of time-critical tasks on Cloud resources has been used as the basis to plan the infrastructure of networked virtual machines for application workflows with multi deadlines [11]. The paper characterises the effect of different greediness in critical path algorithms and enables the customisation of specific heuristics to reach the success of infrastructure planning.

Cotroneo et al. analyse the overloads caused by physical CPU contention in cloud infrastructures from the time-critical perspective of virtual network functions [12]. A general guest-level solution is proposed to protect applications from overloads also in the case of CPU contention. The approach can dynamically adapt the service throughput to the actual system capacity in case of traffic spikes and CPU contention to meet the latency requirements.

Zeng et al. focus on real-time communication in software-defined networking applied to industrial system automation and control [13]. The authors address two key technologies, time synchronisation and time slot-based packet switching, to handle the delay variance in networks. The system contains four modules for time synchronisation, flow classification, time slot assignment, and packet forwarding. The proposed solution can satisfy the requirements of time-sensitive networking for industrial automation with switch delay and achieves a jitter between 12.2μs and 0.12μs.

Simpkin et al. propose cognitive enabled services [14] to tackle the software engineering challenges in constructing distributed time-critical applications. A novel cognitive component uses a symbolic vector architecture to semantically represent service descriptions, workflows, and time-critical QoS and QoE constraints. The viability of this approach is demonstrated by both empirical and analytical proofs. Decentralised fitness functions are used in the solution for encoding time-critical workflow requirements and the enactment of services for on-demand resource discovery and allocation.

Zhang et al. present a novel framework, called Meteor, to optimise the Spark framework for supporting workloads with short-running time [15]. The framework utilises a fine-grained performance model to optimise the scheduling of tasks for new applications. The experiment on Amazon EC2 indicates that Meteor can double the performance of the original Spark.

Ouyang et al. investigate the stragglers problem in a cluster infrastructure and propose a dynamic server blacklisting based solution to avoid QoS violation for time-critical applications [16]. The existing solutions like speculative execution and blacklisting have limited consideration of server performance when performing backup in selecting new dynamic nodes. Ouyang et al. present a dynamic server blacklisting framework to increase the straggler mitigation effectiveness, with consideration of both historical and current behaviour of a server node.

López-Huguet et al. present an automated configuration of a software platform based on the Apache Mesos cluster for data analytics that supports horizontal and vertical elasticity to guarantee specific deadlines [17]. The proposed solution explicitly analyses the deployment times of different instances and include them in the meta information of applications components, software dependencies and configurations to build up the cluster. The approach builds up self-managed hybrid clusters that deal with different workloads and network requirements. The article describes the structure of the recipes, references to public code repositories, and discusses the results and limitations in several experiments.

Cokic et al. tackle the data and communication challenges in smart grid and propose a communication infrastructure to handle different time-critical requests from applications [18]. In this paper, they present a software-defined networking solution for distributed smart grid data aggregation. The proposed solution is validated via several cases, including automatic generation control during peak load, volt/var optimisation during peak load, and steady-state operation with static (background) traffic load.

**Acknowledge:**
We thank all the authors who contribute to this special issue and all the reviewers who support the review processes. Without their support, this special issue would not have been possible. We sincerely thank the editor in chief of the journal prof. Peter Sloot and the special content editor prof. Erik Cambria for supporting the special issue.

# References

1. Zambrano, A.M., Perez, I., Palau, C., Esteve, M., 2017. Technologies of Internet of Things applied to an Earthquake Early Warning System. Future Generation Computer Systems 75, 206–215. https://doi.org/10.1016/j.future.2016.10.009
2. Tudoran, R., Costan, A., Nano, O., Santos, I., Soncu, H., Antoniu, G., 2016. JetStream: Enabling high throughput live event streaming on multi-site clouds. Future Generation Computer Systems 54, 274–291. https://doi.org/10.1016/j.future.2015.01.016
3. Guna, J., Geršak, G., Humar, I., Song, J., Drnovšek, J., Pogačnik, M., 2019. Influence of video content type on users’ virtual reality sickness perception and physiological response. Future Generation Computer Systems 91, 263–276. https://doi.org/10.1016/j.future.2018.08.049
4. Koulouzis, S., Martin, P., Zhou, H., Hu, Y., Wang, J., Carval, T., Grenier, B., Heikkinen, J., Laat, C., Zhao, Z., 2019. Time-critical data management in clouds: Challenges and a Dynamic Real-Time Infrastructure Planner (DRIP) solution. Concurrency Computat Pract Exper e5269. https://doi.org/10.1002/cpe.5269.

5. Yuan, W., He, K., Han, G., Guan, D., Khattak, A.M., 2019. User behaviour prediction via heterogeneous information preserving network embedding. Future Generation Computer Systems 92, 52–58. https://doi.org/10.1016/j.future.2018.09.036.
6. Imran, M., Durad, M.H., Khan, F.A., Derhab, A., 2019. Toward an optimal solution against Denial of Service attacks in Software Defined Networks. Future Generation Computer Systems 92, 444–453. https://doi.org/10.1016/j.future.2018.09.022.
7. Štefanič, P., Cigale, M., Jones, A.C., Knight, L., Taylor, I., Istrate, C., Suciu, G., Ulisses, A., Stankovski, V., Taherizadeh, S., Salado, G.F., Koulouzis, S., Martin, P., Zhao, Z., 2019. SWITCH workbench: A novel approach for the development and deployment of time-critical microservice-based cloud-native applications. Future Generation Computer Systems 99, 197–212. https://doi.org/10.1016/j.future.2019.04.008
8. Knight, L., Štefanič, P., Cigale, M., Jones, A.C., Taylor, I., 2019. Towards extending the SWITCH platform for time-critical, cloud-based CUDA applications: Job scheduling parameters influencing performance. Future Generation Computer Systems 100, 542–556. https://doi.org/10.1016/j.future.2019.05.039.
9. Taal, A., Wang, J., de Laat, C., Zhao, Z., 2019. Profiling the scheduling decisions for handling critical paths in deadline-constrained cloud workflows. Future Generation Computer Systems 100, 237–249. https://doi.org/10.1016/j.future.2019.05.002.
10. Abrishami, S., Naghibzadeh, M., Epema, D.H.J., 2013. Deadline-constrained workflow scheduling algorithms for infrastructure as a Service Clouds. Future Generation Computer Systems 29, 158–169. https://doi.org/10.1016/j.future.2012.05.004.
11. Wang, J., Taal, A., Martin, P., Hu, Y., Zhou, H., Pang, J., de Laat, C., Zhao, Z., 2017. Planning virtual infrastructures for time-critical applications with multiple deadline constraints. Future Generation Computer Systems 75, 365–375. https://doi.org/10.1016/j.future.2017.02.001.
12. Cotroneo, D., Natella, R., Rosiello, S., 2019. Overload control for virtual network functions under CPU contention. Future Generation Computer Systems 99, 164–176. https://doi.org/10.1016/j.future.2019.04.007
13. Zeng, P., Wang, Z., Jia, Z., Kong, L., Li, D., Jin, X., 2019. Time-slotted software-defined Industrial Ethernet for real-time Quality of Service in Industry 4.0. Future Generation Computer Systems 99, 1–10. https://doi.org/10.1016/j.future.2019.04.009.
14. Simpkin, C., Taylor, I., Bent, G.A., de Mel, G., Rallapalli, S., Ma, L., Srivatsa, M., 2019. Constructing distributed time-critical applications using cognitive enabled services. Future Generation Computer Systems 100, 70–85. https://doi.org/10.1016/j.future.2019.04.010.
15. Zhang, H., Huang, H., Wang, L., 2019. Meteor: Optimising spark-on-yarn for short applications. Future Generation Computer Systems 101, 262–271. https://doi.org/10.1016/j.future.2019.05.077.
16. Ouyang, X., Wang, C., Xu, J., 2019. Mitigating stragglers to avoid QoS violation for time-critical applications through dynamic server blacklisting. Future Generation Computer Systems 101, 831–842. https://doi.org/10.1016/j.future.2019.07.017.
17. López-Huguet, S., Pérez, A., Calatrava, A., de Alfonso, C., Caballer, M., Moltó, G., Blanquer, I., 2019. A self-managed Mesos cluster for data analytics with QoS guarantees. Future Generation Computer Systems 96, 449–461. https://doi.org/10.1016/j.future.2019.02.047.

18. Cokic, M., Seskar, I., 2019. Software-defined network management for dynamic smart GRID traffic. Future Generation Computer Systems 96, 270–282. https://doi.org/10.1016/j.future.2019.02.022.