Service-oriented tools for automating digital twin development

Roman Kostromin, Alexander Feoktistov and Mikhail Voskoboinikov

Matrosov Institute for System Dynamics and Control Theory of SB RAS, Lermontov St. 134, Irkutsk, 664033, Russia

Abstract
The paper represents a prototype of service-oriented tools for developing digital twins. These tools automate most of the stages in preparing and carrying out a computational experiment reducing the possibility of human error. Computational experiments based on simulation modeling are performed using the proposed tools. As an example in applying the represented tools, a digital twin of a heat pump used as environmentally-friendly equipment of an infrastructure object on the Baikal natural territory is considered. Owing to the growing anthropogenic load, special attention is paid to the objects located at the coast of Lake Baikal. During the simulation modeling, both the heat pump performance characteristics and retrospective meteorological data are used to predict climatic conditions and select optimal control parameters. In addition, the computational experiment has obviously shown that heat pump use significantly reduces the harmful effect on the environment.

Keywords
digital twin, simulation modelling, microservices, environmentally-friendly equipment

1. Introduction

Nowadays, environmental problems, in particular the reduction of CO_2 emissions, are the subject of research by the scientific community around the world [1]. One of the directions of such research is the introduction of environmentally-friendly equipment. In the era of the development of Industry 4.0 technology, a digital twin is an effective tool for analyzing the benefits of using the equipment. Malakuti et. all [2] describe a digital twin as "a formal digital representation of some asset, process or system that captures attributes and behaviors of that entity suitable for communication, storage, interpretation or processing within a certain context".

As a rule, the object behavior study using a digital twin is based on the simulation modeling of object operation processes [3]. Such a simulation often requires applying the Monte Carlo method to provide the reliability of stochastic simulation results, often generates big data, and thereby leads to the need for High-Performance Computing (HPC) and Virtual Machines (VMs) use [4]. Moreover, within the simulation modeling, it becomes necessary to vary the values of the structural and operational parameters of the object in order to determine the optimal ones. Thus, the problem arises in organizing and carrying out parameter sweep computation [5]. In addition, current and retrospective data on the object equipment operation are necessary for verification and validation of simulation models, as well as to object diagnostic and control [6].

In terms of system architecture, there is a tendency to move from monolithic and general-purpose applications to compact and specialized web services [7].

To this end, we have developed a prototype of new microservice-oriented tools for constructing sets of digital twins representing various subsystems of an infrastructure object. In addition, we have created a system of intelligent agents who are delegated the rights and obligations of subjects related to ensuring the object operation and consuming the resources and services provided to them.

Unlike the well-known toolkits considered in [8-10], the developed tools provide comprehensive support for the following operations:

- Partial automation of the model design process;
- Verification and validation of models;
- Adaptation of models to a specific subject area; support for hybrid modeling;
- Collection and processing of semi-structured data;
- Converting data to the formats required in models;
- Multi-criteria analysis of simulation results;
- Automation of the formation of a web interface for performing the above-listed tools.

We applied the developed tools in the study of the heat supply system that uses heat pumps for infrastructure objects of the Baikal natural territory.

2. Tools for Automating Digital Twin Development

Within the proposed approach, the software of a digital twin is a distributed applied software package. Such software includes a set of applied modules (for example, simulation models). Separate models are represented by services. Planning and executing a workflow-based service composition are supported.

A feature of multiple runs of simulation models is the necessity to save different versions of the experiment and the ability to run a specific version for computations. Therefore, our tools were developed using the concept of parent and inherited services. Parent services are implemented applying the Node.js language within the REST approach. They contain the basic operations required for all stages in preparing and carrying out experiments. Among such operations are forming a web interface, generating model specifications, parsing generated specifications, creating and queuing VM-based jobs, configuring and monitoring a computing environment, and analyzing simulation modeling results.-jobs generation, and configuration, environment monitoring, analyzing simulation modeling results.

An inherited service is considered a new computational experiment that inherits the basic capabilities of the parent service, but it is used in a new configuration. This service configuration is a separate project that can be published for use in other experiments. A detailed description of model parameters, experiments, and requirements to the environment is indicated in [11], where the specification example in the JSON format is presented.

The creation of a new project includes the processes of importing and parsing the simulation model specification, as well as preparing its input data. Also, it is possible to generate input data variants based on the model specification. The VM lifecycle management is implemented using an add-on over OpenStack, in which VM-based jobs are transferred for execution to the queue system of the computing cluster [12]. All standard GPSS simulation reports from the experiment and other output files are collected for processing and analysis.
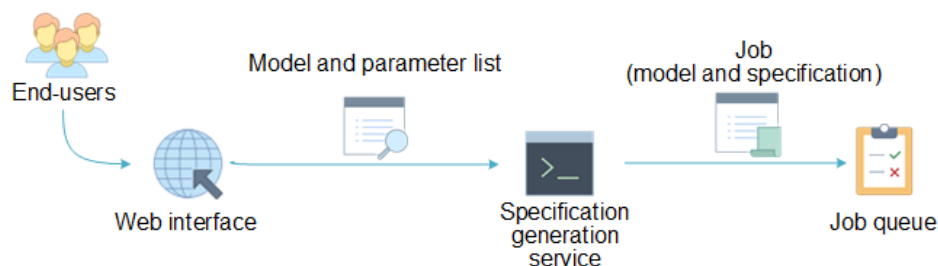

Figure 1: The scheme of the user's interaction with tools

The general scheme of the user's interaction with our tools is shown in Figure 1. The user creates a new project using the web interface and loads the simulation model and list of its parameters. Next, the specification generation service forms a VM-based job that contains the model and its specification. This job is transferred to the queue of the resource management system located on the computing cluster.

Figure 2 presents a diagram of parent service operations. New jobs and their specifications from the queue are processed by the service for processing the queue and starting simulation modeling. If an additional VM needs to be started, a request is made to the OpenStack add-on management service. As a result, the required number of VMs are launched in the distributed environment. Each instance of the model runs in a specific VM and receives its own set of input data from the retrospective DataBase (DB) with means of the monitoring system. The experiment results are stored in the computational DB. Remote command execution under Windows operation systems is implemented using SSH and Cygwin. VM configurations are managed by the Ansible system [13].
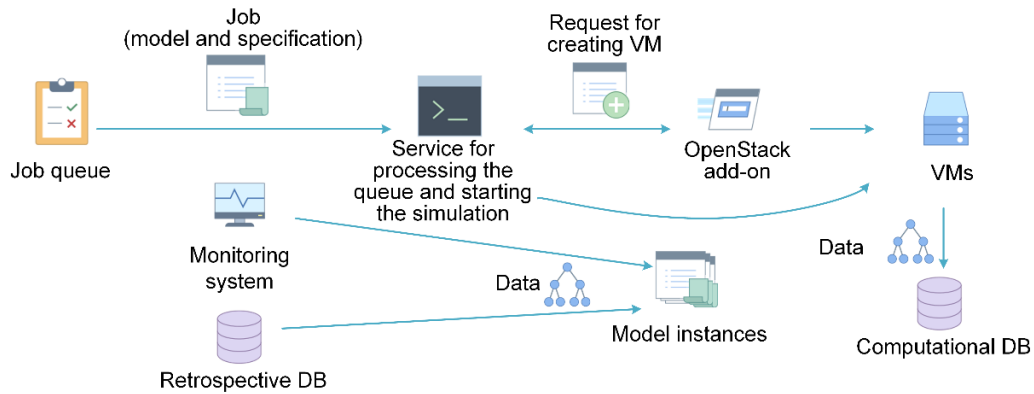

Figure 2: Operation Scheme of the parent service of processing the queue and starting the simulation modeling

3. DT of heat pump

In this section, we demonstrate applying the developed tools in the study of the heat supply system that uses heat pumps for infrastructure objects of the Baikal natural territory. The heat pump makes it possible to reduce the cost of electricity and decrease the harmful effect on the environment due to the use of natural heat sources (soil or water) [14]. The models are developed by problem domain experts. Within the digital twin, DevOps is supported in relation to the developed models. [15].

In our approach, the digital twin of the heat pump is essentially a modular application. A module is a sequential or parallel program. In our example, a module is a model on the GPSS language. Therefore, we launch VMs that contain instances of GPSS-model and tie them to different processor cores of computing environment nodes because GPSS-models are run under Windows operation systems in single-threaded mode only.

Data for executable modules are transferred from different sources. Accumulated meteorological data and equipment operation information are stored in the retrospective DB (Figure 2). The implementation of their collection and primary processing requires the use of additional software and hardware. We use microcomputers and software agents located on them [16]. The agents perform monitoring data received from sensors installed on heat pumps, pre-processing the obtained data, and periodic uploading the pre-processed data to a central database. It is important to note that during the modeling process, simulation models (package modules) interact with agents and request the necessary information via the REST-API.

For each digital twin, the specification is generated based on the simulation models and their input parameters. It contains the data sources, formats of input and output files, parameters for generating variables, and additional system information.

The input data are the results of meteorological observations in the Baikal coastal area over the past 8 years. This data makes it possible to simulate climate change and control the heat pump operation to ensure a comfortable indoor temperature.

We studied the effect of reducing CO_2 emissions by means of simulation in replacing a coal-fired boiler with one of six pump types (Corsa 55, Corsa 70, MOTEN-18D 57, MOTEN-18D 70, BROSK Mark Prom 58, and BROSK Mark Prom 71) from three Russian manufacturers taking into account capital investment. Within the experiment, 24 variants of input data were prepared to simulate the

heat pump operation. Each variant contains different combinations of values for the pump type, the need for drilling, and daily object service capacity in the number of persons. The two criteria for optimization are CO_2 emission reduction and capital investments with respect to the following optimality conditions: maximized quantity and minimized size correspondingly.

In the process of modeling, the object service capacity is 60 and 80 persons. Figure 4 (Figure 5) demonstrates the CO_2 emission reduction (a) and capital investments (b) for object service capacity equal to 60 (80) persons.
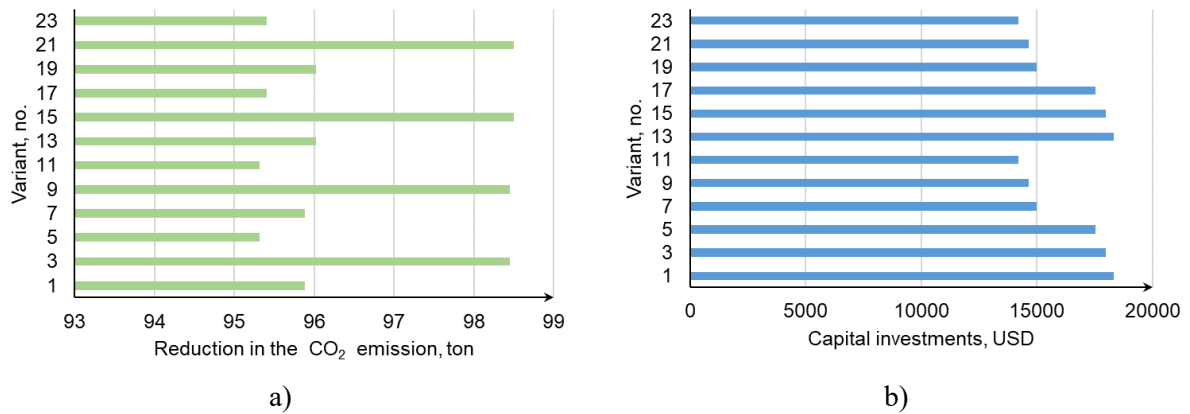

a)

b)

Figure 3: Reduction in the CO_2 emission (a) via capital investments for the variants with 60 person (b)
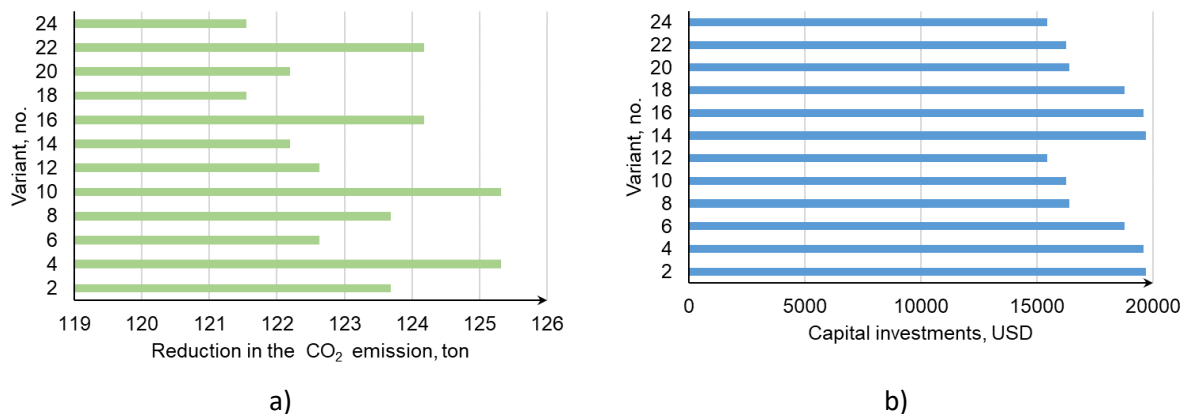

a)

b)

Figure 4: Reduction in the CO_2 emission (a) via capital investments for the variants with 80 person (b)

For selecting optimal parameters for operating environmentally-friendly equipment of the studied objects, we apply the following three methods of discrete multi-criteria analysis from [17]: Lexicographic (1), Majority (2), and Pareto-optimal (3). The Lexicographic method provides selecting the optimal variant by sequentially comparing the criteria, sorted by their significance. In our example, we can prefer to increase the CO_2 emission reductions, or vice versa, save capital investments. When the criteria are of equivalent importance, the majority method gives us the variant that includes the largest number of these criteria with the best values. With the means of the Pareto-optimal selection, we find all the variants that are incomparable with each other, and at the same time, surpass all the remaining ones. The results of applying the aforementioned methods are shown in Table 1.

Table 1
Results of multi-criteria analysis

Method	Capacity in persons	Variant, no. number	CO_2 emission reduction, ton	Capital investments, USD	Heat pump type
1	60	16	124.18	19605	Corsa 70
1	80	4	125.32	19605	Corsa 70
1	60-80	22	124.18	16271.67	Corsa 70
2	60	9	98.45	14661.67	Corsa 55
2	80	15	98.5	17995	Corsa 55
2	60-80	22	124.18	16271.67	Corsa 70
3	60	18	121.55	18800	MOTEN-18D 70
3	80	12	122.63	15466.67	MOTEN-18D 70
3	60-80	24	121.55	15466.67	MOTEN-18D 70

Thus, we can see that the selection of specific equipment depends on both its structural and operational parameters and methods of multi-criteria analysis.

4. Conclusions

In the paper, we address the development and applying the prototype of service-oriented tools for creating and carrying out DTs of infrastructure objects. These tools make it possible to perform simulation modeling based on parameter sweep computing followed by a multi-criteria selection of parameters for the operation of environmentally-friendly equipment of the studied objects. The advantages in applying the developed tools are the automation of software deployment and configuration, significant reduction of the large-scale experiment makespans, and increase of the computation reliability.

5. Acknowledgements

6. References

[1] A. Mardani, D. Streimikiene, F. Cavallaro, N. Loganathan, M. Khoshnoudi, Carbon dioxide (CO2) emissions and economic growth: A systematic review of two decades of research from 1995 to 2017, Science of the total environment 649 (2019) 31–49. doi: 10.1016/j.scitotenv.2018.08.229.

[2] S. Malakuti, P. van Schalkwyk, B. Boss, C. Ram Sastry, V. Runkana, S.W. Lin, S. Rix, G. Green, K. Baechle, N.C. Nath, Digital Twins for Industrial Applications. Definition, Business Values, Design Aspects, Standards and Use Cases: White Paper, in: Industrial Internet Consortium: Milford, MA, 2020, pp. 1–19.

[3] A.V. Petrov, Simulation as the basis of digital twin technology, Proceedings of Irkutsk State Technical University 22(10) (2018) 56–66. doi: 10.21285/1814-3520-2018-10-56-66. (In Russian).

[4] C.A. Varotsos, V.F. Krapivin, A new big data approach based on geoecological information-modeling system, Big Earth Data 1(1-2) (2017) 47–63. doi: 10.1080/20964471.2017.1397405.

[5] G. Henley, A. Card, T. Haup, Using High Performance Computing for Parameter Sweep of S3D Applications, in: IEEE Electric Ship Technologies Symposium (ESTS), IEEE, Washington, 2019, pp. 273–277. doi: 10.1109/ESTS.2019.8847871.

[6] Y. Lu, C. Liu, K.I. Wang, H. Huang, X. Xu, Digital Twin-driven smart manufacturing: Connotation, reference model, applications and research issues, Robotics and Computer-Integrated Manufacturing 61 (2020) 101837. doi: 10.1016/j.rcim.2019.101837.

[7] A. Kovtunenko, A. Bilyalov, S. Valeev, Distributed streaming data processing in IoT systems using multi-agent software architecture, Lect. Notes Comput. Sc. 11118 (2018) 572–583. doi: 10.1007/978-3-030-01168-0_51.

[8] M. Marrone, L.M. Kolbe, Impact of IT service management frameworks on the IT organization. An empirical study on benefits, challenges and processes, Business & Information Systems Engineering 3(1) (2011) 5–18. doi: https://doi.org/10.1007/s11576-010-0257-8.

[9] S. Ivanov, K. Nikolskaya, G. Radchenko, L. Sokolinsky, M. Zymbler, Digital Twin of City: Concept Overview, in: Proceedings of the Global Smart Industry Conference (GloSIC), IEEE, Chelyabinsk, 2020, pp. 178–186. doi: 10.1109/GloSIC50886.2020.9267879.

[10] W. Kritzinger, M. Karner, G. Traar, J. Henjes, W. Sihn, Digital Twin in manufacturing: A categorical literature review and classification, IFAC-PapersOnLine 51(11) (2018) 1016–1022. doi: https://doi.org/10.1016/j.ifacol.2018.08.474.

[11] R. Kostromin, A. Feoktistov, Agent-Based DevOps of Software and Hardware Resources for Digital Twins of Infrastructural Objects, in: Proceedings of the 4th International Conference on Future Networks and Distributed Systems (ICFNDS 2020), ACM, Saint Petersburg, 2020, pp. 1–6.

[12] A. Feoktistov, I. Sidorov, A. Tchernykh, A. Edelev, V. Zorkalzev, S. Gorsky, R. Kostromin R, I. Bychkov, A. Avetisyan, Multi-Agent Approach for Dynamic Elasticity of Virtual Machines Provisioning in Heterogeneous Distributed Computing Environment, in: Proceedings of the International Conference on High Performance Computing and Simulation, IEEE Press, Orleans, 2018, pp. 909–916. doi: 10.1109/HPCS.2018.00145.

[13] R. Kostromin, Survey of software configuration management tools of nodes in heterogeneous distributed computing environment, in: Proceedings of the the 2nd Intern. Workshop on Information, Computation, and Control Systems for Distributed Environments (ICCS-DE 2020), CEUR-WS Proceedings, Aachen, 2020, 2638, pp. 156–165. doi: 10.47350/ICCS-DE.2020.15.

[14] B. Baeten, F. Rogiers, L. Helsen, Reduction of heat pump induced peak electricity use and required generation capacity through thermal energy storage and demand response, Applied Energy 195 (2017) 184–195. doi: 10.1016/j.apenergy.2017.03.055.

[15] A. Feoktistov, S. Gorsky, I. Sidorov, A. Tchernykh, Continuous Integration in Distributed Applied Software Packages, in: Proceedings of the 42th Int. Convention on information and communication technology, electronics and microelectronics, IEEE, Riejka, 2019, pp. 1775–1780. doi: 10.23919/MIPRO.2019.8757002.

[16] I.A. Sidorov, R.O. Kostromin, A.G. Feoktistov, System for monitoring parameters of functioning infrastructure objects and their external environment, in: Proceedings of the the 2nd Intern. Workshop on Information, Computation, and Control Systems for Distributed Environments (ICCS-DE 2020), CEUR-WS Proceedings, Aachen, 2020, 2638, pp. 252–264. doi: 10.47350/iccs-de.2020.23.

[17] L.A. Sholomov, Logical research methods of discrete selection models, Nauka, Moscow, 1989, 288 p. (in Russian).