

Validation of Control Systems with Heterogeneous Digital Models and Virtualization Technologies


Kirill Semenov ¹, Vitaly Promyslov², Alexey Poletykin³

Abstract: A modern instrumentation and control system is a cyberphysical system that combines hardware and software components in a network aware environment. The virtualization technologies become common in the development and operation of control systems: they can be used as a tool for system life-cycle extension or for the creation of a comprehensive digital model of the system, a digital twin. The paper deals with the problem of the design of digital twins of cyberphysical systems of an Industry 4.0 era. It analyses and compares the properties of the model in implementation to the modeling of the digital and physical components of a cyberphysical system. A heterogeneous digital model combining virtual machines and emulators and some real software and hardware is recommended for the purposes of testing and verification of complex cyberphysical systems. The role and tradeoff of virtualization environment for control system simulation are discussed. A real use-case of digital modeling is described and discussed. The practical aspects of assessing the performance characteristics in virtual environment and technologies for keeping the synchronous operation of the components the digital twin and the problems of timekeeping within the heterogeneous digital model are discussed.

1 Introduction

The concept of Industry 4.0 and the term itself were phrased at the first time in 2011, by the Germany working group on the vision of industry development prospects [KLW11]; the other countries had performed researches of the same kind [HPO15] as well. In a short period, the concept and the term have become widespread. The authors of the concept [KWH13] consider the industrial environment of the future as a flexible and adaptable cyberphysical system (CPS) that unites manufacturing, warehousing, and logistics through the medium of Internet of Things (IoT). The list of digital technologies that must be used by a manufacturing company of Industry 4.0 includes (see, for example [HPO15], [SS16]) cloud and fog computing, virtualization, artificial intelligence, new data transmission protocols for the IoT and many others.

Such systems are very complex, and often have an expensive and extensive life cycle. A way of decreasing the complexity of the design and verification and validation (V&V)

¹ V.A. Trapeznikov Institute of Control Sciences of the Russian Academy of Sciences, Moscow, Russia, semenkov@mail.ru,  <https://orcid.org/0000-0003-0865-9072>

² V.A. Trapeznikov Institute of Control Sciences of the Russian Academy of Sciences, Moscow, Russia, vitalionics@gmail.com

³ V.A. Trapeznikov Institute of Control Sciences of the Russian Academy of Sciences, Moscow, Russia, poletik@ipu.ru

of such systems is the creation of their complete digital model. In practice, a problem occurs: a digital twin built on a top of pure digital technology often does not well fit to the expectations due to incomplete formal description of the CPS or bad compromise between discrete nature of the computer model and continuous time in actual system [Ok19]. The possible solution is building of a heterogeneous model that combines original and simulated digital parts with some actual analog components. The heterogeneous digital model itself is actually a CPS. Those heterogeneous models have significant advantages [Qu16] compared to purely computer or purely imitation models. Compared with the first ones, they have the ability to more accurately simulate the object behaviour, especially its timing characteristics; compared with the second ones, they are more flexible and easier to use. The heterogeneous digital models are applicable to many aspects of the industrial system development as performance validation, safety, and security assessment.

In the work, we discuss several questions and technical problems in development and application of digital models for the Instrumentation and Controls (I&C) system: the selection of the proper architecture for the digital model and balance between the original digital and analogue parts of the model. We argue the question of having simultaneously discrete and continuous time scale and share the experience received in synchronizing the components of the digital twin using NTP protocol [Mi91].

2 The Review of I&C Modelling Approaches

In this section, we briefly review the main types of model and describe our approach for the design of a digital model for a real I&C system. In general, an I&C system consists of a set of sensors, actuators and the computers with software implementing the control algorithms and human-machine interfaces (HMI). So, an instrumentation and control system itself is a CPS. The validation of the CPS is usually performed in the conditions as close as possible to the operational conditions. The test cases should cover both normal and stress scenario of system functioning. The validation process has some difficulties:

- The absence of the actual control object or its components during design and partially in other stages.
- The impossibility to validate some modes due to risk of the object physical destruction or high costs of testing.

The modelling in many cases resolves these problems. The usability of the modelling depends on the used framework. Tab. 1 shows the comparison of different models and the regard on their potential of use for the physical (Phys.) or digital (Dig.) component assessment.

The most comprehensive and rigorous type of the models is full-scale model. The full-scale test prototypes can be built to try-out and validate the interaction of system elements. There

Tab. 1: Comparison of properties of models. The “+” or “±” signs mean whether the model suits or does not suit for a specific purpose

Type of model		Motivation	Phys.	Dig.
1	Analytical	Description of a physical object behavior	+	-
		Verification and validation of algorithms (system grey/white-box design)	-	+
		Verification of timing characteristics ⁴	+	±
		Staff training	+	±
		Design of the control (system black box design)	+	+
2	Statistical	Reliability and stability estimation	+	+
3	Functional	System design	+	+
4	Data and data flow	Data representation and system logics without regard to real-time system behavior	-	+
5	Full-scale	Validation of system design	+	+
		Validation of models (same as 1–4)	+	+
		Validation of time behavior	+	+
		Validation of system safety and security	+	+
		Staff training	+	+
6	Digital twin (virtual model, pure digital model)	Validation of system logical structure and interfaces	+	+
		Validation of discrete (state-by-state) time behavior	+	+
		Validation of system security	-	±
		Staff training	+	+
7	Heterogenous: virtual and some real components	Combines all advantages of the models of types 5 and 6	+	+

the preliminary check-out is performed with the real equipment; the final integration with the physical object is carried out at the commissioning stage. The process consumes time and resources. With the progress of the computational facilities, a concept of digital twins of CPS has been gaining popularity as alternative of the full-scale models. For example, Lemay et al. ([LFK13]), using a number of virtual machines running within a computational cluster, created a digital model for a SCADA (supervisory control and data acquisition) system of a power plant. They used some simulations of physical processes and PLCs, sensors, and actuators. The model gave the possibility to imitate a bunch of cyberattacks to a reference SCADA system, albeit not showed high productivity. Alves et al. ([A118]) for security measures tests applied modularity principles to the design and construction of a digital model of a SCADA system. They implemented every element of the SCADA (server,

workstation, PLC, sensor et cetera) as a separate software module, that allows model scaling in a wide range.

However, it is necessary to understand and take into account the limitations of digital models. First, with a digital model, we hardly (if ever) can obtain any data about the productivity of a real system. The system productivity depends on specific models of installed computers and controllers, network capacity, and many other conditions. The model allows getting just some productivity estimates like algorithm performance. Second, the hardware and equipment of any specific manufacturer have its features and restrictions, some internal details of equipment functioning are company secrets, so they cannot be implemented entirely in a digital model. It means the digital model will use some “average,” “neutral” models of the equipment, which also does not allow getting an accurate model. Even if we have a perfect model of a hardware, we shall have in mind the problems of time correlation between the dynamics of an actual CPS and its digital model. Indeed, the processor time of program execution flow is not the real physical time, because a computer represents the change of time by incrementing a hardware-dependent counter. Therefore, the time-related properties of the device digital model can differ from the properties of the real device. Lee and Seshia ([LS17]), in chapter 1 of their book, describe the problem in detail.

So, a digital twin for an I&C system should unite the three kinds of models: models of physical object itself, models of control tools (e.g. actuators), and control software or its model. We summarize these considerations in Tab. 2.

Tab. 2: The principles of I&C digital twin design.

	Physical object	Controllers/sensors	Software
Type of model	analytical (equations)	emulation; black-box software models	—
Representation of time	an abstraction: an argument of the equations	a counter of ticks (hardware dependent)	a counter of ticks (hardware dependent)
Implementation within a digital twin	the equations are solved separately, the twin uses the results and can approximate them	a piece of software running on a real or virtual computer	a piece of software running on a real or virtual computer
Technologies	numerical simulation, clusters, supercomputers	emulation, simulation	virtualization, cloud computing

The maintaining of a single timescale is of great importance for I&C systems. The digital twins are often implemented in a cloud environment with a large number of virtual machines, they have multiple concurrent context-switching events between the processes and virtual machines. It gives rise to the problem of time synchronizations for the components within a model. Currently, the issue of synchronization of a system of virtual machines has little been studied, but the problem is recognized. Thus, VMware [Vm11] says that differences

between virtual and real machines “can still sometimes cause timekeeping inaccuracies and other problems in software running in a virtual machine.” Therefore, there is always a risk that the necessary synchronization accuracy would never be achieved in the digital model of the cyberphysical system.

We suppose the restrictions described above can be partially removed in heterogeneous models that combine a digital model with some real parts. The hybrid model compare to the only digital, or full scale model has advantages: better simulation of the functionality of the actual I&C system and timing characteristics correspond to the purely digital model and reduced costs and simplified maintenance compare to the full scale prototype.

For example, a researcher can include in the model a few real sensors and PLCs. The sensors would transmit their output to the virtual medium; the model of the physical process would calculate input signals for PLCs and pass the signals to the standard interfaces of the PLCs. Partial integration of that kind would allow us to check the temporal characteristics of some control signals and facilitate the commissioning process.

3 The Practice of Digital Twin Design and Operation

We designed and implemented a heterogeneous digital model of the upper-level control system (ULCS) of I&C system of a nuclear power plant ([Po17]). The ULCS is a part of the I&C system and is used to provide HMI functions, gather and integrate information from I&C subsystems and perform self-diagnostics. The ULCS consists of servers, active and passive network equipment, workstation, auxiliary equipment of the cabinets (uninterruptible power sources, printers et cetera). The information is transferred over Ethernet LAN; all key nodes and data paths are redundant and work in parallel providing hot-spare redundancy. The ULCP software works under industrial Linux-based operating system LICS OS ([LI19]). The precise time sources provide a unified time scale within the system.

We implemented a heterogeneous digital model (Fig.1) that includes some real hardware of ULCS system and a few dozens of virtual components representing ULCS computers and network devices running in virtual environment. The real hardware includes a timeserver, an Ethernet switch cabinet, a workstation cabinet with operator terminal, server cabinet. The real timeserver acts as the time synchronization source via the Network Time Protocol (NTP). Switch cabinet and server cabinet hardware elements are used, first, to integrate the model with some real hardware and software running on a “bare-metal” computer and, second, for I&C self-diagnostics subsystem verification. The virtual machines for every server and workstation physically run on a host server under LICS OS and QEMU/KVM hypervisor ([KY20], [QE20]); the workstation physical computer runs under LICS OS as well. The software imitators represent the adjacent subsystems of the NPP I&C system. The modelling of network interaction was a significant part of the task. The I&C is a distributed system, the components communicate with one another via TCP/IP and UDP/IP protocol. As we have told, the operational servers, workstation and Ethernet network are redundant,

and the model must reflect both logical and physical redundancy of the system. As the model is a heterogeneous one, it is assumed to interact with external real hardware via some network protocols. We implemented the ULCP network structure and topology in the virtual environment, but made the interaction with the real hardware possible. Virtual switch software OpenVSwitch [Op20] is a tool for network topology construction within the model: every real switch is mapped to a software switch, and the commutation of virtual machines and virtual switches within the virtual network corresponds to the ULCS network topology. VLAN assignment within software switches allows to model physical separation of redundant Ethernet channels. Thus, we model network redundancy where any single failure in the network path does not break the system connectivity.

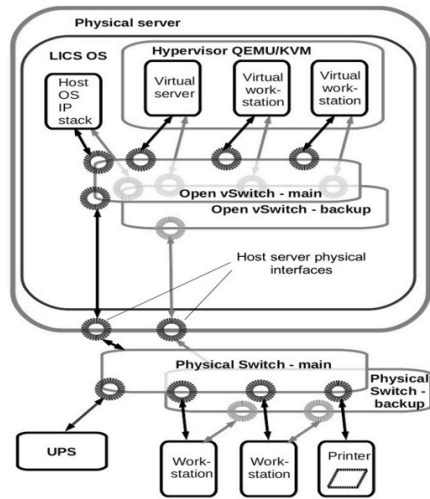


Fig. 1: The outline of heterogeneous digital model for the upper-level control system of an I&C system of NPP

To work with operator graphical environment, we pass the graphics via Spice network protocol [Sp20] to the real workstation; both single and multiple monitor workstation was modelled. Spice client software was also installed onto the virtual machines.

The created heterogeneous digital model allows us to test the interaction of ULCS software components, prepare software deployment to the real complex, test software updates before the deployment, verify security controls. However, the operation process showed some drawbacks in the heterogeneous model design. In the beginning, the computer hardware emulated by the hypervisor (like controllers, network adapters) was chosen to be maximally close to the real one. But the performance of software emulation of the devices did not satisfy the needs of I&C system, the model showed poor network throughput. So we had to switch over to paravirtualized devices (for the description of software emulated and paravirtual devices see e.g. [SMP19], [Go11]) and to give up the idea of maximal similarity between the emulated and real hardware. Graphical mode and HMI is another weak point

of virtual models. Virtual environment usually does not allow achieving the same graphics productivity as bare-metal solutions. During the modelling, we did not intend to achieve the graphical productivity of a real system but we were going to test and verify the HMI. We managed to achieve acceptable graphics throughoutput and latency for our purposes. In spite of the restrictions and drawbacks, the designed and constructed model became a good platform for the software complex testing and verification of an I&C system. The heterogeneous digital model allows significantly decrease duration of the validation for the ULCS. The cloud-based platform for the model provides effective and secure remote access for the personal involved in system design during the pandemic lockdown restriction imposed by government officials.

4 Some Aspects of Clock Synchronization in a Virtual Environment

The synchronization of virtual system clock and timekeeping in virtual models is a problem of great importance and presently investigated poorly. The developers of virtualization systems and timekeeping hardware limit themselves to general vague recommendations (see, for example [Bu17]). Below we are going to describe our approach to synchronization that we used during creation of the heterogeneous digital model. A virtual server or workstation is a virtual machine (VM) of x86-64 architecture, the virtual machines and hypervisor work under operating system (OS) LICS, a specialized Linux distribution; QEMU/KVM is used as the virtualization software. A GPS/GLONASS NTP timeserver serves as a synchronization source for the virtual model.

Computer clocks usually should conform to the following requirements: they should not stop; they should not go backwards; their resolution must be sufficient for the application purposes; application software must be able to read the clock data; clock access time should be small. Historically the computers of x86 architecture used a set of times (PIT, RTC, HPET and others), but now the main timer is the TSC, Tick Step Counter, a counter of CPU cycles. In modern processors, this register increases evenly and does not depend on the dynamically changing CPU frequency; in addition, in multiprocessor systems, the TSC registers in all processor cores and processors within the same motherboard are synchronized.

TSC emulation is available in Linux-based virtual machines; however, for the stable timer operation during virtual machine migration and CPU frequency changes, major Linux developers (see e.g. [Re20]) recommend using the paravirtual (that is, interacting directly with the host clock) clock driver “kvm-clock”. This is the default clock driver for QEMU/KVM.

However, the experiment shows that the paravirtual driver does not allow to achieve acceptable synchronization accuracy for our virtual model of the I&C system. The experiments show that, first, the virtual machine clock is constantly lagging behind the timeserver clock and, second, the PLL (phase-locked loop) cannot adjust the virtual machine clock. The saw-tooth kind of the curve means the NTP protocol after a period of linear growth sets

the virtual machine clock forcibly. NTP statistics also shows that for this case the clock rate exceeds $500 \cdot 10^{-6} \cdot 10^{-6}$ sec/sec, or 43 sec/day. We assume that simultaneous use of paravirtual clock driver on a number of competing VMs causes delays in hypervisor response to a VM request, so the VM clock latency increase up to 100 times.

After the abandon of the paravirtual driver and switching to the TCS clock in emulation mode in the VM (i.e., refusing to hard link the VM clock to the hypervisor clock), the situation becomes normal (see Fig. 2).

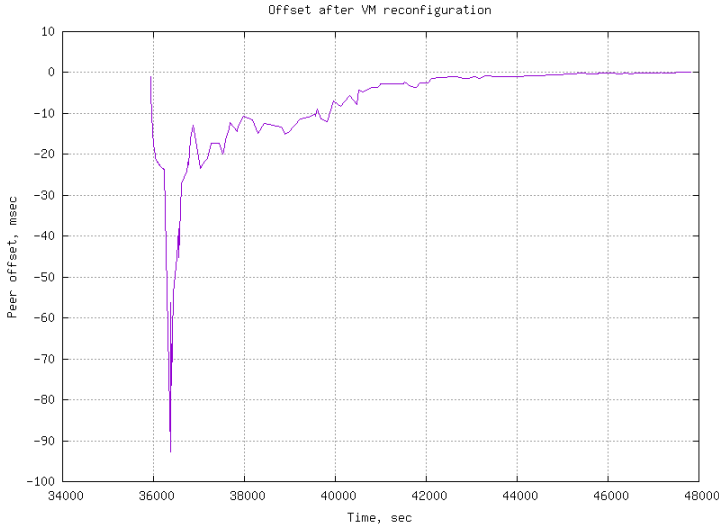


Fig. 2: Clock offset between a virtual machine and NTP server when the virtual machine uses TSC-emulation clock driver

As we can see, in this case, NTP managed to synchronize the timeserver and the VM within a few milliseconds. So, NTP seems to be suitable for application in virtual models but for now its application needs further researches.

5 Conclusions

In the paper, we deal with the problem of the designing the architecture and practical implementation of the digital model of the instrumentation and control systems (I&C systems). I&C system, being a cyberphysical systems (CPS), includes many hardware digital and analogue components. It makes time landscape of the I&C system comprehensive because it combines and mixes discrete and continuous time scale.

The two problems of modelling are considered: the extension of the model life span over whole life cycle of the real objects and accurate simulation of the performance and timing

characteristics of the I&C system. The first problem is solved in a framework of the digital twin approach and heterogeneous model. The digital twin is a useful tool for working out the interaction of system elements, their functional relations and software quality and hardware reliability estimation. To achieve high reference level between the I&C system and model the actual software components of I&C system are wrapped using the virtualization layer. To mitigate the second problem, real hardware devices are used in a time critical functions simulation. However, such digital twin, being deployed in virtual environment, do not provide the ability to measure the temporal characteristics of digital components I&C system and that question still challenging. A useful approach dealing with I&C system performance assessment using deterministic queuing theory is provided in [BP19]. We demonstrated the efficiency of a proposed heterogeneous digital model of I&C systems in the course of works on the test and verification of and upper-level control system for a nuclear power plant. The digital components of the model are about 40 virtual machines for servers and workstations, software switches; the real components of the model are some devices (servers, workstations, auxiliary hardware). This solution allows us to significantly reduce the amount of hardware at the test site, the complexity and labour intensity of configuration change during the development, the total cost of test site construction while maintaining a high reference of obtained results. The cloud-based platform used for system's digital model successfully provides secure and moderate development environment for the engineers and system designers. The existence of the model allows uninterrupted development and verification process during the lockdown period without violation of social distance restrictions.

In the process of model implementation, we put attention to the timekeeping problem. For the modelled system, the timekeeping is performed by means of the NTP (Network Time Protocol). We conclude that paravirtual clocks for virtual machines don't fit for timekeeping in high load environment with intensive exchange between virtual machines. The use of emulated TSC (Tick Step Counter) processor register allows achieve time synchronization. However, the search of optimal way of timekeeping is under research.

That heterogenous digital model solves problems of an accurate performance validation of the CPS and verification of the components with incomplete mathematical description. The balance between original software and hardware parts and simulated ones of the digital model is a challenging question as well as architecture solutions of such systems. The experience gained during the creation of the digital model in cloud-based environment can help to solve the problem of the life span prolongation of the software when hardware is outdated and not more available on the market. This already have been done for the upper level components of an I&C system [SMP19] and might be feasible for the controllers and even smart sensor level.

The reported study (Sections II) was partially funded by RFBR, project number 19-29-06044.

Bibliography

- [Al18] T. Alves, R. Das, A. Werth, and T. Morris, "Virtualization of SCADA testbeds for cybersecurity research: A modular approach", *Computers and Security*, 77, pp. 531–546, 2018.
- [Bu17] Burnicki, M., *Time Synchronization in Virtual Machines* // https://kb.meinbergglobal.com/kb/time_sync/time_synchronization_in_virtual_machines, 2017.
- [BP19] A.A. Baybulatov, and V.G. Promyslov "Control System Availability Assessment via Maximum Delay Calculation", *Proceedings of the 2019 International Conference on Industrial Engineering, Applications and Manufacturing (ICIEAM)*. Sochi: IEEE, 2019.
- [Go11] Y. Goto, "Kernel-based Virtual Machine Technology," *FUJITSU Sci.Tech. Journal.*, vol. 47, # 3, pp. 362–368, 2011.
- [HPO15] M. Hermann, T. Pentek, and B. Otto, "Design Principles for Industrie 4.0 Scenarios: a Literature Review", https://www.researchgate.net/publication/307864150_Design_Principles_for_Industrie_40_Scenarios_A_Literature_Review, 2015.
- [KLW11] H. Kagermann, W-D. Lukas, and W. Wahlster, "Industrie 4.0: mit dem Internet der Dinge auf dem Weg zur 4. industriellen Revolution," *VDI Nachrichten*, No. 13, 2011 (in German).
- [KWH13] H. Kagermann, W. Wahlster, and J. Helbig, "Recommendations for implementing the strategic initiative INDUSTRIE 4.0: securing the future of German manufacturing industry; final report of the Industrie 4.0 working group", *Forschungsunion*, Berlin, 2013.
- [KY20] KVM Contributors, <https://www.linux-kvm.org/page/Documents> (last access on Apr. 26 of 2020).
- [LFK13] A. Lemay, J. Fernandez, and S. Knight, "An isolated virtual cluster for SCADA network security research", *ICS-CSR*, 2013.
- [LI19] LICS, RF registration number 2019618036, <https://www1.fips.ru/publication-web/publications/document?type=doc&tab=PrEVM&id=07B0B75D-B08F-4A7B-BF76-011ED855B976>, 2019. (in Russian).
- [LS17] E.A. Lee, and S. Seshia, "Introduction to Embedded Systems – A Cyber Physical Systems Approach", Second Edition, MIT Press, 2017.
- [Mi91] Mills, D.L. Internet time synchronization: the Network Time Protocol. *IEEE Trans. Communications COM-39*, 10 (October 1991), 1482-1493.
- [Ok19] Oks, Sascha Julian & Jalowski, Max & Fritzsche, Albrecht & Moeslein, Kathrin. Cyber-physical modeling and simulation: A reference architecture for designing demonstrators for industrial cyber-physical systems. *Procedia CIRP*. 84. 257-264. 10.1016/j.procir.2019.04.239, 2019
- [Op20] OpenVSwitch Team, <http://docs.openvswitch.org/en/latest/>, (last access on Apr. 26 of 2020).
- [Po17] A. Poletykin, E. Zharko, N. Mentgazetdinov, and V. Promyslov, "The new generation of upper levels systems and industry 4.0 conception in NPP APCS", *Proceedings of the 10th International Conference "Management of Large-Scale System Development" (MLSD)*. Piscataway, USA: IEEE, vol. 1. pp. 1-5, 2017.

-
- [QE20] QEMU Contributors, https://wiki.qemu.org/Main_Page (last access on Apr. 26 of 2020).
- [Qu16] Quadri, Imran Rafiq et al. “Modeling Methodologies for Cyber-Physical Systems : Research Field Study on Inherent and Future Challenges.”, 2016.
- [Re20] RedHat, KVM Guest Timing Managemeng // https://access.redhat.com/documentation/en-us/red_hat_enterprise_linux/7/html/virtualization_deployment_and_administration_guide/chap-kvm_guest_timing_management, 2020.
- [SBK17] M. Schütze, S. Bondorf, and M. Kreider, “Verification of the FAIR Control System Using Deterministic Network Calculus”, ICALEPS2017 Proceedings, pp. 238-245, DOI: 10.18429/JACoW-ICALEPCS2017-TUCPL06, 2017.
- [SMP19] K.V. Semenov, N.E. Mengazetdinov, and A.G. Poletykin, “Extending Operation Lifespan of Instrumentation and Control Systems with Virtualization Technologies”, Proceedings 2019 International Russian Automation Conference (RusAutoCon) <https://ieeexplore.ieee.org/document/8867595>, 2019.
- [Sp20] Spice project team, <https://www.spice-space.org/>, (last access on Apr. 26 of 2020).
- [SS16] A. Schumacher, S. Erol, and W. Sihn, “A maturity model for assessing Industry 4.0 readiness and maturity of manufacturing enterprises”, *Procedia CIRP*, 52, pp. 161–166, 2016.
- [Vm11] VMware Inc., “Timekeeping in VMware Virtual Machines. Information Guide”, <https://www.vmware.com/content/dam/digitalmarketing/vmware/en/pdf/techpaper/Timekeeping-In-VirtualMachines.pdf>, 2011, (last access on Apr. 26 of 2020).