Building the DataBench Workflow and Architecture

Todor Ivanov[1], Timo Eichhorn[1], Arne Jørgen Berre[2], Tomás Pariente Lobo[3], Ivan Martinez Rodriguez[3], Ricardo Ruiz Saiz[3], Barbara Pernici[4], Chiara Francalanci[4]

[1] Frankfurt Big Data Lab, Goethe University Frankfurt, Frankfurt, Hessen, Germany
todor,timo@dbis.cs.uni-frankfurt.de
[2] SINTEF AS, Norway
Arne.J.Berre@sintef.no
[3] ATOS SPAIN SA, Spain
tomas.parientelobo@atos.net, ivan.martinez@atos.net,
ricardo.ruizsaiz@atos.net
[4] Politecnico di Milano, Italy
barbara.pernici@polimi.it, chiara.francalanci@polimi.it

Abstract. In the era of Big Data and AI, it is challenging to know all technical and business advantages of the emerging technologies. The goal of DataBench is to design a benchmarking process helping organizations developing Big Data Technologies (BDT) to reach for excellence and constantly improve their performance, by measuring their technology development activity against parameters of high business relevance. This paper focuses on the internals of the DataBench framework and presents our methodological workflow and framework architecture.

Keywords: DataBench · Big Data · Benchmarking

1 Introduction

Organisations rely on evidence from the benchmarking domain to provide answers on how their processes are performing. There is extensive information on how and why to perform technical benchmarks for the specific management and analytics processes, but there is a lack of objective, evidence-based methods to measure the correlation between Big Data Technology (BDT) benchmarks and an organisation's business benchmarks and demonstrate return on investment (ROI). New benchmarking approaches are being developed in particular in the big data domain, which presents new technological challenges. To address these challenges new benchmark initiatives focusing on machine learning and artificial intelligence like MLPerf [10, 9, 13] and AIBench [5, 4] are in development. Also, there are comprehensive studies [7, 6] on the existing Big Data benchmarks that compare and discuss the different types of benchmarks and assessment metrics. However, to the best of our knowledge, these existing benchmarks focus on technological aspects and not on business indicators. The DataBench project addresses this significant gap in the current benchmarking community's activities,

by providing verifiable benchmarks and evaluation schemes of BDT performance of high business impact and industrial significance.

The approach followed by DataBench starts with performing a comparative analysis of existing benchmarking initiatives and technologies. In fact, the goal of DataBench is not to create another benchmark, but to support an approach for efficient usage, evolution, extensions, and synergy of the available Big Data benchmarks from the international Big Data benchmarking community related to industrial requirements. Based on that, the project will proceed to develop a methodology and an economic and market analysis to assess the European and industrial significance of the BDT to be benchmarked. Industrial significance will be assessed through the investigation of the main Big Data use cases, that will allow the correlation of Big Data technical performance with business processes. Relying on all these inputs, the project will build the DataBench Toolbox, a tool which will connect and evaluate external benchmarking initiatives. Using the DataBench Toolbox and the methodology and metrics previously defined, evaluation and benchmarking will be carried out considering both business relevance and technical aspects. We foresee at least three different groups of potential users. The first group are people with a technical background that are interested in benchmarking a relevant BDT or application in their company. The second one are business people that would like to assess the usage of BDTs and applications from a business point of view. And the third one are providers of Big Data benchmarks that would like to offer their benchmarks to a broader audience of users. Currently, the Alpha version of the DataBench Toolbox has been released as a first attempt to showcase the main functions related to the Big Data technical benchmarking. More details about the project can be found in our vision paper [8] or on our DataBench website [2]. In another paper [11], we reported our initial findings on the relationship between business and technical performance indicators.

This paper focuses on describing the internal DataBench architecture. In particular, we divided it into three abstract layers: the methodological workflow, the framework architecture and the components implementation. In this paper, we look only at the first two. The Methodological Workflow (Section 2) describes the internal main processes and operations, while the Framework Architecture (Section 3) defines the logical components in which DataBench is implemented.

The paper is organized as follows: Section 2 presents the internal DataBench Methodological Workflow. Section 3 looks at the logical DataBench framework components. Finally, Section 4 summarizes our paper.

2 Methodological Workflow

This part describes in detail the internal processes and operations taking place in the DataBench framework and the logic behind this approach.

Figure 1 shows a schema of processes intended to illustrate different elements of the tooling support to be provided in DataBench to different set of users. A single user may have different roles, initially the following:

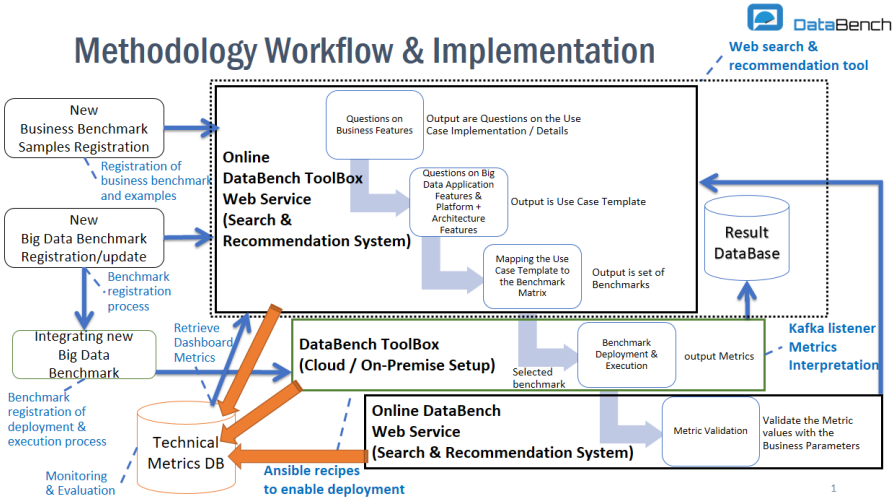

Fig. 1. DataBench Methodological Workflow

- **Benchmarking Providers:** Organizations that own a particular benchmark. They can be the actual developers of the benchmark or the organizations that maintain them. These users can register and update their benchmarks.
- **Technical Users:** Users that would like to search and potentially execute a technical benchmark. This includes the possibility of searching, downloading, executing and giving the results of the execution back to the Toolbox.
- **Business Users:** Users that would like to search and understand the business value of specific big data solutions. These users would not need to run technical benchmarks, but rather search for similar cases, business indicators, etc.
- **DataBench Admin:** People in charge of the administration of the Toolbox.

There are several processes depicted in Figure 1. On the left-hand side of the figure, the three boxes represent the registration process of two different kinds of benchmarks:

- The registration of data related to business-oriented big data benchmarks. The idea of the component located in the upper left corner of the figure (New Business Benchmark Samples Registration) is to capture domain and industry specific best practices and blueprints associated to concrete business key performance indicators (KPIs).
- The registration of technical benchmarks. The two remaining components on the left represent the way the DataBench Toolbox will capture the necessary meta data and features about technical benchmarks to enable the search and recommendation processes (New Big Data Benchmark Registration/Update

component), and to enable the automation of the deployment and the interpretation of the results of the execution of the benchmarks (Integrating new Big Data Benchmark component). Note that the registration of the automation provided by the second component is optional, in the sense that it requires the provision of deployment recipes and rules of interpretation of the results of the execution of the benchmarks which could prove a difficult task for some of the benchmarks analyzed so far. However, the aim in DataBench is to automate as many as possible technical benchmarks, so the documentation of the process to integrate the automation will be also a key part for future extensibility to other benchmarks.

The components in the center of the Figure 1 show the full process from searching to executing and visualizing the results of benchmarks. This process is divided into the following steps:

- **Search and Recommendation System:** The upper central box shows the steps to define the search criteria a user could pose to the system with the aim to select a benchmark that suits their needs. Based on those criteria (technical, business, application or platform features), the system will offer a set of potential benchmarks that could fulfill the user needs, as well as associated material (blueprints, best practices in sectors, etc.) that might facilitate the decision of the selection of the right benchmark.
- **The DataBench Toolbox setup:** The middle central box (in green in Figure 1) represents the process of deploying and enabling the execution either in cloud or in-premise of the selected benchmark. This could only happen if the registration of that benchmark provided the necessary recipes to allow the deployment. After the execution, the results of the benchmark will be sent back to the Toolbox for post-processing.
- **The validation of the metrics:** This process will allow in certain cases the matching of the technical metrics with business insights or key performance indicators (KPIs). The results of the benchmarks will be then visualized and compared to others, giving the user a clear added-value in comparison with the mere technical results that the execution of a technical benchmark may provide.
- **Monitoring and Evaluation:** This process gathers multiple metrics and internal component information with the goal to offer monitoring and evaluation capabilities to the different users of the DataBench framework. All the gathered information is stored in a central Technical Metrics Database. The data is prepared, integrated, processed and visualized into a dashboard web service that can be accessed by the different users. The key functionality of this process is to enable both DataBench administrators, technical and business users to monitor how the DataBench framework evolves in time as well as perform an evaluation of the current framework state.

At the point of writing this document, partners are in the process of defining and prototyping the look and feel of the different processes listed in this section. The initial alpha version of the DataBench Toolbox is currently implemented and will be described in detail in deliverable D3.2 [3].

3 Framework Architecture

To realize the processes described in the DataBench methodological workflow it was necessary to define and implement functional modules presented as part of the DataBench Framework Architecture in deliverable D3.1 [1]. The proposed modular framework is based on templates which are complemented with a web interface from where the user can decide and choose the metrics needed. The web interface will also act as a dashboard where the results of the executions will be gathered and shown to the user, as seen in Figure 2.

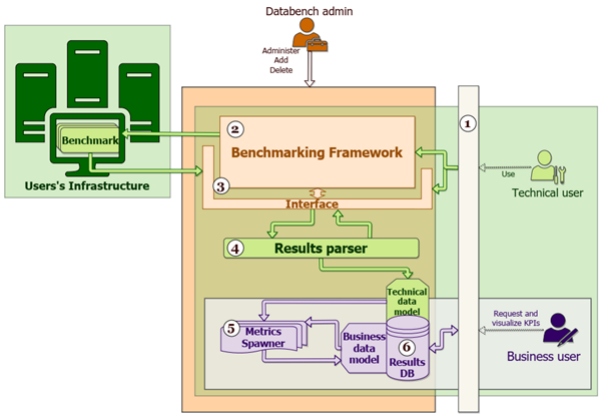

Fig. 2. DataBench Framework Architecture

The proposed modular DataBench Framework Architecture is composed of the following six interconnected modules described in detail in deliverable D3.1 [1]. The remaining two modules the Metrics DB and Metrics Dashboard are introduced in deliverable D5.3 [12] covering the functionality necessary for the DataBench monitoring and evaluation process.

1. **Web Interface** connects to the backend of the Toolbox and provides the different users with the functionality to choose which benchmarks they want to run. It is also in charge of providing a layer of configuration that the users can fill in to pre-configure the templates and the benchmarks to be run later on. The web module is also used to show in a dashboard the results of the executions and the derived metrics and business insights.
2. **Benchmark Framework Interface** module will be the main point of interaction for the administrator with the Benchmarking Framework, since he will be in charge of handling the integration, addition and deletion of the new, updated or modified benchmarks.
3. **Results Interface** enables the transfer of benchmark results to the framework either automatically by the benchmark run or manually by the user.

4. **Results Parser** converts the benchmark results into standardized data model to enable calculation of the business metrics in the next steps.
5. **Metrics Spawner** connects to the results DB module, so that it can parse the corresponding results from the technical data model and calculate the defined KPIs and at the end, write them back to the results DB.
6. **Results DB** is a place where the **Result Parsers** can store the data into and also have a place from where the web interface can read the results to show them in the dashboard.
7. **Metrics DB** is very similar to the **Results DB** module with the difference that it will store persistently the collected technical metrics. The goal is to reuse as much of the available functionality as possible, which means that the **Metric Spawner** and the **Results Parser** will be adapted to gather and prepare the metrics for the dashboards.
8. **Metrics Dashboards** offer the monitoring and evaluation functionality of the DataBench framework, represented as Platform, User (Profile) and Administrator Metrics Dashboards mapped to the different user functionality and privacy criteria.

The **Platform metrics** describe the key feature parameters of the DataBench framework that are used for static monitoring and evaluation. Examples for such metrics are total number of registered platform users, available registered benchmarks, number of use case scenarios, number of benchmark runs and others. These metrics will be available to all the different platform users to perform independent monitoring and evaluation of the platform environment.

The **User (Profile) metrics** are generated for each specific user and describe his/her activities when using the platform. Example metrics are the number of benchmark searchers, number of downloaded benchmarks, number of submitted benchmark results and history log of all operations performed by the user in the last 30 days. These metrics will be used by both business and technical users to monitor their usage of the platform as well as to have a convenient history of the latest operations.

The **Administrator metrics** are in a way combination of the above two categories. The goal of this type of metrics is to enable the full monitoring of the DataBench framework from the static platform metrics to the user actions and operations performed in the different profiles. The administrator view will enable the performance of end-to-end platform analysis on the utilization of the platform. It will help to discover patterns and trends in the user searches and most executed operations.

4 Conclusions and Future Work

This paper presents an initial overview of the internal DataBench Toolbox design. We introduce the DataBench Methodological Workflow followed by the Framework Architecture components as two abstract layers that describe in detail the functionalities in terms of internal processes, supported operations and user interfaces. By defining the technical functionality of each framework

component, the next implementation step of picking the most suitable software technologies and frameworks becomes clear and easier to realize. The latest news about the DataBench development are available on the project webpage [2] together with extended documentation of the internal architecture and design presented in this paper.

Acknowledgements

This work has been partially funded by the European Commission H2020 project DataBench - Evidence Based Big Data Benchmarking to Improve Business Performance, under project No. 780966. This work expresses the opinions of the authors and not necessarily those of the European Commission. The European Commission is not liable for any use that may be made of the information contained in this work. The authors thank all the participants in the project for discussions and common work.

References

1. D3.1 DataBench Architecture, D.T.: https://www.databench.eu/wp-content/uploads/2018/10/databench-d3.1-ver.1.0.pdf (2018)
2. DataBench: www.databench.eu (2019)
3. D3.2 DataBench Toolbox Alpha including Support for Reusing of Existing Benchmarks, D.T.: https://www.databench.eu (2019)
4. Gao, W., Luo, C., Wang, L., Xiong, X., Chen, J., Hao, T., Jiang, Z., Fan, F., Du, M., Huang, Y., Zhang, F., Wen, X., Zheng, C., He, X., Dai, J., Ye, H., Cao, Z., Jia, Z., Zhan, K., Tang, H., Zheng, D., Xie, B., Li, W., Wang, X., Zhan, J.: Aibench: Towards scalable and comprehensive datacenter AI benchmarking. In: Benchmarking, Measuring, and Optimizing - First BenchCouncil International Symposium, Bench 2018, Seattle, WA, USA, December 10-13, 2018, Revised Selected Papers. pp. 3–9 (2018)
5. Gao, W., Tang, F., Wang, L., Zhan, J., Lan, C., Luo, C., Huang, Y., Zheng, C., Dai, J., Cao, Z., Zheng, D., Tang, H., Zhan, K., Wang, B., Kong, D., Wu, T., Yu, M., Tan, C., Li, H., Tian, X., Li, Y., Shao, J., Wang, Z., Wang, X., Ye, H.: Aibench: An industry standard internet service AI benchmark suite. CoRR **abs/1908.08998** (2019), http://arxiv.org/abs/1908.08998
6. Han, R., John, L.K., Zhan, J.: Benchmarking big data systems: A review. IEEE Trans. Services Computing **11**(3), 580–597 (2018). https://doi.org/10.1109/TSC.2017.2730882
7. Ivanov, T., Rabl, T., Poess, M., Queralt, A., Poelman, J., Poggi, N., Buell, J.: Big data benchmark compendium. In: Performance Evaluation and Benchmarking: Traditional to Big Data to Internet of Things - 7th TPC Technology Conference, TPCTC 2015, Kohala Coast, HI, USA, August 31 - September 4, 2015. Revised Selected Papers. pp. 135–155 (2015)
8. Ivanov, T., Zicari, R.V., Lobo, T.P., de Lama Sanchez, N., Berre, A., Hoffmann, V., Stevens, R., Cattaneo, G., Schwenk, H., Ostberg-Hansen, C., Pepato, C., Pernici, B., Francalanci, C., Geronazzo, A., Polidori, L., Giacomazzi, P., Grobelnik, M., Hodson, J.: Databench: Evidence based big data benchmarking to

improve business performance. In: KDD 2018 Project Showcase Track (2018), https://www.kdd.org/kdd2018/files/project-showcase/KDD18_paper_1805.pdf
9. Kumar, S., Bitorff, V., Chen, D., Chou, C., Hechtman, B.A., Lee, H., Kumar, N., Mattson, P., Wang, S., Wang, T., Xu, Y., Zhou, Z.: Scale mlperf-0.6 models on google tpu-v3 pods. CoRR **abs/1909.09756** (2019), http://arxiv.org/abs/1909.09756
10. Mattson, P., Cheng, C., Coleman, C., Diamos, G., Micikevicius, P., Patterson, D.A., Tang, H., Wei, G., Bailis, P., Bittorf, V., Brooks, D., Chen, D., Dutta, D., Gupta, U., Hazelwood, K.M., Hock, A., Huang, X., Jia, B., Kang, D., Kanter, D., Kumar, N., Liao, J., Narayanan, D., Oguntebi, T., Pekhimenko, G., Pentecost, L., Reddi, V.J., Robie, T., John, T.S., Wu, C., Xu, L., Young, C., Zaharia, M.: Mlperf training benchmark. CoRR **abs/1910.01500** (2019), http://arxiv.org/abs/1910.01500
11. Pernici, B., Francalanci, C., Geronazzo, A., Lucia, P., Stefano, R., Leonardo, R., Berre, A.J., Todor, I., et al.: Relating big data business and technical performance indicators. In: Conference of the Italian Chapter of AIS. pp. 1–12 (2018)
12. D5.3 Assessment of Technical Usability, Relevance, S., Complexity: https://www.databench.eu (2019)
13. Verma, S., Wu, Q., Hanindhito, B., Jha, G., John, E.B., Radhakrishnan, R., John, L.K.: Demystifying the mlperf benchmark suite. CoRR **abs/1908.09207** (2019), http://arxiv.org/abs/1908.09207