Heterogeneous Distributed Computing Environment for Vulnerability Analysis of Energy Critical Infrastructures

Alexei V. Edelev[1], Ivan A. Sidorov[2], Alexander G. Feoktistov[3]

[1] *Melentiev Energy Systems Institute of Siberian Branch of the Russian Academy of Sciences,*
Lermontov str., 130
Irkutsk, Russia
E-mail: flower@isem.sei.irk.ru

[2,3] *Matrosov Institute for System Dynamics and Control Theory of Siberian Branch of the Russian Academy of Sciences,*
Lermontov str., 134
Irkutsk, Russia
[2] *E-mail: ivan.sidorov@icc.ru*
[3] *E-mail: agf@icc.ru*

Abstract

In the paper, we propose a new approach to developing the subject-oriented heterogeneous distributed computing environment for the vulnerability analysis of the energy critical infrastructures of the national level. We provide a framework for creating such environment and developing applied software packages for the vulnerability analysis. Unlike the known tools for developing subject-oriented environments, our framework ensures integrating the models of Grid and cloud computing, as well as selecting the optimal environment resource configuration for performing experiments of various scales.

Keywords: Energy Critical Infrastructures, Modeling, Application, Distributed Computing.

1. Introduction

Nowadays, energy security is one of the global challenges [1]. Winzer [2] defines energy security as a state of protection of all socio-economic spheres of human activity from a deficit in providing its energy needs. Obviously, the studying an energy system vulnerability, which can lead to the energy deficit, is very relevant.

To this end, we consider energy systems as national critical infrastructures that include a large number of various components integrated into a single network. There are different types of threats (technological failures, natural disasters, terrorist acts, etc.) for critical infrastructures. Thus, we should analyze the vulnerability of energy critical infrastructures taking into account a huge number of combinations of their parameters (states, threats, development directions, actions for energy security protection, etc.). A combinatorial nature of the analysis leads to necessary of the high-performance computing use.

Analysis of the known tools of modeling the functioning of energy systems based on the paradigms of parallel and distributed computing shows that they allow us to study only certain aspects of the above-mentioned problems [3]. In this regards, the problem arises of developing and applying a subject-oriented distributed computing environment for the effective analysis of the vulnerability of critical infrastructures in the energy sector.

In the paper, we propose a new approach to the analysis of the vulnerability of the national level energy critical infrastructures. Within this approach, the principles of operation and architecture of the subject-oriented distributed computing environment were developed.

Unlike the above-listed specialized environments for supporting object-oriented research, it provides the development of distributed application software packages, integration of models for Grid and cloud computing, and selection of the optimal resource configuration for performing experiments of various scales.

The environment is created using the specialized framework. It includes tools, which are the base software of the public access computer center "Irkutsk Supercomputer Center of the SB RAS" for parallel and distributed programming [4, 5].

The rest of this paper is structured as follows: Section 2 briefly reviews projects related to developing subject-oriented computing environment in various fields of research. Section 3 addresses the vulnerability analysis of the national energy critical infrastructures. The proposed approach to developing the subject-oriented heterogeneous distributed computing environment for solving the aforementioned problem is represented in Section 4. Section 5 concludes the paper.

2. Related Work

Nowadays, projects related to developing and applying specialized environments for distributed computing in various subject domains are highly relevant. Widely known examples of such environments are the follows ones:

- Distributed Grid-infrastructure for supporting experiments in the field of high-energy physics at the accelerator of the Large Hadron Collider [6],
- Open web-based platform for genomic research [7],
- Cloud platform for rapid analysis of biomedical data [8],
- High-level toolkit for the development of distributed scientific applications [9],
- E-Science infrastructure for data-driven computing [10],
- Service-oriented distributed computing environment for simulation modeling warehouse logistics [11],
- System for modeling critical infrastructures of the electric power industry in a distributed computing environment [3].

Usually, the problem-solving scheme in such environments is closely correlated with the workflow concept that is used in describing the interrelated sub-problems. Workflow stages are implemented by the program modules. Known systems for developing specialized subject-oriented computing environments supporting workflow are Askalon, Karajan, Kepler, Pegasus, Taverna, and Triana [12]. Unfortunately, the possibilities of the above-listed systems usually restrict potential opportunities of the distributed applications within the certain computational model.

In this regard, the research field related to the integration of various computational models (e.g., models of Grid and cloud computing, as well as dedicated and non-dedicated resources) becomes very topical [13].

3. Vulnerability Analysis of National Energy Critical Infrastructures

The energy sector consists of critical national infrastructures which are the most vital for ensuring health, safety, security and well-being of the society, effective functioning of government and economy of a country. The resilience is a structural property of critical infrastructure as the ability to resist to internal drifts and cascading failures, and recover back to the initial operation state. A lot of attention should be paid to the protection and resilience of critical national infrastructures with regard to different kinds of disturbances, such as: natural disasters, man-made catastrophes, technical failures, international crime and terrorism [14]. According Yusta et al. [15] the definition of energy security nowadays evolves into a broader concept that includes cost-effective strategies on the infrastructure protection.

While modeling and simulation tools have provided insights within an individual critical infrastructures, the field of critical infrastructures integration/interdependencies [16] is still relatively immature and has not been studied comprehensively [17]. Also this field is better investigated on the urban, rural, or regional scales rather than on the national level [18].

According Zio [19] the vulnerability analysis provides the quantification of the critical national infrastructures resilience, the identification of their critical elements, and allows the evaluation and comparison of the effectiveness of different protection and recovery strategies. For the vulnerability analysis of critical infrastructures, it is necessary to model them to know their behavior. Taking into account the structural and dynamic complexity of critical national infrastructures, the characterization of the hazardous events and the evaluation of their consequences and probabilities there is not one single modeling approach that captures all aspects of critical infrastructures behavior. A vulnerability analysis framework is needed to integrate a number of existing methods and new analysis approaches capable of viewing the complexity problem from different aspects (topological and functional, static and dynamic), under the large uncertainties exist in the description of the failure behavior of the elements of a complex system, of their interconnections and interactions.

For example, Zio [20] analyzes the power transmission network from four perspectives in order to identify its most important elements.

After comparing the reliability analysis and vulnerability analysis for power systems, Zio and Golea [21] have concluded that the vulnerability analysis should complement various types of the probabilistic risk analysis.

Johansson et al. [22] consider the natural gas transmission network from the following aspects: energy supply security, topology, reliability and controllability.

In [23], the multi-product flow problem and the model for investigating the economic interrelationships of various sectors are combined to quantify the impact of disturbances in the transport sector on the industrial sectors and to assess the importance of network elements.

There are topological, reliability, controllability, sustainability and energy security apects for critical national infrastructures of the energy sector. For each perspective, an indicator is computed to quantify the importance of each critical national infrastructures component. It enables quantitative evaluations of the critical national infrastructures protection and resilience and the identification of the role of each component with respect to different perspectives considered. A comparative evaluation of the perspectives is performed to complete the vulnerability analysis of critical national infrastructures of the energy sector by investigating the consistency and inconsistency between these perspectives.

On the basis of the foregoing, it can be concluded that work on the field of vulnerability analysis methods for critical national infrastructures of the energy sector with stress on their (inter)dependencies and integration is at an early stage. The field of critical national infrastructures (inter)dependencies and integration has not been studied comprehensively despite the substantial scientific interest. The interconnectivity can facilitate the propagation of failure from one critical infrastructure to another. A number of events from around the world have highlighted the potential for failure propagation with critical national infrastructures, resulting in unforeseen, widespread disruptions.

The integration of different types of analysis and methods of system modeling is put forward for capturing the inherent structural and dynamic complexities of critical national infrastructures of energy sector and eventually evaluating their vulnerability characteristics, so that decisions on protections and resilience actions can be taken with the required confidence.

There is a need for new methods and tools for energy critical national infrastructure (inter)dependencies and integration that use data analytics for visualization to identify interdependencies, critical elements and resulting impacts in such complex systems. The benefits for decision makers are:

- To understand the dynamics and complexity of the system and avoid ineffective responses and poor coordination for rescue, recovery, restoration and, mitigation and coordination of such responses,
- To identify obvious and, most importantly, hidden defects or weakness in critical national infrastructure, to manage and reduce them before they manifest as failures.

The used vulnerability analysis approach [24] provides one unified way of modeling systems to represent critical national infrastructures of energy sector of different types. The built models of critical national infrastructures are merged into one system-of-systems model of energy sector [17]. To incorporate the effects of critical national infrastructure interdependencies the energy sector model is based on a quite new energy hub conception [25]. Although that the energy hub concept has been designed for the urban or rural levels the work [26] shows the applicability of that idea at the regional and national levels.

Together it enables solving new problems of the comprehensive vulnerability analysis of critical national infrastructures in energy sector based on a single system modeling approach taking into account higher order (inter)dependencies between them.

4. Subject-Oriented Environment

The new problems of the vulnerability analysis of national energy critical infrastructures that formulated above lead to an increase in the computational complexity of experiments from two to four orders of magnitude in comparison with the known problem formulations. Depending on the parameter dimensions of the used models, these problems can be solved using different computing systems (personal computer, server, cluster, grid, or cloud) for a relatively acceptable time (hours, days), the duration of which is determined by the characteristics of the used computing resources. An integration of the listed systems into a single environment provides both the flexibility in selecting the necessary configuration of the computational infrastructure and speed in implementing experiments of various scales.

Fig. 1 shows the integration scheme of computational systems into the single problem-oriented environment for solving problems of analyzing the vulnerability of national energy critical infrastructures.

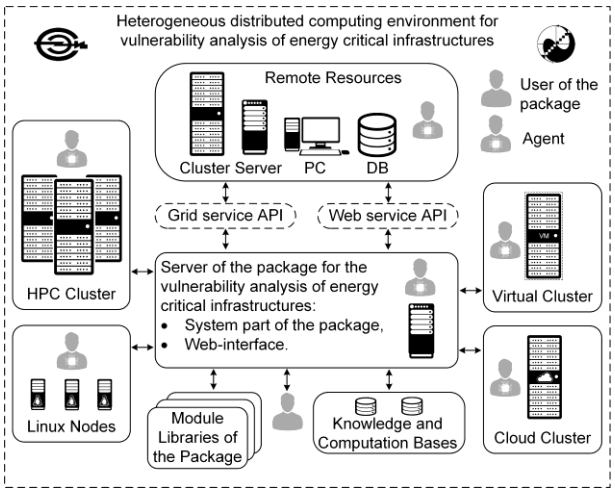

Fig. 1. Integration scheme of computational systems.

The integration is carried out with the help of the above-mentioned tools to develop distributed applied software packages.

The package developer creates it by combining applied software that is designed by him (applied part of the package) and system software of the tools (system part of the package).

The system part of the package is represented by a server that supports the web interface for user access to the components of the package. The basis of the system part is the conceptual model designer, computation schedulers, and problem-solving scheme interpreters.

Module libraries of the package include applied software for solving the problems of analyzing the vulnerability of critical infrastructures in the energy sector.

The knowledge base contains the following knowledge components of the conceptual model:

- Information about modules for solving the subject domain problems, data preprocessing, and data postprocessing;
- Structural properties of algorithms for solving problems in terms of parameters and operations of the subject domain,
- Conditions for selecting the optimal algorithms for solving the problem, depending on the configuration of the environment resources and their states,
- Information about the hardware and software infrastructure (characteristics of nodes, communication channels, network devices, and network topology, as well as data on software and hardware failures),
- Data about distributed computing management systems used in the environment,
- Parameters of administrative policies applied to resources and users.

API for supporting Grid and web-services provides connection of various remote information and computing resources.

The computation database stores the information that is necessary for carrying out the experiments, and the results of solving the problems.

To solve the problem, the user has to formulate the problem on the conceptual model. Based on the problem formulation, the computation scheduler automatically constructs a problem-solving scheme (an abstract program) and creates a job to perform the problem-solving process in the environment. The job includes information about the required resources, executable modules, input and output data, and other necessary information.

Selecting the optimal configuration of the environment for concrete experiments is based on knowledge of both the problem specifics and environment resources. The knowledge about resources includes their computational characteristics and rules for using. All knowledge is reflected in the conceptual model.

We develop and use a multi-agent system to improve the distributed computing management at the environment level. Agents are placed in the control nodes of environment resources. They represent owners of resources and their users. They mitigate conflicts between the preferences of resource owners and the quality criteria of problem-solving processes that are determined by users. To this end, agents use market mechanisms for regulating supply and demand of resources.

We developed the parameter sweep applications to solve complicated practical problems through the combinatorial analysis [4, 5]. These problems related to the energy development with regard to the energy security requirements. The problem solutions ensure to correct the results obtained earlier based on heuristic methods. The implemented experiments showed the advantages of the proposed approach.

5. Conclusions

The paper address the vulnerability analysis of national energy critical infrastructures with high-performance computing in a heterogeneous distributed environment.

Complex technical infrastructures are considered critical if their inoperability or destruction has a significant impact on the society, state, and economy. Critical infrastructures consist of many components integrated into a single network. Energy critical infrastructures are largely exposed to many types of

threats. Thus, the protection of energy infrastructures and their resilience support are topical problems of the national and international scale.

The study of the vulnerability of critical technical infrastructures in the energy sector is based on the analysis of a significant number (tens and hundreds of millions) of disruption scenarios and their calculation requires the use of high-performance computing systems.

Our contribution is multifold. We formulate new large-scale problems of analyzing a vulnerability of critical infrastructures in the energy sector.

We propose a new approach to developing the subject-oriented heterogeneous distributed computing environment for support of such analysis.

We also provide the specialized framework for parallel and distributed programming. In comparison with the known tools for developing subject-oriented environments, it supports the following opportunities in the complex:

- Creating the subject-oriented heterogeneous distributed computing environment,
- Developing applied software packages for the vulnerability analysis,
- Integrating the models of Grid and cloud computing,
- Selecting the optimal environment resource configuration for performing experiments of various scales.

A relevant approach to improving the efficiency of computation management in those systems is the applying of multi-agent technologies. To this end, we develop and use a new subject-oriented multi-agent system that ensures to improve the computation planning and environment resource allocation in the process of solving large-scale problems of energy infrastructure study in practice. The multi-agent system is integrated with the provided framework.

To show the main advantages of the proposed approach in practice, we have solved the problems related to the energy development with regard to the energy security requirements. These problems were solved by applying resources of the public access computer center “Irkutsk supercomputer center of the SB RAS” [27].

Acknowledgements

The study was partially supported by the Basic Research Program of SB RAS, projects № III.17.5.1 and № IV.38.1.1. Part of the work was supported by the Russian Foundation for Basic Research, project № 16-07-00931-a.

References

1. G. Luft and A. Korin (eds.), *Energy Security Challenges for the 21st Century: A Reference Handbook*, (Praeger, 2009).
2. C. Winzer, Conceptualizing energy security, *Energy policy* **46** (2012) 36–48.
3. P. Pederson, D. Dudenhoeffer, S. Hartley, and M. Permann, *Critical Infrastructure Interdependency Modeling: a Survey of US and International Research* (Idaho National Laboratory. 2006).
4. A. V. Edelev and I. A. Sidorov, Combinatorial Modeling Approach to Find Rational Ways of Energy Development with Regard to Energy Security Requirements, *Lecture Notes Comp. Sci.* **10187** (2017) 310–317.
5. A. Edelev, V. Zorkaltsev, S. Gorsky, V. B. Doan, and H. N. Nguyen, The Combinatorial Modelling Approach to Study Sustainable Energy Development of Vietnam, *Commun. Comput. Inf. Sci.* **793** (2017) 207–218.
6. G. Aad et al. The ATLAS Experiment at the CERN Large Hadron Collider, *J. Instrum.* **3** (2008) S08003.
7. J. Goecks, A. Nekrutenko, and J. Taylor, Galaxy: a Comprehensive Approach for Supporting Accessible, Reproducible, and Transparent Computational Research in the Life Sciences, *Genome biol.* **11**(8) (2010) R86.
8. B. Allen, R. Ananthakrishnan, K. Chard, I. Foster, R. Madduri, J. Pruyne, S. Rosen, and S. Tuecke. Globus: A Case Study in Software as a Service for Scientists, in *Proc. 8th Workshop on Scientific Cloud Computing*, eds. K. Chard, A. Costan, B. Nicolae and D. Zhao (ACM, New York, 2017), pp. 25–32.
9. A. Afanasiev, O. Sukhoroslov, and M. Posypkin, A High-Level Toolkit for Development of Distributed Scientific Applications, *Lecture Notes Comp. Sci.* **4671** (2007) 103–110.
10. K. V. Knyazkov, S. V. Kovalchuk, T. N. Tchurov, S. V. Maryin, and A. V. Boukhanovsky, CLAVIRE: e-Science infrastructure for data-driven computing, *J. Comput. Sci.* **3**(6) (2012) 504–510.
11. I. Bychkov, G. Oparin, A. Tchernykh, A. Feoktistov, V. Bogdanova, Yu. Dyadkin, V. Andrukhova, and O. Basharina, Simulation Modeling in Heterogeneous Distributed Computing Environments to Support Decisions Making in Warehouse Logistics, *Procedia Eng.* **201** (2017) 524–533.
12. D. Talia, Workflow Systems for Science: Concepts and Tools, *ISRN Software Engineering* **2013** (2013) 404525.
13. I. Bychkov, A. Feoktistov, I. Sidorov, and R. Kostromin, Job Flow Management for Virtualized Resources of Heterogeneous Distributed Computing Environment, *Procedia Eng.* **201** (2017) 534–542.
14. A. Alessandri and R. Filippini, Evaluation of resilience of interconnected systems based on stability analysis, in *Proceedings of the 7th International Workshop on Critical Information Infrastructures Security*, eds. Hämmerli BM, Svendsen NK, Lopez J (Berlin, Heidelberg: Springer; 2013), pp. 180–90.
15. J. M. Yusta, G. J. Correa and R. Lacal-Arantegui, Methodologies and applications for critical infrastructure protection: State-of-the-art, *Energy Policy* **39**(10) (2011) 6100-6119.

16. S. M. Rinaldi, J. P. Peerenboom and T. K. Kelly, Identifying, understanding, and analyzing critical infrastructure interdependencies. *IEEE Control Systems* **21**(6) (2001) 11-25.
17. S. Thacker, R. Pant and J. W. Hall, System-of-systems formulation and disruption analysis for multi-scale critical national infrastructures. *Reliability Engineering & System Safety* **167** (2017) 30–41.
18. S. Saidi, L. Kattan, P. Jayasinghe, P. Hettiaratchi, and J. Taron, Integrated infrastructure systems—A review, *Sustainable Cities and Society*. **36** (2018) 1-11.
19. E. Zio, Challenges in the vulnerability and risk analysis of critical infrastructures. *Reliability Engineering & System Safety* **152** (2016) 137-150.
20. Zio, E. and Golea, L. R., Analyzing the topological, electrical and reliability characteristics of a power transmission system for identifying its critical elements. *Reliability Engineering & System Safety* **101** (2012) 67–74.
21. J. Johansson, H. Hassel and E. Zio, Reliability and vulnerability analyses of critical infrastructures: comparing two approaches in the context of power systems. Reliability Engineering & System Safety. **120** (2013) 27–38.
22. F. Han, E. Zio, V. Kopustinskas, and P. Praks, Quantifying the importance of elements of a gas transmission network from topological, reliability and controllability perspectives, considering capacity constraints, in *Risk, Reliability and Safety: Innovating Theory and Practice: Proc. of ESREL 2016*, eds. LesleyWalls, Matthew Revie and Tim Bedford, (CRC Press), pp. 2565–2571.
23. M. Darayi, K. Barker and J.R. Santos, Component importance measures for multi-industry vulnerability of a freight transportation network. *Networks and Spatial Economics* **17**(4) (2017) 1111–1136.
24. J. Johansson and H. Hassel, An approach for modelling interdependent infrastructures in the context of vulnerability analysis, *Reliability Engineering & System Safety* **95**(12) (2010) 1335–1344.
25. R. Mohammadi, Y. Noorollahi, B. Mohammadi-Ivatloo and H. Yousefi, Energy hub: From a model to a concept–A review, *Renewable and Sustainable Energy Reviews* **80** (2017) 1512–1527.
26. T. Krause, F. Kienzle, L. Yang and G. Andersoon, Modeling interconnected national energy systems using an energy hub approach, in *Proc. IEEE Power Tech Conf.*, (Trondheim, 2011), pp. 1–7.
27. *Irkutsk Supercomputer Centre of SB RAS*. [Online]. Available: http://hpc.icc.ru.