Astronomical Data Analysis Software and Systems XXVII
ASP Conference Series, Vol. 522
Pascal Ballester, Jorge Ibsen, Mauricio Solar, and Keith Shortridge, eds.

High Performance Astronomical Data Analysis towards Exascale

Giuliano Taffoni, David Goz, Luca Tornatore, Marco Frailis,
Gianmarco Maggio, and Fabio Pasian

INAF Osservatorio Astronomico di Trieste, Trieste, Italy;
`taffoni@oats.inaf.it`

Abstract. To date, most data-intensive HPC jobs in the government, academic and industrial sectors have involved the modelling and simulation of complex physical systems. In more recent times, the scientific data explosion is fuelling the growth of a new trend in HPC: high performance data analysis (HPDA). Some scientific disciplines (e.g. bioscience, weather and climate, security) are exploring the possibilities of HPDA to promote new insight, identifying some problems related to the nature of current HPC systems. In the design and development of the Exascale supercomputing facilities, HPDA is playing a crucial role. In this paper we will present the work done in the ExaNeSt and EuroExa EU-funded projects to build a prototype of a low power Exascale facility able to facilitate the astrophysical community in using HPC resources for data analysis and simulations.

1. Introduction

In the past decade a dramatic growth in supercomputer performance has been achieved despite the fact that processor frequencies have saturated. The chip industry has compensated these limitations by moving from single-core to multi- and many-core processors. In more recent times, the ability of increasingly powerful HPC systems to run data-intensive problems at larger scale, at higher resolution, and with more elements is contributing to the definition of a new research area in HPC, the high performance data analysis. The HPDA ecosystem is mainly driven by the proliferation of larger, more complex scientific instruments and sensor networks, from "smart" power grids to the Large Hadron Collider and Square Kilometer Array (Dewdney et al. 2009). The data ecosystem is also enriched by the growth of iterative problem-solving methods for stochastic and parametric modelling, whose cumulative results produce large data volumes.

Even the expensive hyperscale clusters that dominate today's HPC markets cannot handle the most challenging high performance data analysis workloads efficiently. HPDA requires flexible, reconfigurable and resilient platforms capable to execute (i) a large variety of scientific simulations, (ii) Big Data analytics but also (iii) complex data science workflows that take care of data ingestion, preparation, integration, analysis and eventually on-the-fly visualization. These workloads need ultrafast communication among processing elements and between processing elements and memory. This is particularly difficult to achieve on standard HPC resources, that have a rigid environment. The loosely coupled, physically distributed compute-and-memory nodes of today's high-end clusters turn daunting data analytics problems into I/O bottlenecks that can slow solution times to a crawl. The problem is that the actual HPC systems may spend a small number of compute cycles to compute a result and then spend hundreds of cycles to move the results through the system.

On the other hand, HPDA is playing a crucial role in the design and development of the new generation Exascale supercomputing facilities, capable of executing O(10^{18}) operations per second. In this paper, we present the research activity in the HPDA area done in the framework of ExaNeSt (Katevenis et al. 2016) and EuroExa[1] EU-funded projects, and the role of Astronomical applications in the design and definition of a new generation of supercomputers.

2. EuroExa and ExaNeSt projects

The ExaNeSt European project (call H2020-FETHPC-2014, n. 671553) aims to develop both the system-level interconnect and a fully-distributed NVM (Non-Volatile Memory) storage together with the cooling infrastructure for an European Exascale-class supercomputer based on low-cost, low-power many ARM-cores plus computing accelerators (FPGAs). The Consortium brings together industrial and academic research expertise to develop all the different components of the system, namely: a low latency, high throughput unified RDMA-based interconnect enabling the scalability of both storage and I/O bandwidth, together with the compute capacity and a novel distributed storage architecture based on local storage devices attached to the computing nodes enabling near-data computation and reducing the energy and latency of I/O traffic (based also on NVM).

EuroEXA brings a holistic foundation from multiple European HPC projects and partners together with the industrial SME to co-design a ground-breaking platform capable of scaling peak performance to 400 PFLOP in a peak system power envelope of 30MW; over four times more performance at four times more energy efficiency of today's HPC platforms. Further, it targets a power usage effectiveness parity rating of 1.0 through use of renewables and immersion-based cooling. To achieve the demands of extreme scale and the delivery of exascale, EuroEXA embraces the computing platform as a whole, not just component optimization or fault resilience, and it relies on co-design as a method to design a platform that really responds to the needs of real scientific and industrial applications. EuroEXA co-designs a balanced architecture for both compute- and data-intensive applications and it provides a homogenised software platform offering heterogeneous acceleration based on FPGA with scalable shared memory access.

It targets very big data workflow made by multiple real applications that interact with large input data sets and generate other large data sets as an output. The net effect is a complex web of data access and processing, representing a different degree and range of accesses that traditional storage systems are not built to handle. ExaNeSt and EuroEXA are designing a platform able to answer to the needs of this kind of demanding HPDA applications, for example a new approach towards the filesystem and a novel system software is under design.

[1] EuroEXA project `http://www.euroexa.eu`

3. Astrophysical applications

We selected a set of astrophysical applications both in the area of numerical simulations and in data reduction and analysis of big data as co-design counterpart to the HW. In particular we use three codes: GADGET, MIRIAD and GAIA AV.

GADGET (Springel 2005) is an N-body and hydrodynamical code for large-scale, high-resolution numerical simulations widely used to study formation and evolution of cosmic structures, galaxy formations and astrophysical processes. The code will be ported on the EuroEXA prototype and re-engineered to profit of the exascale capacity of the testbeds.

MIRIAD (Tingay et al. 2013) big data reduction suite is the basis for a working pipeline for analyzing large radio astronomy datasets from the Murchison Widefield Array (low frequency precursor to the SKA). MIRIAD will be redesigned to work on a massive parallel exascale platform, stressing and testing the IO capabilities with an example of HPDA application.

AVU-GSR is used as part of the analysis pipeline of the ESA's (European Space Agency) cornerstone GAIA mission. It is a matrix solver for extremely-large sparse matrices, based on the (least squares) LSQR method, which finds a solution for the astrometric parameters of up to 100 million primary stars. A parallel version of the code is currently used and developed, based on hybrid MPI+OpenMP GAIA and will be tested with the new System software offered by EuroEXA (e.g. OmpSs or PGAS).

4. System software, virtualisation and containers

New supercomputing facilities will be more complex to use than actual HPC resources; to profit of their capabilities it will be necessary to use advanced system software (libraries or computing environments) that will hide the intrinsic complexity of the systems. Framework as COMPSs or OmpSs are now extended to provide simple access to massive parallel architectures and accelerators. Graphical analytics frameworks (e.g. IndeX 3D volumetric visualization system), semantic analysis, knowledge discovery algorithms (e.g. IBM Watson cognitive computing portfolio) will be "embedded" in the new supercomputing machines.

One of the key aspects of HPDA is the use of interactive computing, real-time operations, 3D visualization and deep neural analysis. For example, to find on the fly patterns that we do not expect and that requires an immediate reaction either in terms of visualization or machine learning analysis. Interactive computing is achieved integrating system software with interactive environments. We test the use of pyCOMPSs with Jupyter[2] running on a parallel cluster.

Given the complexity, the use of virtualisation and containers is necessary. We are experimenting the use virtual machines based on KVM and containers, in particular Docker and Singularity. In general virtualization, by inserting an additional layer between the application and the hardware, can negatively impact the user-perceived performance. However, in ExaNeSt it has been demonstrated that, thanks to an advanced pass-through technology, it is possible to enable close to bare-metal performance for common HPC APIs as MPI and I/O (in particular for what regards parallel filesys-

---

[2] `http://www.jupyter.com`

tems). We will evaluate the benefit of this technology in the framework of the Euclid SDC-IT (the Italian Science Data Center) (Laureijs et al. 2012). We performed tests on containers and VMs for data analysis on a (currently conventional) HPC infrastructure. We focused on the Euclid data reduction and analysis software which requires each scientific workflow of the science ground segment to be run in any of the 9 SDCs of the Euclid Consortium. To this purpose, key architecture choices have been the isolation of the pipeline software from an SDC computing infrastructure, through the so called Infrastructure Abstraction Layer (IAL), and the adoption of a virtualized common development and execution environment.

5. Conclusions

We discuss our work towards the use of exascale supercomputing facility designed for HPDA. In particular we present the ongoing work we are doing in the framework fo two European Projects: EuroEXA and ExaNeSt. Applications and in particular computing demanding astronomical applications, has been ported on the prototypes with successful results and now we are re-designing them to use the new environment and the system software developed in the framework of the European projects.

Acknowledgments. This work has received funding by the European Union's Horizon 2020 Research and Innovation Programme under the ExaNeSt project (Grant Agreement No. 671553) and EuroEXA project (Grant Agreement No. 754337)

References

Dewdney, P. E., Hall, P. J., Schilizzi, R. T., & Lazio, T. J. L. W. 2009, IEEE Proceedings, 97, 1482

Katevenis, M., Chrysos, N., Marazakis, M., & et al. 2016, in 2016 Euromicro Conference on Digital System Design (DSD), 60

Laureijs, R., Gondoin, P., Duvet, L., & et al. 2012, in Space Telescopes and Instrumentation 2012: Optical, Infrared, and Millimeter Wave, vol. 8442, 84420T

Springel, V. 2005, MNRAS, 364, 1105. astro-ph/0505010

Tingay, S. J., et al. 2013, Publications of the Astronomical Society of Australia, 30, e007