# Towards a demonstrator of the Sigma Data Processing Architecture for BDEC 2

Gabriel Antoniu[1], Alexandru Costan[1], Ovidiu Marcu[1], Maria S. Pérez[2], Nenad Stojanovic[3]

[1]Univ Rennes, Inria, CNRS, IRISA

[2]Universidad Politécnica de Madrid

[3]Nissatech

1 May 2019

At the first BDEC 2 workshop held in Bloomington we presented a white paper introducing the Sigma data processing architecture, which aims to enable unified data processing across hybrid infrastructures combining edge, cloud and HPC systems. To do so, it aims to relevantly leverage and combine stream processing and batch processing in situ and in transit. We now introduce several software libraries and components based on which the Sigma architecture can be implemented.

**Our vision in a nutshell.** Due to an ever-growing digitalization of the everyday life, massive amounts of data start to be accumulated, providing larger and larger volumes of historical data (**past data**) on more and more monitored systems. At the same time, an up-to-date vision of the actual status of these systems is offered by the increasing number of sources of real-time data (**present data**). Today's data analytics systems correlate these two types of data (past and present) to predict the future evolution of the systems to enable decision making. However, the relevance of such decisions is limited by the knowledge already accumulated in the past. Our vision consists in improving this decision process by enriching the knowledge acquired based on **past data** with what could be called **future data** generated by simulations of the system behavior under various hypothetical conditions that have not been met in the past. This can enable hybrid modeling combining simulation models (running on HPC systems) and data-driven models (running on clouds or on cloud+edge infrastructures), to enable higher-precision data analytics (**Hybrid Analytics**). To support such analytics we advocate for unified data processing on converged, extreme-scale distributed environments thanks to a novel data processing architecture able to relevantly leverage and combine stream processing and batch processing in situ and in transit: **the Sigma architecture for data processing**.

**Reminder of the Sigma Architecture for Data Processing.** Traditional *data-driven analytics* relies on *Big Data processing* techniques, consisting of *batch processing* and *real-time (stream) processing*, potentially combined using for instance the *Lambda architecture*. This architecture uses batch processing to provide comprehensive and accurate views of batch data, while simultaneously using real-time stream processing to provide views of online data.

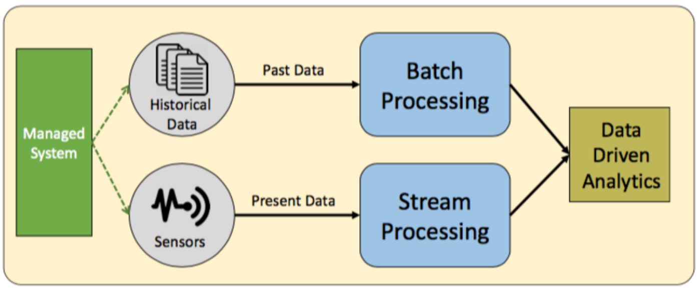


*Figure 1. The Lambda data processing architecture.*

On the other side, *simulation-driven analytics* is based on computational (usually physics-based) simulations of complex phenomena, which often leverage HPC infrastructures. The need to get fast and relevant insights from massive amounts of data generated by extreme-scale simulations led to the emergence of *in situ* and *in transit* processing approaches [Bennet2012]: they allow data to be visualized and processed interactively in real-time as data are produced, while the simulation is running.

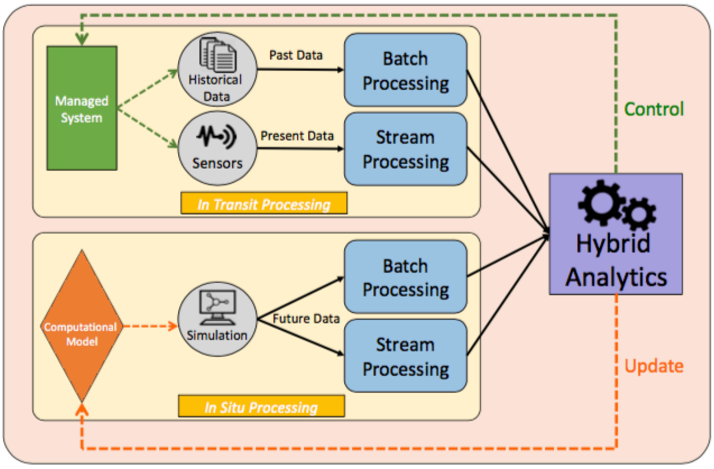

*Figure 2. The Sigma data processing architecture.*

To support hybrid analytics and continuous model improvement, we propose to combine the above data processing techniques in what we will call the **Sigma architecture**. It combines batch-based and stream-based Big Data processing techniques (i.e., the Lambda architecture) with in situ/in transit data processing techniques inspired by the HPC (Figure 2). This allows to collect, manage and process extreme volumes of past, real-time and simulated data. The architecture relies on two layers:

- The *data-driven layer* leverages the unified Lambda architecture to *update the simulation model* dynamically using the past and real-time data through a continuous learning loop [Marcu2017].
- The *computation-driven layer* uses in situ and in transit processing to *proactively control in real-time* the targeted systems. To enable this capability, in situ processing platforms need to be enhanced with Big Data analytics support. This will lead to a very efficient management of the targeted systems, based on new sets of services, not existing in current systems, like: *proactive controlling* (e.g., real-time response to unknown anomalies) based on performing data-analytics on past, present and future data; *on-demand computational simulations triggered in real-time*, driven by data analytics, in order to find opportunities for improvement, tuning and for learning forecasting system models from past data. This is expected to reduce uncertainty in prediction and to improve decision making.

To provide a reference implementation of the Sigma architecture, our approach consists in jointly leveraging two existing software components: the Damaris middleware for scalable in situ/in transit processing and the KerA unified system for ingestion and storage of data for scalable stream processing.

**Damaris: scalable in situ and in transit processing on HPC systems.** Damaris is a middleware for scalable I/O management and in situ/in transit visualization on large-scale HPC systems. It relies on dedicated cores/nodes for asynchronous I/O or in situ/in transit processing. Developed at Inria in the framework of a collaboration within the JLESC international lab, it scaled up to 16,000 cores on Oak Ridge's leadership supercomputer Titan (first in the Top500 supercomputer list at the time of the experiments, in 2013) before being validated on other top supercomputers. Active development is currently continuing at Inria, where it serves as a basis for several strategic emerging collaborations with industry (e.g. Total).

**KerA: unified ingestion and storage for low-latency, high-throughput stream processing on clouds and cloud+edge infrastructures.** KerA is a unified architecture for stream ingestion and storage aiming to optimize data processing in Big Data applications (low-latency and high throughput). It minimizes data movement within the analytics architecture, finally leading to better utilized resources. Its design principles include: a unified data model for objects and streams; eliminating data redundancies between ingestion and storage; dynamic partitioning for flexible and elastic management of stream partitions. We implemented and preliminarily evaluated a software prototype for KerA with the goal of illustrating its efficient handling of diverse access patterns: low-latency access to streams and/or high throughput access to unbounded streams and/or objects.

**Damaris+KerA: towards a framework for scalable in situ/in transit analytics on hybrid infrastructures (HPC+cloud+edge).** Damaris has already been integrated with HPC storage systems (e.g., HDF5) to enable scalable collective writing of data generated by simulations running on tens of thousands of cores (e.g., at the end of each iteration). Such data can further be analyzed (typically through offline analysis). We developed for Damaris a new real-time storage backend based on KerA: leveraging Damaris dedicated cores and shared memory components, we are able to asynchronously write in real-time the data output of a simulation (that corresponds to multiple iterations) as a series of events, logically aggregated by a stream, with metadata describing the stream's sub-partitions for each iteration. This is an important step towards a complete data processing and analytics workflow: by leveraging real-time streaming analytics (e.g., Apache Flink) efficiently coupled with KerA through shared memory, scientists can further apply machine learning techniques *in real time, in parallel with the running simulation.*

**Questions and Answers**

1. *What innovative capabilities/functionalities will the proposed candidate platform demonstrate (e.g. transcontinuum workflow, edge computing, data logistics, distributed reduction engine, etc.)?*

The above integration illustrates how data processing techniques and supporting technologies from the HPC area (in situ/in transit processing) and Big Data worlds (stream processing) can be composed in hybrid scenarios mixing HPC simulations and Big Data analytics. It provides the possibility to explore many scenarios combining simulations and analytics, for instance, to launch other simulations on demand, in parallel with the initial simulation (e.g., based on the output of a previous iteration of a simulation) in order to explore various models. As a further step, we are investigating solutions to extend this approach towards edge infrastructures, to support even more hybrid workflows running on hybrid HPC/Cloud/edge infrastructures.

2. *What applications/communities would/could be addressed?*

Applications exhibiting scenarios that mix simulations and analytics in parallel are concerned.

3. *What is the "platform vision," i.e. what kind of shared cyberinfrastructure (CI) for science would the further research/design/development of this platform lead to?*

A converged HPC+cloud+edge infrastructure supporting workflows that combine simulations and real-time analytics.

4. *How available/ready/complete is the set of software components to be used to build the demonstrator?*

Damaris is distributed as an open-source library [Damaris]. KerA is operational and has been preliminarily evaluated on cluster/cloud testbeds, it is not open-sourced. We are working on an integrated platform based on KerA and Damaris to extend the KerA+Damaris integrated data processing framework to support edge infrastructures (work in progress).

5. *As far as one can tell at this early date, to what extent can this be done with existing and/or otherwise available hardware/software/human resources?*

Building a full, convincing demonstrator requires identification of concrete use cases, (in particular, from industry) joint international collaboration and close interaction with the corresponding partners. We believe substantial extra human resources and dedicated efforts are required.

6. *What is the potential international footprint of the demonstrator?*

High. Both KerA and Damaris have already been developed through international efforts. The Sigma architecture is the result of an international design effort (as illustrated by the list of authors of this document). We plan to jointly submit H2020 project proposals to continue such joint international efforts. We are open to international collaborations within BDEC, enabling integration/connection to other platforms in global, converged scenarios. We are looking for use cases at the international scale. We are also in the process of building a startup that aims to leverage the Sigma architecture.

**References**

[Ibanez2017] R. Ibanez, E. Abisset-Chavanne, J.V. Aguado, D. Gonzalez, E. Cueto, F. Chinesta. A Manifold-Based Methodological Approach to Data-Driven Computational Elasticity and Inelasticity. Archives of Computational Methods in Engineering, 2017.

[Bennet2012] J.C. Bennet, H. Abbasi, P.-T. Bremer, R. Grout et al. Combining in-situ and in-transit processing to enable extreme-scale scientific analysis. In Proc. ACM SC'12, Salt Lake City, Nov. 2012.

[Carbone2015] P. Carbone, A. Katsifodimos, S. Ewen, V. Markl, S. Haridi, K. Tzoumas, Apache Flink: Stream and batch processing in a single engine, Bulletin of the IEEE Computer Society Technical Committee on Data Engineering 36 (4).

[Damaris] The Damaris project. https://project.inria.fr/damaris/

[Dorier2012] M. Dorier, G. Antoniu, F. Cappello, M. Snir, L. Orf. "Damaris: How to Efficiently Leverage Multicore Parallelism to Achieve Scalable, Jitter-free I/O", In Proc. CLUSTER – IEEE International Conference on Cluster Computing, Sep 2012, Beijing, China. URL: https://hal.inria.fr/hal-00715252

[Dorier 2013] M. Dorier, R. Sisneros, T. Peterka, G. Antoniu, D. Semeraro, "Damaris/Viz: a Nonintrusive, Adaptable and User-Friendly In Situ Visualization Framework", Proc. LDAV – IEEE Symposium on Large-Scale Data Analysis and Visualization, Oct 2013, Atlanta, USA. URL: https://hal.inria.fr/hal-00859603

[Marcu2017] O. Marcu, A. Costan, G. Antoniu, M.S. Pérez, R. Tudoran, S. Bortoli and B. Nicolae, Towards a Unified Storage and Ingestion Architecture for Stream Processing. IEEE International Conference on Big Data (Big Data), 2402–2407; 2017. DOI: 10.1109/BigData.2017.8258196