

IT-Infrastructure for an Integrated Visual Analysis of Distributed Heterogeneous Simulations

Tobias Meisen^{1, a}, Rudolf Reinhard^{1, b}, Thomas Beer^{2, c}, Daniel Schilberg^{1, d}, Sabina Jeschke^{1, e}

¹RWTH Aachen University, Institute of Information Management in Mechanical Engineering, Dennewartstraße 27, 52068 Aachen, Germany

²RWTH Aachen University, Institute for Scientific Computing, Seffenter Weg 23, 52074 Aachen, Germany

^atobias.meisen@ima.rwth-aachen.de, ^brudolf.reinhard@ima.rwth-aachen.de, ^cbeer@vr.rwth-aachen.de, ^ddaniel.schilberg@ima.rwth-aachen.de, ^esabina.jeschke@ima-zlw-ifu.rwth-aachen.de

Keywords: IT-Infrastructure, simulation interconnection, application integration, information integration, visualization.

Abstract. Computational simulations are used for the optimization of production processes in order to significantly reduce the need for costly experimental optimization approaches. Yet individual simulations can rarely describe more than a single production step. Hence, a set of simulations has to be used to simulate a contiguous representation of a complete production process. Besides, simulated results have to be analyzed by domain experts to gather insight from the performed computations. In this paper, an IT-infrastructure is proposed that aims at a rather non-intrusive way of interconnecting simulations and domain expert's knowledge to facilitate the collaborative setup, execution and analysis of distributed simulation chains.

The final version of this work has been published in *Advanced Materials Research*, ISSN: 1662-8985, Vols. 328-330, pp 1940-1946, doi:10.4028/www.scientific.net/AMR.328-330.1940, © 2011 Trans Tech Publications Ltd, Switzerland

Introduction

Optimizations for an increasing efficiency in production are often achieved by making use of simulations. Because of high costs and time needs, it is hardly possible to realize such simulations by experiments of the complete production process. Therefore, the usual approach consists in the use of computational simulations, which minimizes the experimental effort [3][42]. Unfortunately, these simulations mostly regard single aspects of the process and do not consider simulation results from previous steps in the simulated process. Interactions of effects on different scales are oversimplified or conditionally included in existing simulation approaches. Consequently, key factors, such as the material's history through the process, are neglected. To overcome these weaknesses of typical simulation approaches, an IT-infrastructure is presented including components that allow the interconnection of several simulation tools as well as an integrated analysis of the simulation results. This infrastructure has been successfully adopted to implement the AixViPMaP® (Aachen (Aix) Virtual Platform for Materials Processing) [33].

The paper is structured as follows: In the next section, the current state of the art will be outlined in order to provide a foundation for the following section, in which one of the implemented use cases is presented in detail. This brief introduction is followed by a section in which the components of the simulation platform are described. In the last section, a conclusion and outlook will be drawn from the insights generated in this paper.

State of the art

In the field of simulations, the process development time using manually linked simulations incurs an overhead of 60%, caused solely by the manual operation of the individual simulation components and conversion tools of a simulation chain [25].

At the same time, integration problems belong to the most frequented topics with reference to finding answers to questions which are raised across application boundaries [17][41]. The complexity of such integration problems arises by reason of the many topics that have to be regarded to provide a solution. Besides application interconnection on the technical level, the data has to be propagated and consolidated. Furthermore, user interfaces are required to model the underlying process and a unified visualization of the data for the purpose of analysis. The necessary visualization depends on the requirements of the user and therefore has to consider the user's background. In addition, the integration of data requires the understanding and thus the comprehension of domain experts. Because of those reasons, integration solutions are often specialized and highly adapted to the specific field of application. One example for such a solution is the Cyber-Infrastructure for Integrated Computational Material Engineering (ICME) [18] concerning the interconnection of MATLAB applications. Other examples are solutions that require making adjustments on the source level of the application, like CHEOPS [34] or the FlowVR toolkit [1]. Yet others require the implementation of standards like SimVis [21]. Realizing a flexible solution, the technical, the data and the analysis level have to be taken into account.

The interconnection on the technical level has been researched intensively. Several solutions have been presented during the last years [10][19][30]. In particular the use of middleware technologies has been established to provide a solution to such kind of problems. Middleware solutions used in the field of simulation interconnecting often require linking of hardware resources, in addition to the associating of the simulations. This issue is addressed in the field of grid computing. Popular grid middleware agents include Globus (www.globus.org) [11], g-lite (glite.cern.ch), UNICORE (www.unicore.eu) [37] and Condor (www.cs.wisc.edu/condor) [39].

Concerning the data and information level, a conversion of the data syntax is not sufficient. Instead, the structure and the semantics of data have to be considered in the conversion process [12][13][35]. For such processes, the usage of schema- and ontology-based integration methods [9][31][38] has been established as a solution. Thereby, research mainly focuses on data schemas based on the relational or XML data model. In this respect, knowledge bases containing background knowledge about the application domain are often used to facilitate the semantic integration of data [12][14]. There are various research projects in this field which have produced different solutions for schema- and ontology-based integration, like COMA++ [2][8] and FOAM [7]. Both systems feature algorithms analyzing the structure of the schema and do not regard the stored dataset. Hence, these systems are unable to identify different dataset semantics within one schema.

For the analysis of simulation results, first of all, an adequate visualization is required. Current simulation tools often comprise in-house solutions tailored to specific simulations. However, these are only suitable for visualizing the datasets generated by this particular simulation software. Hence, generic visualization solutions, such as ParaView (www.paraview.org) [23] or OpenDX (www.opendx.org), are better suited to this purpose, mainly because of their wide support of meshes generated by finite-element-simulations. However, these applications are designed to extract and display visual features, such as isosurfaces, from individual datasets. Furthermore, they lack the means of interaction required to deal with all the simulation results of a simulated process chain in a common context.

Use Case

Within this paper, the production of a line-pipe will be stressed as an example use case. During the production process, several simulation tools are used that are specialized for the simulation of the needed production steps. The goal consists in the simulation of the whole process, whereby the

results of each specialized tool will be considered across the whole simulated process. The production process which will be used to exemplify the use case is illustrated in Fig. 1.

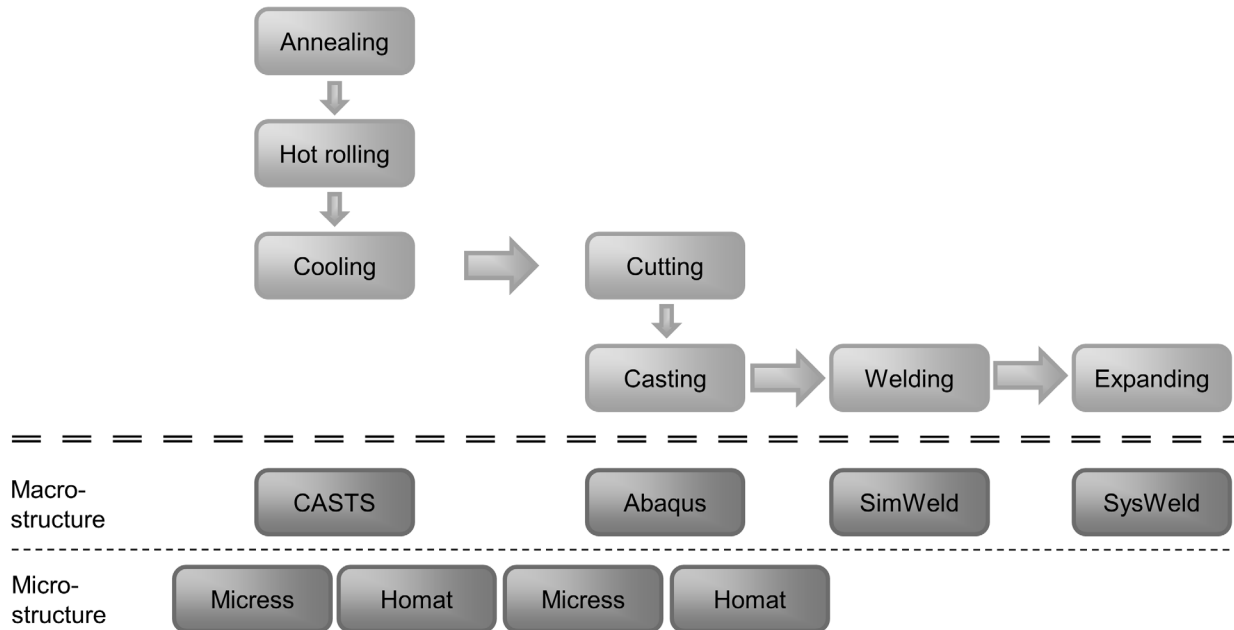


Fig. 1: Production process of a line-pipe and involved simulation tools

The use case starts with a simulation of the annealing, the hot rolling as well as the controlled cooling of the components via CASTS, an application developed by Access e. V.

The next step consists in representing the cutting and the casting with the aid of Abaqus (Dassault Systems), whereas the welding and the expanding of the line-pipe will be simulated via SimWeld, a tool which was developed by the Welding and Joining Institute (ISF) of RWTH Aachen University, and via SysWeld, a software product contrived by the ESI-Group [32]. Furthermore, the simulation of modifications in the microstructure of the assembly will be realized by making use of Micress [26] and Homat [27], which were both developed by Access e. V. All in all, the use case contains six different kinds of simulations, each based on different data formats and models. Thereby, the required integration solution has to take different requirements into account [29]. Two requirements which turned out to be central with reference to the presented IT-infrastructure are, on the one side, the possibility of data propagation focusing the semantic data exchange between the applications and, on the other side, the necessity of a process-oriented data consolidation. Both of them are used to facilitate the subsequent visualization and analysis of the collected data. Another important requirement which has to be fulfilled is the easy adaption of the whole simulation (i.e. the simulation of the production process) concerning changes of the process (e.g. adding or replacing simulation tools or extending the process by further manufacturing methods). In addition to this use case, the provided solution has been used to model four additional use cases that are consolidated in the domain of material processing.

Distributed Collaborative Simulation Platform

The overall architecture is based on a GRID approach. It utilizes the middleware Condor for centralized resource management and distribution of simulation tasks and data, thus using this middleware as a single point of contact for the platform users and to hide the complexity of the distributed system (more details about the underlying GRID-based system can be found in [5]). Beside the simulation resources, an Information Integrator serves as a special resource in its role as translator between simulation resources. An Interactive Workflow Manager has been developed to define the simulation process and to select the involved simulation tools. Finally, the integrated analysis is realized by a visualization component.

Interactive Workflow Manager. The workflow manager is a web 2.0 application that enables the user to visually design simulation workflows. These workflows are automatically translated into description files required by the Condor middleware and can then be executed from within the web interface. The resulting jobs are performed on the different distributed resources. Because the entire process is managed by a server-side component, there is no need to install any software on client machines in order to design and execute simulation chains. Fig. 2 depicts the editor view of the Workflow Manager.

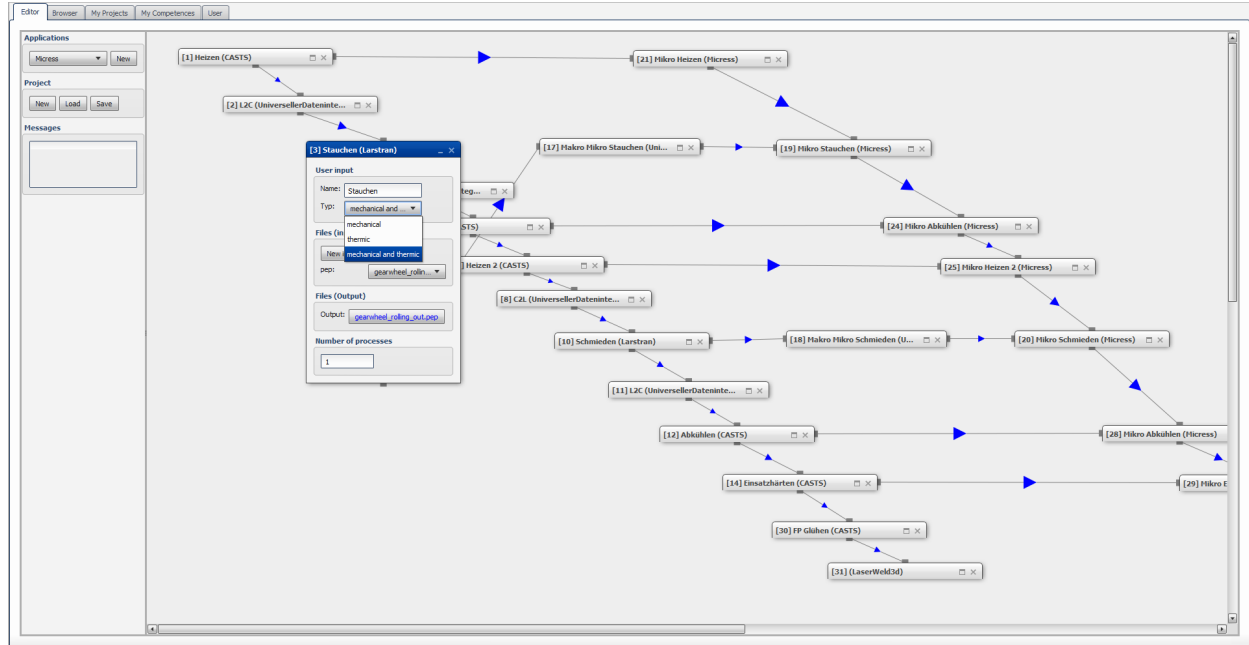


Fig. 2: Simulation workflow design in the editor view of the Workflow Manager

Workflows are built by adding simulation steps represented by windows that hold configuration details for the appropriate simulation tool. These form the nodes of the underlying process graph. Dependencies between simulation steps are modeled by graph edges that define data input and output. For tools that require a translation step in-between, the Information Integrator service, described below in detail, is placed between two simulation resources. To provide the feature of collaboration, the application server supports a rights management system as well as project locking. With these tools at hand, a project coordinator can create new simulation projects, whereby a domain expert can browse available and associated simulation tools to build up a team to further refine complex simulation chains. Finally, instances of created simulation workflows can be executed and monitored, whereas results from the finished instances can be accessed through the web interface by all team members of a simulation project.

Adaptive Information Integration. As already described previously, an integration solution is required to interconnect the different simulations on the syntactical, structural and semantic level. In the presented approach, the translation is realized by the Information Integrator service. The Integrator itself makes use of services to provide the needed functionality. Hence, the architecture of the Information Integrator is based as well upon a service-oriented approach. Thereby, a service bus [6] realizes the communication between the different services. The main purpose of these services is to overcome the heterogeneity of the data. In literature, different kinds of heterogeneity are distinguished [15][22][28]. In this paper, the well-established kinds of heterogeneity listed in [28] are used. Thus the syntactical and the structural heterogeneity are overcome by making use of adapter services. In contrast to a common adapter-based approach, the adapter services do not need any domain-specific knowledge to fulfill their task. Their main purpose is to transfer the data stored in a file or database into a Central Data Storage (CDS) without considering the structure or the semantics of the data. The main concept is exemplarily sketched in Fig. 3.

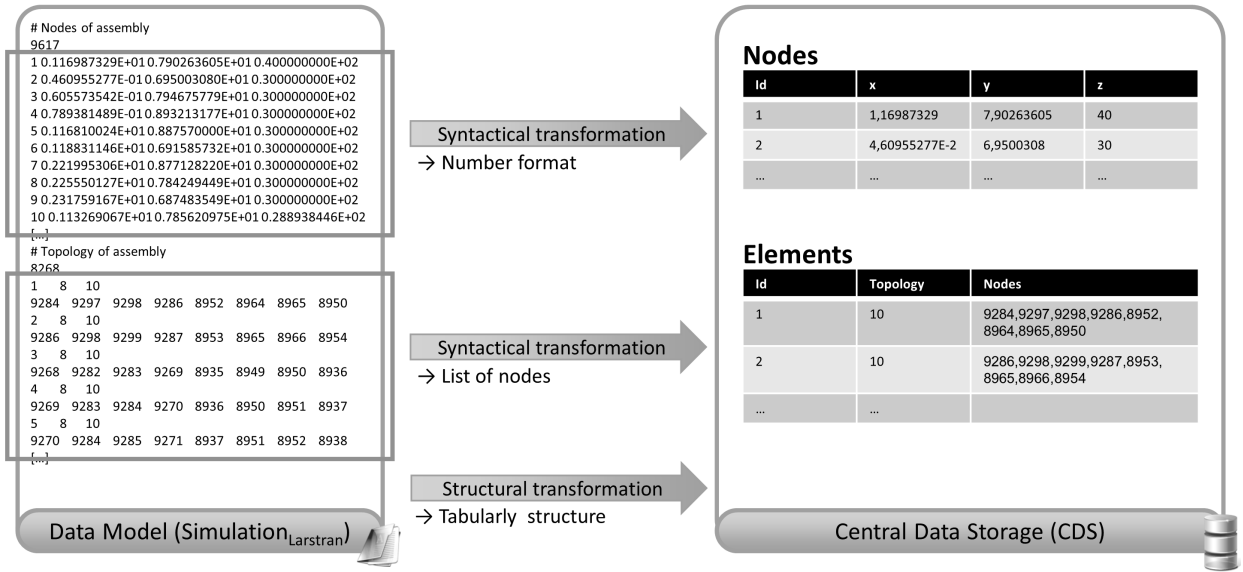


Fig. 3: Integration of data into the Central Data Storage (CDS)

Because of the huge data volumes that are examined in examined use cases, a relational database is used as CDS. For example, the data generated executing the presented use case exceeds more than ten gigabytes. Using a relational database, a very well scaling and the usage of index algorithms, like the R-Tree [16] for three dimensional data or the X-Tree [4] for multi-dimensional data, become possible. The Information Integrator distinguishes two kinds of adapter services: Integration services are used to transfer data into the CDS, whereas extraction services are employed to transfer data from the CDS into a file or database.

After the data has been integrated into the CDS, it is further processed with analysis, planning and transformation services. Analysis services identify features that are fulfilled by the integrated process data (as-is state). In this manner, a feature is a domain-specific property or characteristic of data. Furthermore, the service analyzes which features have to be fulfilled so that the data can be extracted into a defined data format (target state). This information is forwarded to a planning service to determine a translation process that transforms the given set of features into the needed one. Finally, transformation services process the identified translation. The features as well as the transformations and their preconditions and effects, are described in a knowledge base. This knowledge base is expressed in the Web Ontology Language (OWL) [36] and considers domain-specific knowledge. Fig. 4 shows an example of the determination of a translation process by using analysis and planning services.

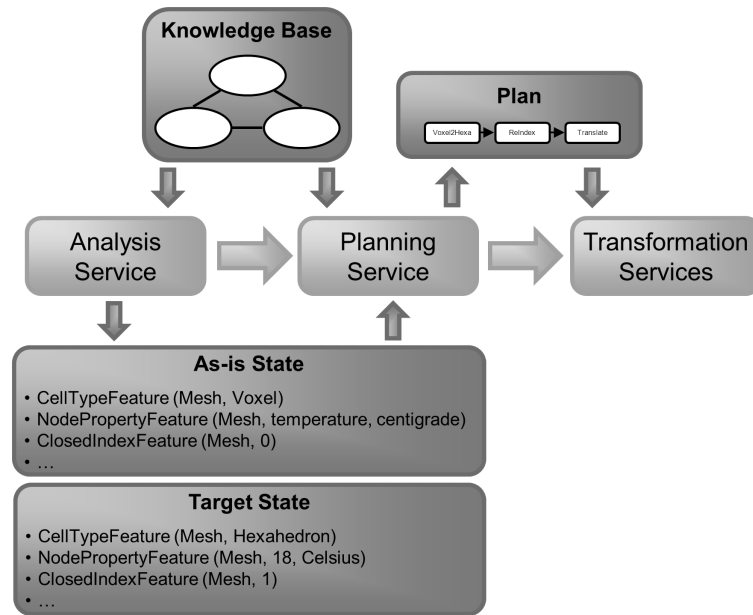


Fig. 4: Determination of a transformation process using services

In the depicted example, the initial mesh consists of voxel elements and has a node property named temperature, which is measured in centigrade. In addition, the indexing of the nodes and elements is closed, which means there are no gaps in the numbering of the index, starting by zero. For extracting the data into the desired target format, the mesh has to fulfill a set of features, defined in the target state. These features determine that the mesh has to consist of hexahedrons and the temperature has to be presented by a property named 18. The indexing of nodes and elements has to be closed and started by one.

As result, the planning service determines a translation process, described by a plan, which translates the data into the desired target state. Following the example, the translation process contains three steps: VoxelToHexahedron, ReIndex and Translate. Each of these transformation steps is realized as a separated service within the Information Integrator (cf. Fig. 5).

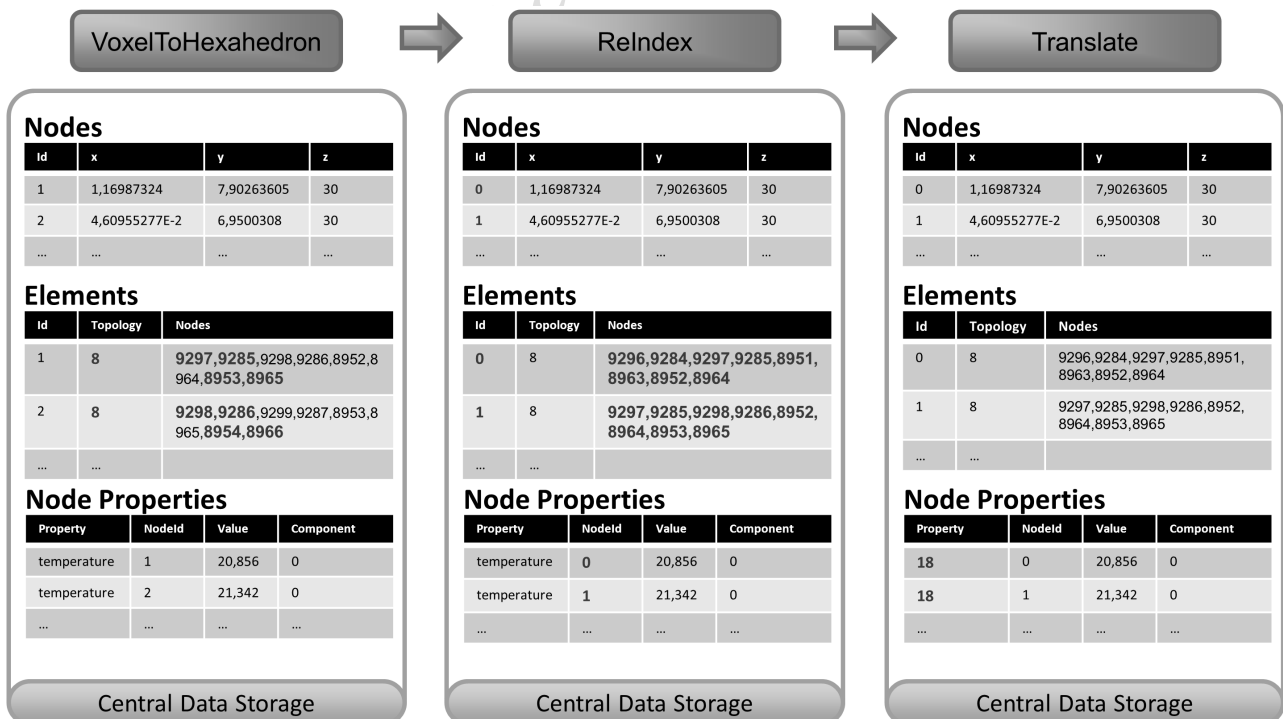


Fig. 5: Transformation process

Because of this new approach, the Information Integrator is robust against changes in the requirements of data formats according to the structure and the semantics. Besides that, new data formats can easily be integrated by defining the integration and extraction service as well as the structural and semantic requirements. Transformation services are automatically included after they have been published in the knowledge base.

Virtual Integrated Simulation Analysis. The ultimate goal of performing simulations is a deeper understanding of the underlying processes and effects. Therefore, efficient means of data analysis are a mandatory component in the overall workflow. Hence, as the simulated phenomena can rarely be detected by generic computational methods, paradigms and tools, scientific visualizations serve to transform the vast amount of numerical results into a readily accessible visual representation. After the simulations have successfully delivered their results and thus formed a contiguous representation of material, visualizations of the different production steps have to be merged into a common context. This context involves not only multiple scales, but also different temporal and spatial resolutions as well as different data fields (e.g. temperatures at macro level, grain identifiers at micro level).

In a pre-processing step, the Information Integrator is used to translate all simulation data into a format suitable for visualizations. Because of the widespread distribution and wide acceptance, the PolyData-Format of the Visualization Toolkit software library (www.vtk.org) [24] has been chosen. Moreover, a unified post-processing is executed. Because a visualization pipeline that filters semantically important information out of the heterogeneous simulation results from different domains is not yet fully automatable, it is realized as a manual preparation step. Within the post-processing, standard visualization objects, like object surfaces, cut-planes or isosurfaces, can be extracted for each dataset interactively. For performance reasons, just the surface meshes of the datasets are exported instead of driving the complete visualization pipeline in real-time during the analysis process. Besides performance issues, it is also a memory problem to keep full unstructured volumetric data for all simulations in-core at once. Besides, a wide range of display systems is targeted and thus only different amounts of data can be handled interactively. Hence it is necessary to use decimation algorithms to enable an interactive visual analysis process.

Decimation algorithms do have a long history and many different approaches are known [20]. The presented solution proposes the use of a pragmatic decimation method based on approximated centroidal Voronoi diagrams (ACVD) [40] that results in reasonably uniform mesh resolutions. This way, data values on flat surfaces that would be rejected by traditional geometric error metrics are preserved, while not generating too much computational load. For most datasets in the considered simulations chains, this approach delivers good quality. For the few other datasets, the automated decimation tool is configured to use a traditional decimation algorithm [9] as provided by VTK. Which algorithm is used can either be set globally for a decimation session or on a per-dataset basis by adding a special tag to the metadata structure used for the visualization.

After the visualization data and the corresponding decimated version have been made available to the visualization front end, the integration into a common visualization context in a concerted manner enables the analysis and the discussion of phenomena along the whole process between domain experts. Thereby, multiple datasets are representing material and work piece states at different spatial scales, e.g. grain level (magnitude of μm) and work piece scale (magnitude of m), spatial locations, temporal scales (ranging from seconds to hours for sub-processes) and temporal locations. In order to avoid a lost-in-data situation and to help the analyst keep track of the spatiotemporal relations the datasets form in the simulated process chain, a hybrid 2D/3D user interface called the dataset sequencer is proposed. It focuses on interaction with the temporal setup of a visualization environment during the analysis process. With this component of the visualization system, a 2D “desktop” and 3D “immersive” means of interaction is offered, which both can be used on a wide range of target systems. While a detailed discussion would be out of the scope for this paper, Fig. 6 gives an impression of how 3D data is arranged and visualized. The figure shows an excerpt of selected results of the different use cases.

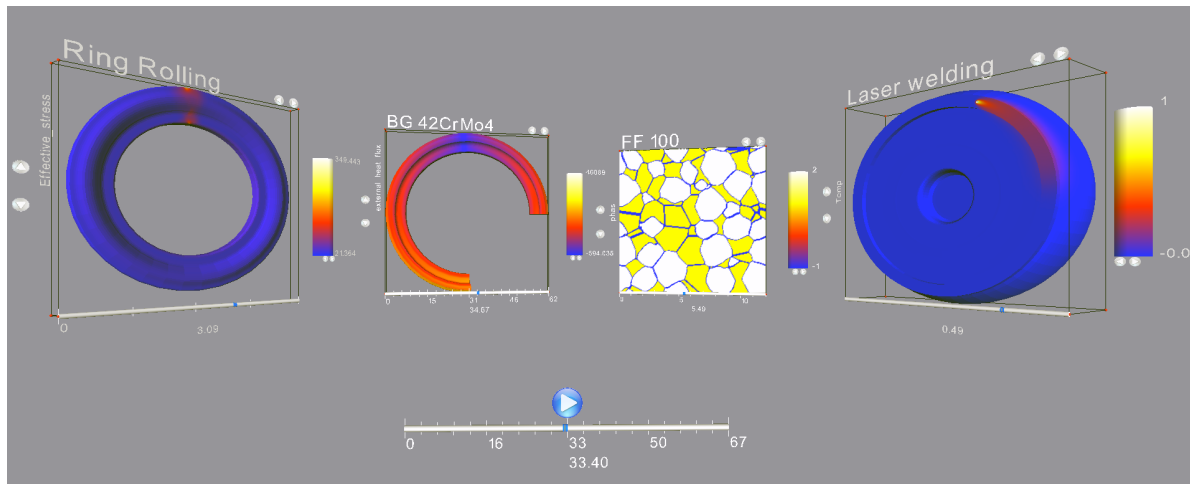


Fig. 6: 3D visualization scene

Conclusion

The presented IT-infrastructure has been successfully adopted to implement the AixViPMaP® that comprises more than ten simulation tools including commercial solutions like Abaqus. It contains general concepts which facilitate the interconnection of heterogeneous simulation resources at all technical layers from low level distributed scheduling and data transfer up to high-level visual design of simulation chains through a web interface. In the process, adaptive information integration assures data consistency at the syntactical, structural and semantic level. With these powerful tools at hand, the AixViPMaP® allows new ways of exploration, investigating simulation data at an inter-simulation tool level. Next, the underlying system will be extended by an analysis layer and an interactive user interface to trigger analysis via the visualization tool to provide a complete solution for simulation interconnection and integrated analysis - a so-called Virtual Production Intelligence solution.

Acknowledgement

The approaches presented in this paper are supported by the German Research Association (DFG) within the Cluster of Excellence “Integrative Production Technology for High-Wage Countries”.

References

- [1] J. Allard, B. Raffin: Distributed Physical Based Simulations for Large VR Applications. Proceedings of the IEEE Virtual Reality Conference, Alexandria, USA, VR '06 (2006), pp. 98-96
- [2] D. Aumueller et. al.: Schema and ontology matching with COMA++. Proceedings of the 2005 ACM SIGMOD international conference on Management of data (2005), pp. 906-908
- [3] T. Belytschko and T. Black, Elastic Crack Growth in Finite Elements with Minimal Remeshing. International Journal for Numerical Methods in Engineering, 45 (5) (1999), pp. 601-620
- [4] S. Berchtold, D. A. Keim, H. Kriegel: The X-Tree: An Index Structure for High-Dimensional Data. VLDB Conference, (1996)
- [5] P. Cerfontaine, T. Beer, T. Kuhlén, C. Bischof: Towards a Flexible and Distributed Simulation Platform. International Conference on Computational Science and its Applications (2008), pp. 867 – 882
- [6] D. Chappell, Enterprise Service Bus: Theory in Practice, O'Reilly Media (2004)
- [7] M. Ehrig: FOAM - framework for ontology alignment and mapping. Proceedings of the Workshop on Integrating Ontologies, vol. 156 (2005), pp. 72-76

- [8] D. Engmann, S. Maßmann, Instance Matching with COMA++. BTW Workshops (2007), pp. 28-37
- [9] J. Euzenat and P. Shvaiko: Ontology matching. 1. ed. Berlin; New York: Springer (2007)
- [10] I. Foster, C. Kesselman: The Grid 2, Second Edition: Blueprint for a New Computing Infrastructure. Elsevier (2003)
- [11] I. Foster: Globus Toolkit Version 4: Software for Service-Oriented Systems. IFIP International Conference on Network and Parallel Computing, Springer-Verlag LNCS 3779 (2006), pp. 2-13
- [12] M. Gagnon: Ontology-based integration of data sources. Information Fusion, 10th International Conference on Information Fusion, (2007), pp. 1-8
- [13] M. García-Solaco, F. Saltor and M. Castellanos: A Structure Based Schema Integration Methodology. Proceedings of the Eleventh International Conference on Data Engineering, Taipei, Taiwan, (1995), pp. 505-512
- [14] F. Giunchiglia, P. Shvaiko and M. Yatskevich: Discovering missing background knowledge in ontology matching. Proceeding of the 2006 conference on ECAI (2006), pp. 382-386, IOS Press
- [15] C., H. Goh: Representing and reasoning about semantic conicts in heterogeneous information systems. PhD thesis, Massachusetts Institute of Technology (1997)
- [16] A. Guttman: R-Trees: A Dynamic Index Structure for Spatial Searching. Proc. ACM SIGMOD International Conference on Management of Data, (1984)
- [17] A. Halevy, A. Rajaraman, and J. Ordille: Data integration: the teenage years. Proceedings of the 32nd international conference on Very large data bases (VLDB) (2006), pp. 9-16
- [18] T. Haupt: Cyberinfrastructure for Integrated Computational Material Engineering. Proceedings of Magnesium Technology (2010)
- [19] M. Henning: The rise and fall of CORBA. Communications of the ACM, Vol. 51 No. 8 (2008), pp. 52-57
- [20] T. K. Heok and D. Daman: A Review on Level of Detail. IEEE Computer Graphics, Imaging and Visualization '04 (2004), pp. 70-75
- [21] O. Kalden, P. Fritzen, S. Kranz: SimVis - A Concurrent Engineering Tool for Rapid Simulation Development. Proceedings of Recent Advances in Space Technologies, RAST '07 (2007), pp. 417-422
- [22] W. Kim, J. Seo: Classifying schematic and data heterogeneity in multidatabase systems. Computer, vol. 24, no. 12 (1991), pp. 12-18
- [23] Kitware Inc: The ParaView guide, ISBN 1-930934-21-4, (2008)
- [24] Kitware, Inc.: The Visualization Toolkit, ISBN 978-1930934-12-2, (2006)
- [25] U. Lang: Process Chain Integration for Simulation and Data Processing. Space Grid Workshop, Frascati, Italy (2003)
- [26] G. Laschet, J. Neises, I. Steinbach: Micro- Macrosimulation of casting processes. 4ième écoled' été de 'Modélisation numerique en thermique', vol. C8 (1998), pp. 1-42
- [27] G. Laschet, M. Apel: Thermo-elastic homogenization of 3-D steel microstructure simulated by phase-field method. Steel Resarch int., vol. 81, n. 8 (2010), pp.637-643
- [28] U. Leser: Informationsintegration: Architekturen und Methoden zur Integration verteilter und heterogener Datenquellen, Dpunkt-Verl., Heidelberg, 1st ed., (2008)
- [29] T. Meisen, D. Schilberg, K. Henning: Planner Based Data Integration for Simulation Chains in Virtual Production. Proceedings of International Conference on Science, Technology and Innovation for Sustainable Well-Being, Thailand, Khon Kaen, 2009
- [30] Y. Natis: Service-Oriented Architecture Scenario. Gartner Research Note AB-19-6751 (2003)
- [31] E. Rahm and P. Bernstein: A survey of approaches to automatic schema matching. The VLDB Journal, vol. 10, nr. 4 (2001), pp. 334-350
- [32] E. Rossiter, O. Mokrov: Integration des Simulationspaketes SimWeld in FEM Analyseprogramme zur Modellierung von Schweißprozessen, Sysweld Forum (2007)

- [33] G. Schmitz and U. Prah: Toward a virtual platform for materials processing. JOM Journal of the Minerals, Metals and Materials Society, vol. 61 (2009), pp. 19-23
- [34] G. Schopfer, A. Yang L. v. Wedel and W. Marquardt: CHEOPS: a tool-integration platform for chemical process modelling and simulation. International Journal on Software Tools for Technology Transfer, 6(3) (2004), pp. 186-202
- [35] S. Spaccapietra and C. Parent: Conflicts and Correspondence Assertions in Interoperable Databases. SIGMOD Record, vol. 20, nr. 4 (1991), pp. 49-54
- [36] S. Staab, R. Studer: Handbook on ontologies, Springer, Berlin, 2nd edition, (2009)
- [37] A. Streit, M. Romberg and D. Mallmann (Editors): UNICORE Summit 2010. Proceedings of 6th UNICORE Summit 2010, Jülich, Germany, (2010)
- [38] P. Shvaiko and J. Euzenat: A Survey of Schema-Based Matching Approaches. Journal on Data Semantics IV, vol. 3730, Springer Berlin / Heidelberg (2005), pp. 146-171
- [39] D. Thain, T. Tannenbaum and M. Livny: Distributed Computing in Practice: The Condor Experience. Concurrency and Computation: Practice and Experience, Vol. 17, No. 2-4 (2008), pages 323-356
- [40] S. Valette, J. M. Chassery: Approximated Centroidal Voronoi Diagrams for Uniform Polygonal Mesh Coarsening. Computer Graphics Forum, (2004)
- [41] C. White: Data Integration: Using ETL, EAI and EII Tools to Create an Integrated Enterprise. tech. rep. (2005), The Data Warehousing Institute
- [42] T. Zacharia et al.: Modeling of Fundamental Phenomena in Welds. Modelling and Simulation in Materials Science and Engineering, 3 (1995), pp. 265-288