Model-driven Engineering for Big Data

Belachew Regane

School of Graduate Studies, Software Engineering Track

Addis Ababa University, Ethiopia

belachew.regane@aau.edu.et

ORCID: https://orcid.org/0000-0002-2011-4107

Abstract

Accessing heterogeneous and huge amount of data through different sources heavily impacts users of data nowadays worldwide. Thus, Big Data has now become a hot emerging paradigm in computing environments. Issues in scalability, interoperability, platform independency, adaptability and reusability in big data systems are considered the main current challenges. This raises the need for appropriate software engineering approaches to develop effective and efficient Big Data system models, i.e. an approach which reduce investment cost and development time.

Keywords: Model-driven Engineering, Big Data, Model-driven Approaches, Emerging Paradigm, Software Engineering.

Introduction

Today, software engineering has emerged advanced methodologies to solve problems from different perspectives, while still further research is needed to overcome new challenges raised in emerging technologies, i.e. Big Data (Osvaldo et al., 2017; Venters et al., 2017).

According to Bhadani and Jothimani (2016), Klein et al. (2016) and Kepner (2016), the increment of data in terms of volume, variety, and velocity from time to time indicates a serious need for scalable systems to fit Big Data. Existing software development approaches are not fully appropriate for the development of software applications that handle highly increasing and heterogeneous data as in Big Data systems. There is an urgent need for developing comprehensive engineering principles, methodologies, and tools that support the entire software development lifecycle of MDE applications to overcome challenges in Big Data systems. This pushes software engineers to come with better approaches to handle challenges raised in such emerging complex systems. As discussed in Mussbacher et al. (2014) and Kumar and Alencar (2016), developing and maintaining stable and scalable big data systems are still open issues that need further research to overcome the associated challenges. Thus, the author believes model-driven engineering technique is the appropriate approach to alleviate complexities in modeling Big Data system.

Model-Driven Engineering and Debating issues

In complex systems like Big Data systems, model-driven engineering plays a great role to manage software development by using models, notations and transformation rules for abstraction. Models, metamodels, modeling languages, model transformation and supporting software tools are basic components which an approach focuses on to achieve the development, maintenance and evolution of the system with low cost and high quality (Osvaldo et al., 2017; Godena et al.,2015).

Model-driven engineering (MDE) is an approach to software development where models rather than programs are the principal outputs of the development process (Sommerville, 2011).

In model-driven engineering (MDE), models are categorized into three types, which are computational-independent models, platform-independent models, and platform-specific models. As discussed in Sommerville (2011) and Ciccozzi and Spalazzese (2018), model-based engineering is still at an early stage of development and it is unclear whether it will have a significant effect on software engineering practice. The main arguments for and against MDE are:

1. For MDE, Model-based engineering allows systems at a high level of abstraction. This results in reduced errors speeded up the design and implementation process and allow the creation of reusable, platform-independent application models. System implementations can be generated automatically using different tools for different platforms from the same model. Thus, by writing a translator for platforms, it is easy adapting the system to any new platform technology.
2. Against MDE, even if models are a good way of facilitating discussions about software design, the abstractions supported by models are not always the right abstractions for implementation. Furthermore, it is unclear whether the costs and risks of model-driven approaches outweigh possible benefits.

Why MDE for big data?

Being in the debate, why is the need for model-driven engineering as a software development approach for big data systems? The basic reasons behind the need for a model-driven software engineering approach for Big Data systems have been identified as follows:

a) MDE is a platform-independent method, which indicates that the system operation is performed without reference to its implementation. It is described using UML models that show the static system structure and how it responds to internal and external events.
b) MDE is a programming language-independent method. Different service frameworks and application programming interfaces (APIs) have been developed using different programming languages and run on different platforms, helping developers to implement services. Therefore, a graphical representation of service may have several implementations (i.e., source code). This source code may also use different programming languages and may run on different platforms.
c) MDE is a modeling language-independent method. Thus, it is possible to use different existing modeling languages and different modeling editors.
d) MDE is domain-independent, in which it is applicable for all domains. This can be done by developing different abstractors for different target languages.
e) In software development projects, scope, budget and time are triple challenging constraints to achieve the goal. However, with MDE, the development of software systems can be done automatically, so that it reduces investment cost and development time.

Conclusion

With the opportunities presented for the MDE approach, MDE modeling techniques should be considered in developing a Big Data system to overcome challenges in scalability, interoperability, platform independency, adaptability and reusability of models.

Thus, the author believes that using computational-independent models, it is possible to design and develop a generic and scalable system model for Big Data through model-driven engineering, to overcome challenges raised in scalability, interoperability, platform independency, adaptability and reusability of models.

References

1. Bhadani, A., Jothimani, D. (2016). Big data: Challenges, opportunities, and realities.
2. Colin C. Venters, Rafael Capilla, Stefanie Betz, Birgit Penzenstadler, Tom Crick, Steve Crouch, Elisa Yumi Nakagawa, Christoph Becker, Carlos Carrillo, (2017). Software Sustainability: Research and Practice from a Software Architecture Viewpoint. The Journal of Systems & Software.
3. Federico Ciccozzi and Romina Spalazzese, (2018). MDE4IoT: Supporting the Internet of Things with Model-Driven Engineering, accessed in September 2018, from http://www.es.mdh.se/pdf_publications/4386.pdf.
4. Giovanni Godena, TomažLukman, MarjanHeričko, StankoStrmčnik, (2015).The Experience of Implementing Model-Driven Engineering Tools in the Process Control Domain, Information technology, and control.
5. Ian Sommerville (2011). Software Engineering, (Ninth Edition), Addison-Wesley.
6. Jeremy Kepner, (2016). Enabling Scale-Up, Scale-Out, and Scale-Deep for Big Data, Massachusetts Institute of Technology.
7. John Klein, Ian Gorton, Laila Alhmoud, Joel Gao, Caglayan, (2016). Model-Driven Observability for Big Data Storage, IEEE.
8. John Klein, Ross Buglak, David Blockow, Troy Wuttke, Brenton Cooper, (2016). A Reference Architecture for Big Data Systems in the National Security Domain, 2nd International Workshop on BIG Data Software Engineering.
9. Kumar, V. D., &Alencar, P. (2016). Software engineering for big data projects: Domains, methodologies, and gaps. In Big Data (Big Data), 2016 IEEE International Conference on (pp. 2886-2895).
10. Mussbacher, Gunter and Amyot, Daniel and Breu, Ruth and Bruel, Jean-Michel and Cheng, Betty and Collet, Philippe and Combemale, Benoit and France, Robert B. and Heldal, Rogardt and Hill, James and Kienzle, Jörg and Schöttle, Matthias and Steimann, Friedrich and Stikkolorum, Dave and Whittle, Jon, (2014). The relevance of model-driven engineering thirty years from now, In 17th International Conference ACM/IEEE - Conference on Model Driven Engineering Languages and Systems (MODELS), Valencia, Spain.
11. S. Sousa Osvaldo Jr., Denivaldo Lopes, Aristófanes C. Silva, ZairAbdelouahab, (2017). Developing software systems to Big Data platform based on the MapReduce model: An approach based on Model-driven Engineering, Information and Software Technology 92 (2017) 30–48.