ISBN: 978-3-95450-209-7 ISSN: 2226-0358 doi:10.18429/JACoW-ICALEPCS2019-MOPHA086

THE DESIGN OF EXPERIMENTAL PERFORMANCE ANALYSIS AND VISUALIZATION SYSTEM

J. Luo†, L. Li, Z. Ni, X. Zhou, Institute of Computer Application,
China Academy of Engineering Physics, Mianyang City, China
Y. Gao, Stony Brook University, New York, USA

Abstract

The analysis of experimental performance is an essential task to any experiment. With the increasing demand on experimental data mining and utilization, methods of experimental data analysis abound, including visualization, multi-dimensional performance evaluation, experimental process modeling, performance prediction, to name but a few. Oriented to the high power laser experimental facility, we design and develop an experimental performance analysis and visualization system, consisting of data source configuration component, algorithm management component, and data visualization component. It provides us feasibilities such as experimental data extraction and transformation, algorithm flexible configuration and validation, and multi-viewing presentation of experimental performance. It will bring great convenience and improvement for the analysis and verification of experimental performance.

INTRODUCTON

Large-scale physical experiments are conducted intending to capture experimental data and make further scientific analysis, which plays a significant role in multiple discipline researches, like Hefei Light Source (HLS), European X-ray Free-Electron Laser (XFEL), Large Hadron Collider (LHC), International Thermonuclear Experimental Reactor (ITER), National Ignition Facility(NIF), to name but a few [1-6].

Taking the high power laser facility as an example, a shot experiment often produces hundreds of gigabytes data, ranging from energy scalars to focal spot images. After a shot is conducted, all diagnostic subsystems start to acquire these experimental data and further archive them into the database. Subsequently, scientists utilize the visualization subsystem to observe the ordinary report of the experiment, and make statistical analysis of the experimental data captured offline, to evaluate the performance of this experiment and the status of the facility. Furthermore, fresh experimental data contributes to modify the performance model of each laser beam.

With the development of the equipment, the demand of experimental data analysis increases rapidly, resulting in a variety of data analysis and visualization programs. These programs are conventionally customized for different application scenarios, by realizing different analysis algorithms and different graphical forms. On one hand, they are well-adapted to specific scientific scenarios; on the other hand, this program development pattern leads to a great consumption of manpower involved in program development, a messy data manage framework, a high redundancy of experimental data storage, and also a narrow application range of each program.

To address these issues mentioned above, we designed a unified data management scheme; on this basis, we further realized an experimental performance analysis and visualization system. This system is under commissioning at the high power laser facility.

The next section provides an overview of current data management architecture. Problems of experimental data application are discussed in section III. Section IV illustrates the design of experimental performance analysis and visualization system. Finally, we conclude this paper in section V.

OVERVIEW OF CURRENT DATA MANAGEMENT ARCHITECTURE

The experimental equipment produces a large variety of data. These data are multi-folds, mainly including process data that indicates the real-time circumstances of the facility entities, progress data which supports the automatic scheduling and management of experimental tasks, physical phenomenon data that describes the initial input and final output of an experiment. This paper focuses on the process of the physical phenomenon data, it is also called the experimental data.

As mentioned in the former section, after an experiment is conducted, all diagnostic subsystems start to measure these experimental data and further archive them into the database. Subsequently, these experimental data are mainly utilized for online visualization, offline statistical analysis, and decision supportive computations like the modification of the performance model (see Fig. 1).

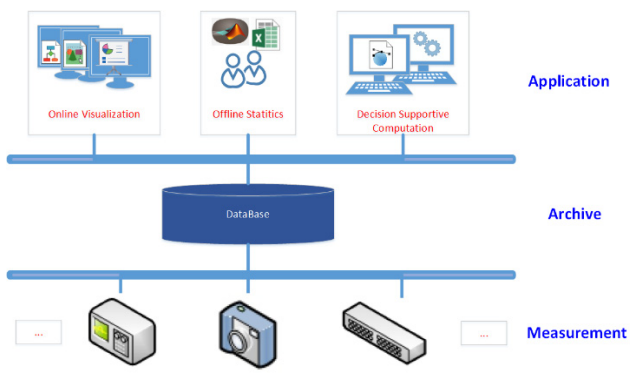

Figure 1: Flow of experimental data.

† email address: luoj1987@caep.cn

ISBN: 978-3-95450-209-7 ISSN: 2226-0358 doi:10.18429/JACoW-ICALEPCS2019-MOPHA086

PROBLEMS OF EXPERIMENTAL DATA APPLICATION

Operators utilize the experimental data to evaluate the performance of this experiment, while scientific researchers utilize them for numerical analysis or simulation, mostly by off-line methods. With the increasing demand on experimental data mining and utilization, many efforts have been made, resulting in a variety of data analysis and visualization programs. However, these programs are conventionally customized for different application scenarios, by realizing different analysis algorithms and different graphical forms.

Despite a variety of data analysis and visualization programs, the application of experimental data reveals many shortages and problems, which can be summarized as follows:

1. The experimental data coming from a large variety of measurements are heterogeneous. At the very beginning of database archiving system design, we mainly concern for the performance of data storage, while overlook the performance of data utilization. For some raw data in the database, due to their data formats, the extraction of useful information is a time-consuming task. It greatly affects the efficiency of data utilization.
2. The previous products of data analysis and visualization programs mostly utilize the C/S framework, so that each data analysis and visualization program are only applied in a narrow user range. The degree of information sharing is quite low.
3. Different scientists analyze the experiment data from different backgrounds and views, which results in different data integration patterns, different analysis algorithms, and different data visualization forms. This leads to meaningless repetitions of manpower development, and a large amount of meaningless maintaining jobs.

Hence, this motivates us to raise one question: ***Can we design a unified framework for experimental data utilization, in which the experimental data processing algorithm and development environment can be configured? Such that different experimental data application program can be produced by simple configuration, or a few codes are needed.***

To answer this question and address the above issues, the efforts as well as the contributions of this paper are: a unified data management scheme is designed; on this basis, a web-based experimental performance analysis and visualization system is further realized, which presents an effective reference for other experimental data related applications. The feature of "web-available" here is primarily to support an easy access of experimental performance information in the whole facility [7].

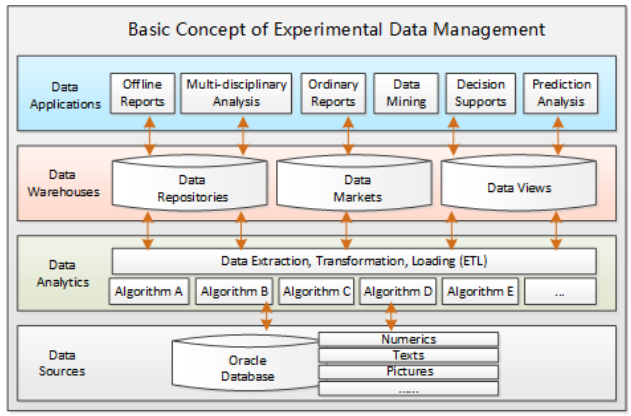

Figure 2: Basic concept of experimental data management.

EXPERIMENTAL PERFORMANCE ANALYSIS AND VISUALIZATION

We designed a unified data management scheme; on this basis, we further realized a web-based experimental performance analysis and visualization system. The main insight is to enable the unified management of data process algorithms and the personalization for data visualization.

The basic concept of the unified experimental data management scheme is depicted in Fig. 2, which consist of 4 layers, namely data applications, data warehouses, data analytics, and data sources. The original experimental data is measured and stored in Oracle databases, with a variety of formats. Data analytics realizes the statistical and other analysis computation. Data warehouses store the analysis results. Data applications supports the process of simple but sophisticated operational data visualization, by configuration according to their usages.

Based on the basic concept of the unified experimental data management scheme, we realized a web-based experimental performance analysis and visualization system, whose main task is to provide a web-available graphical experiment report, multi-dimensionally analyzing the performance of an experiment. The overview technical architecture is described in Fig. 3. It mainly includes 3 components: data source configuration, data process algorithm management, and result data visualization.

Data Source Configuration

The component of data source configuration mainly defines the data correlated to the input of analytics algorithms and data visualization. It covers a wide range, including files and databases.

For algorithm input data definitions, we currently realize the method of writing SQL sentences, not provide GUI interfaces yet, which means that SQL sentences should be written through the algorithm management platform. For visualization component, the data source configuration and correlation is operated via graphical interfaces.

ISBN: 978-3-95450-209-7 ISSN: 2226-0358 doi:10.18429/JACoW-ICALEPCS2019-MOPHA086

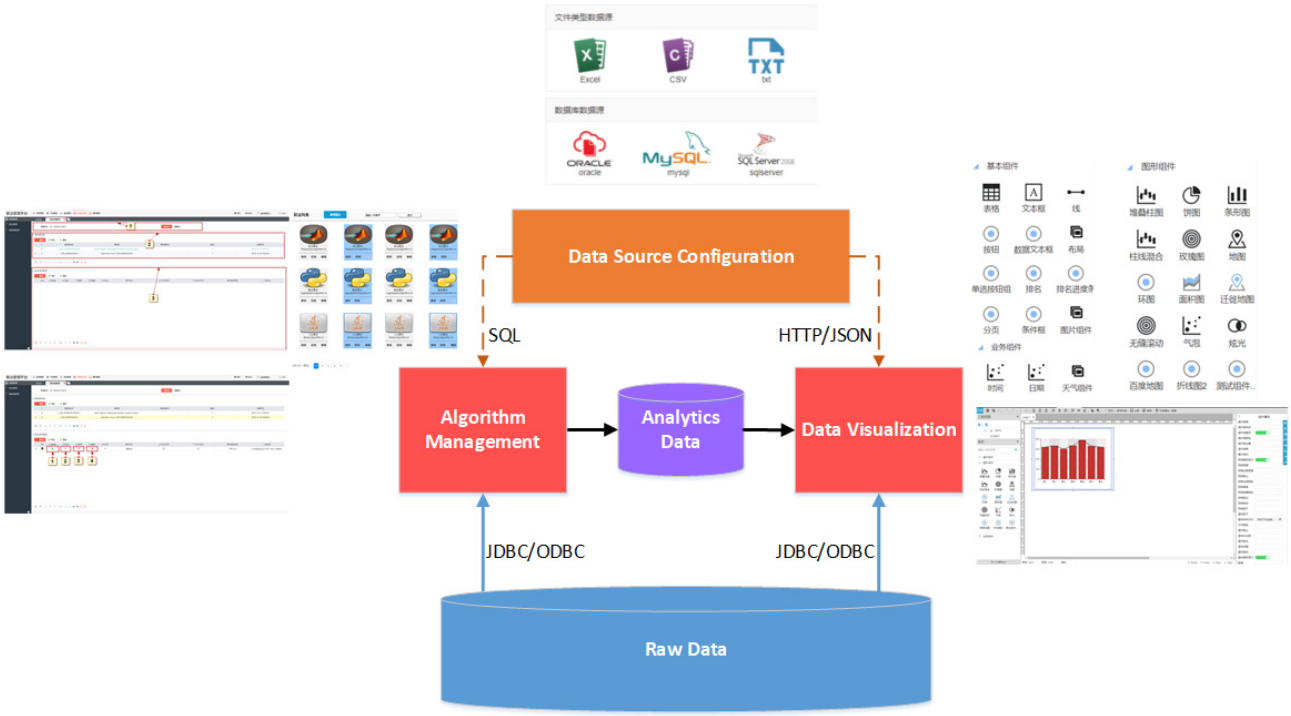

Figure 3: Technical architecture of the web-based experimental performance analysis and visualization system.

Algorithm Management

The component of algorithm management provides functionalities of algorithm integration, data input definitions, computation task deployment, and analytics computation service.

The motivation behind this component is that: different scientists analyze the experiment data from different backgrounds and views, using different computation tools; we expect to provide this algorithm management platform, such that different analysis computation environments can be integrated and run their functions under this unified framework. Hence we just concern the data deliveries and transfers, while the data process logics are implemented by physical scientists who are more professional than us.

What's more, this platform also enables the validation of some immature algorithms. This will greatly help some theoretical analysis researches.

Data Visualization

The component of data visualization reuses a commercial product, which provides the basic graphical dashboards and the editing functionality.

The main tasks of this component consists of painting the graphical layout of web pages, binding data to the dash boards, and editing the presentation forms of dash boards.

All visualization systems are web-available.

CONCLUSION

We designed a unified data management scheme; on this basis, we further realized a web-based experimental performance analysis and visualization system. This system is under commissioning at the high power laser facility, providing the analysis and visualization of experimental performance.

Our future work is to implement the GUI interfaces of algorithm data input definitions, and the integration of more databases and computation environments.

REFERENCES

[1] Y. F. Song, C. Li, K. Xuan, and G. F. Liu. "Automatic Data Archiving and Visualization at HLS-II," *Nuclear Science Technology*, vol. 29, pp. 128-133, 2018.

[2] W. Long, H. Sun, X. Li, B. Jiang, and Y. Liu. "The CSNS RCS Remote Control System Based on EPICS," *Proceedings of the 2nd International Conference on Systems and Informatics*, Shanghai, China, Nov. 2014, pp. 89-93. doi:10.1109/ICSAI.2014.7009265

[3] S. A. Bobkovl, *et al.*, "Software Platform for European XFEL: Towards Online Experimental Data Analysis," *Lobachevskii Journal of Mathematics*, vol. 39, pp. 1170-1178, 2018.

[4] A. Wallander, F. Maio, J. Y. Journeaux, W. D. Klotz, P. Makijarvi, and I. Yonekawa, "Baseline Architecture of ITER Control System," *IEEE Transaction on Nuclear Science*, vol. 58, Aug. 2011.

[5] A. D. Casey *et al.*, "Flexible Data Driven Experimental Data Analysis at the National Ignition Facility", in *Proc. ICALEPCS'13*, San Francisco, CA, USA, Oct. 2013, paper TUPPC072, pp. 747-750.

[6] R. W. Carey, K. W. Fong, and R. J. Sanchez. "Large-scale CORBA Distributed Software Framework for NIF Controls," *Proceedings of 8th International Conference on Accelerator and Large Experimental Physics Control Systems*, San Jose, USA, Nov. 2001, paper THAI001, pp. 425-429.

[7] L. Dong, M. Dylan, and M. Elder, "Web Applications for Experimental control at CLS," *Proceedings of IEEE 8th International Conference on E-science*, Chicago, IL, USA, Oct. 2012. doi:10.1109/eScience.2012.6404470