Computers in Industry 65 (2014) 521–535

Contents lists available at ScienceDirect

Computers in Industry

journal homepage: www.elsevier.com/locate/compind

Collaborative simulation and scientific big data analysis: Illustration for sustainability in natural hazards management and chemical process engineering

Jean-Pierre Belaud [a,b,*], Stéphane Negny [a,b], Fabrice Dupros [c], David Michéa [c], Benoît Vautrin [d]

[a] Université de Toulouse, INP-ENSIACET, 4, allée Emile Monso, F-31432 Toulouse Cedex 04, France
[b] CNRS, LGC (Laboratoire de Génie Chimique), F-31432 Toulouse Cedex 04, France
[c] BRGM, 3 Avenue Claude Guillemin, 45060 Orléans, France
[d] OXALYA – OVH.COM Group, 6 bis, rue Riquet, 75019 Paris, France

ARTICLE INFO

Article history:
Received 21 December 2012
Received in revised form 19 December 2013
Accepted 13 January 2014
Available online 7 February 2014

Keywords:
ICT platform
Computer-aided design
3D visualization
Big data
Natural hazards management
Chemical process engineering

ABSTRACT

Classical approaches for remote visualization and collaboration used in Computer-Aided Design and Engineering (CAD/E) applications are no longer appropriate due to the increasing amount of data generated, especially using standard networks. We introduce a lightweight and computing platform for scientific simulation, collaboration in engineering, 3D visualization and big data management. This ICT based platform provides scientists an "easy-to-integrate" generic tool, thus enabling worldwide collaboration and remote processing for any kind of data. The service-oriented architecture is based on the cloud computing paradigm and relies on standard internet technologies to be efficient on a large panel of networks and clients. In this paper, we discuss the need of innovations in (i) pre and post processing visualization services, (ii) 3D large scientific data set scalable compression and transmission methods, (iii) collaborative virtual environments, and (iv) collaboration in multi-domains of CAD/E. We propose our open platform for collaborative simulation and scientific big data analysis. This platform is now available as an open project with all core components licensed under LGPL V2.1. We provide two examples of usage of the platform in CAD/E for sustainability engineering from one academic application and one industrial case study. Firstly, we consider chemical process engineering showing the development of a domain specific service. With the rise of global warming issues and with growing importance granted to sustainable development, chemical process engineering has turned to think more and more environmentally. Indeed, the chemical engineer has now taken into account not only the engineering and economic criteria of the process, but also its environmental and social performances. Secondly, an example of natural hazards management illustrates the efficiency of our approach for remote collaboration that involves big data exchange and analysis between distant locations. Finally we underline the platform benefits and we open our platform through next activities in innovation techniques and inventive design.

1. Introduction and scientific context

Sustainability is a paradigm for thinking about the future in which environmental, societal and economic considerations are equitable in the pursuit of an improved lifestyle. Most of the economies are developing with breakneck velocities and are becoming epicenters of unsustainable global growth. Immense utilization of natural resources, waste generation and ecological

Abbreviations: API, Application Programming Interface; CAD, Computer-Aided Design; CAD/E, Computer-Aided Design/Engineering; CSV, Comma-Separated Values; GIS, Geographical Information System; GUI, Graphical User Interface; GMT, Generic Mapping Tools; HTTP, HyperText Transfer Protocol; HTTPS, Hyper Text Transfer Protocol Secure; HPC, High Performance Computing; ICT, Information and Communication Technologies; IT, Information Technologies; LAN, Local Area Network; MS, Microsoft; SaaS, Software As A Service; SME, Small Medium Enterprise; SOA, Services Oriented Architecture; SVG, Scalable Vector Graphics; TCP, Transmission Control Protocol; VTK, Visualization ToolKit; X3D, Extensible 3D; XML, eXtreme Markup Language; WAN, Wide Area Network; WYSIWIS, What You See Is What I See.

* Corresponding author at: Université de Toulouse, INP-ENSIACET, 4, allée Emile Monso, F-31432 Toulouse Cedex 04, France. Tel.: +33 05 34 32 36 48.
E-mail addresses: jeanpierre.belaud@ensiacet.fr (J.-P. Belaud), stephane.negny@ensiacet.fr (S. Negny), f.dupros@brgm.fr (F. Dupros), d.michea@brgm.fr (D. Michéa), benoit.vautrin@oxalya.com (B. Vautrin).

0166-3615/$ – see front matter
http://dx.doi.org/10.1016/j.compind.2014.01.009

irresponsibility are the reasons for such a dire situation. With the world in majority debating over issues like climate change, water resources, food security, energy efficiency for the last few decades, it is evident that sustainability and green thinking have taken root in all approaches and dialogs. Governments are rethinking their developmental paths adapted to ensure a sustainable lifestyle. Industry, academic institutions, public sectors are taking serious advancement to implement the same. A brief highlight of sustainability takes into consideration three pillars i.e. economic, social and environment. Engineering domains have to develop innovative solutions according to this new paradigm. One first industry to come under scrutiny was the chemical processes and heavy industry sector; however this has tended to evolve to cover other sectors and different sizes of industry. Efforts in manufacturing and chemical industries have been moving from ''end of pipe'' technological solutions to limit or control pollution, to the integration of the environmental preoccupation in early stage of product or process (preliminary) design at an industrial park level. Pollution control, eco-efficiency, life cycle thinking and industrial ecology are the main steps of sustainable manufacturing practices [1].

At present scientific area of chemical process engineering and natural hazards management is recognized as a method to integrate an efficient sustainability analysis and strategy. Those two engineering domains provide handful solution to manage systems by enabling the use of modeling, simulation, optimization, planning and control in order to develop a more sustainable product and process. In this context scientific simulation based on big data and collaborative work has to be developed for succeeding Computer-Aided Design/Engineering (CAD/E) of sustainable system.

In scientific simulation based High Performance Computing (HPC) area, pre and post-processing technologies are the keys to make the investments valuable. Besides, the data size and data model increase make it mandatory for industrial and academic users to have access to sufficient power on a remote and collaborative way. Our aim is to develop an open technological web platform that provides HPC, collaboration and 3D visualization capabilities to end users and software developers for product design by simulation. A lot of research works, systems and toolkits have been proposed for distributed and remote scientific visualization of large data sets over scientific networks. A good overview of different solutions for distributed and collaborative visualization can be found in Brodlie et al. [2] and in Grimstead et al. [3]. Nevertheless, to our knowledge, a multi-domain collaborative platform for decision-making in simulation for complex systems, orchestrating transparently a set of advanced pre and post processing scientific visualization services is a real innovation. That is a fact that the collaborative part of the different existing systems is often reduced to basic tools such as ''shared-display''. Current collaborative techniques have no advanced communication of visual objects and advanced person–machine interface dedicated to remote collaboration. As a consequence, the design of our platform will require innovations into: (i) pre and post processing semantic visualization services in distributed and parallel environments, (ii) 3D large scientific data set scalable compression and transmission methods, (iii) advanced collaborative virtual environments for 3D data, and (iv) Computer-Aided Design/Engineering (CAD/E) usage for multi-domains of engineering.

We discuss hereafter the three first items. Section 2 deals with last item for the needs for collaboration in CAD/E systems and illustrates some current industrial needs from BRGM experience. The Section 3 presents the platform without giving any deep technical information. This platform is now available as an open project with all core components licensed under LGPL V2.1 and advanced technical information can be found from the platform documentation.[1] Before drawing conclusion, the Section 4 provides two examples of usage of the platform in CAD/E for sustainability engineering from one academic application and one industrial case study.

1.1. Pre and post processing visualization services

The aim is to provide engineers and researchers with tools to operate on their meshes remotely. Mesh generation, optimization and adaptation are a topic highly studied in the literature [4]. The current solutions that actually enable such an analysis have two major drawbacks: they do not offer the possibility to distribute transparently the processing; they are mainly local solutions. The state of the art on scientific visualization environment has deeply evolved with the design of user-friendly solutions such as AVS, IBM/Data Explorer, Avizo, Covise, Ensight, VTK, Cassandra and many others. Recently, distributed and parallel visualization solutions such as the open source platform Paraview and the commercial package Ensight Gold/DR appeared. Nevertheless, there are still efforts to do for the deployment of these solutions on a large computing grid or HPC center.

1.2. Remote scientific 3D visualization

Due to the large volume of data handled, a lot of compression and progressive transmission methods have been proposed in the past to deliver in real time 3D content. These different methods can be classified into four main approaches:

- Image based streaming: the 3D data is stored on the server and only the 2D rendered images are streamed in real time to the server. It is the approach chosen by many solutions because it can be easily implemented and it ensures the best use of the network bandwidth.
- Object streaming: a 3D object is compressed and is progressively transmitted over the network. Specific file formats such as X3D and MPEG4-BIFS are generic formats for 3D object streaming.
- Scene streaming: the data delivery is extended to the entire scene. This approach is widely used for famous Massively Multiplayer Online Games such as Active Worlds or Second Life.
- Scientific visualization streaming: the large volume of the scientific data, their time dependent deformation and the accuracy of model representation are important features that imply specific streaming methods.

In addition to this 3D visualization requirement, two optional needs for our platform are considered: (i) major 3D compression techniques; both mesh geometry and mesh connectivity compression techniques; to transmit 3D data over the internet and (ii) digital watermarking as a potential efficient solution for copyright protection. Those two capabilities are not developed in this article which is focused on the usage of the platform in engineering.

1.3. Collaborative environments and techniques

The DIS/HLA IEEE standard, on which most of military tactical simulations are based, illustrates how a distributed simulation manages a set of several entities which interact and communicate in ''real time'': when an action is executed, the related information or outputs are dispatched/broadcasted as ''quickest'' as possible in the network. The more commonly used algorithms [5] deals with the concept of ''referentials'' and ''proxys'', as we can find in

[1] http://forge.collaviz.org/documentation.