Big Data analysis and reduction on Exascale infrastructure: the AVU-GSR application in the EuroExa project

U.Becciani[1], F. Vitello[1], E. Sciacca[1], A. Calanducci[1], A. Costa[1]
S. Riggi[1], A. Vecchiato[2]
[1]INAF - Catania Astrophysical Observatory
[2]INAF – Turin Astrophysical Observatory

EuroEXA – H2020 EU FETHPC-2016

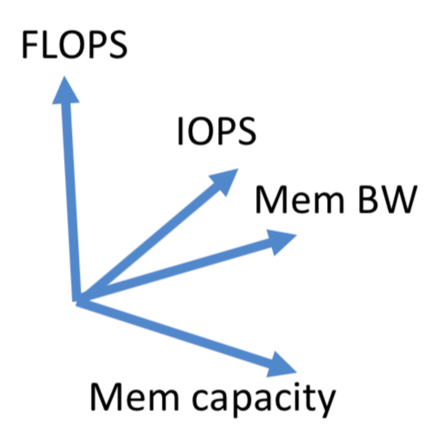
Starting in September 2017, a new H2020 EU initiative named EuroExa - funded under the FETHPC-2016 program (call H2020-FETHPC-2016, n. 754337) - will build upon ExaNeSt results to deliver a world-class, ARM-based HPC platform prototype.

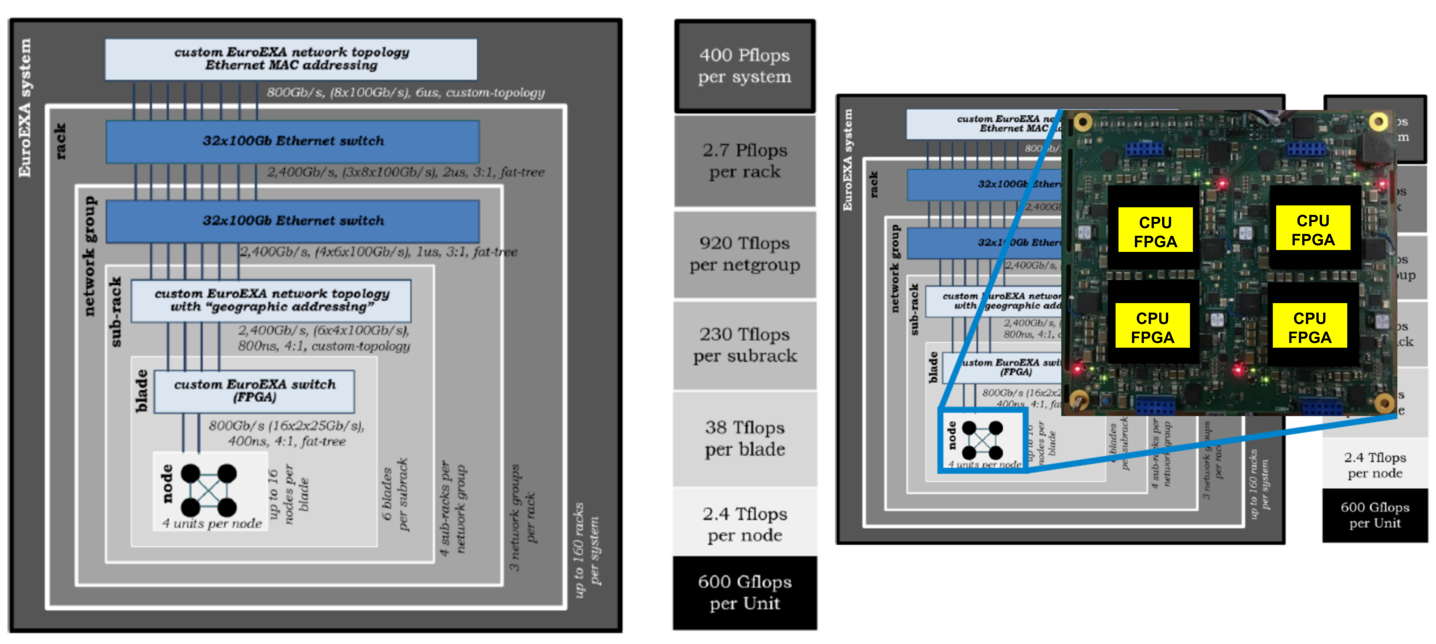
The main deliverable of the project is a **co-designed platform** capable of scaling to a peak performance to **400 PFlops** in a system with a peak **power envelope of 30 MW** achieved through a customized ARM-based processing unit, the adoption of FPGAs for data-flow acceleration and the integration at rack level of a low latency, high throughput ExaNest-based network architecture.

EuroExa will provide a homogenised software platform with **advanced runtime capabilities** supporting novel parallel programming paradigms, dataflow programming, heterogeneous acceleration and scalable shared memory access porting and optimizing a rich mix of both traditional and next-generation HPC applications

EuroEXA: main project characteristics

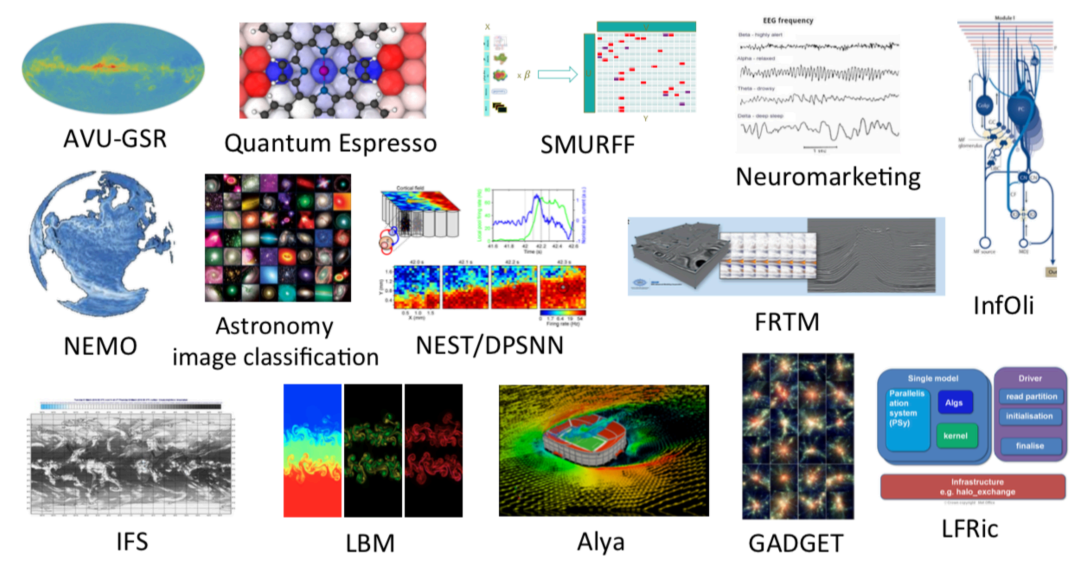
EuroEXA will **co-design a balanced architecture** for both compute- and data-intensive applications using a cost-efficient, modular-integration approach enabled by novel inter-die links and the tape-out of a resulting EuroEXA processing unit with integration of FPGA for data-flow acceleration.





Technology from FORTH (QFDB board):

- 12 cm x 13 cm
- 4 ARM Processors and 4 FPGA Accelerators
- M.2 SSD
- 4 x SODIMMs + Onboard RAM
- Daughterboard style
- 160 Gb/s of I/O

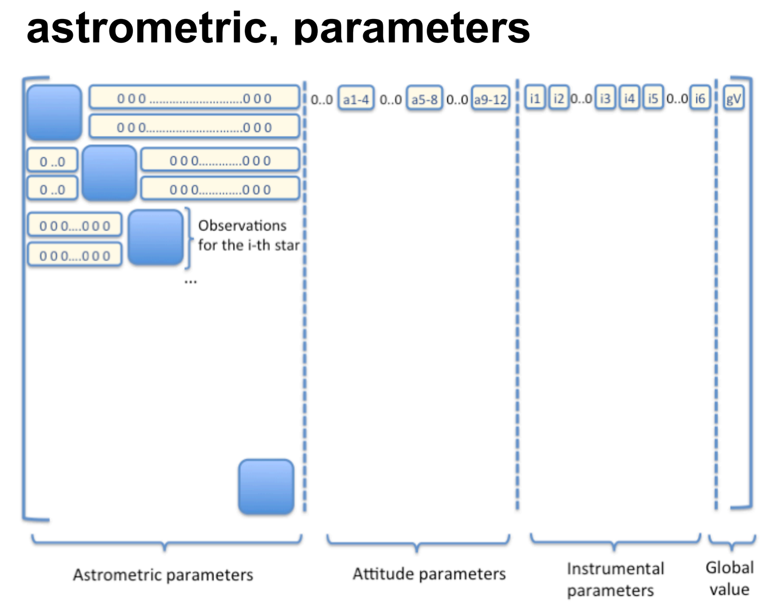


Working together with **a rich mix of key HPC applications** from across climate/weather, physics/energy and life-science/ bioinformatics domains we will demonstrate the results of the project through the deployment of an integrated and operational peta-flop level prototype.

AVU-GSR OmpSs+MPI Application @ EuroEXA

The Gaia ESA mission will estimate the astrometric and physical data of more than one billion objects, providing the largest and most precise catalog in the history of astronomy. An astrometric verification unit (AVU) was instituted by the Gaia Data Processing and Analysis Consortium (DPAC). One of its task is to implement and operate an independent. global sphere reconstruction (GSR) We present in this work the preliminary results obtained within the EuroExa project porting the AVU-GSR code to novel HPC infrastructures employing the new paradigm of OmpSs (BSC) with MPI and ARM architecture to exploit the potential of an Exascale platform.

For each observation, the Total Matrix stores the astrometric, parameters



10^8 stars, each star observed 720 times

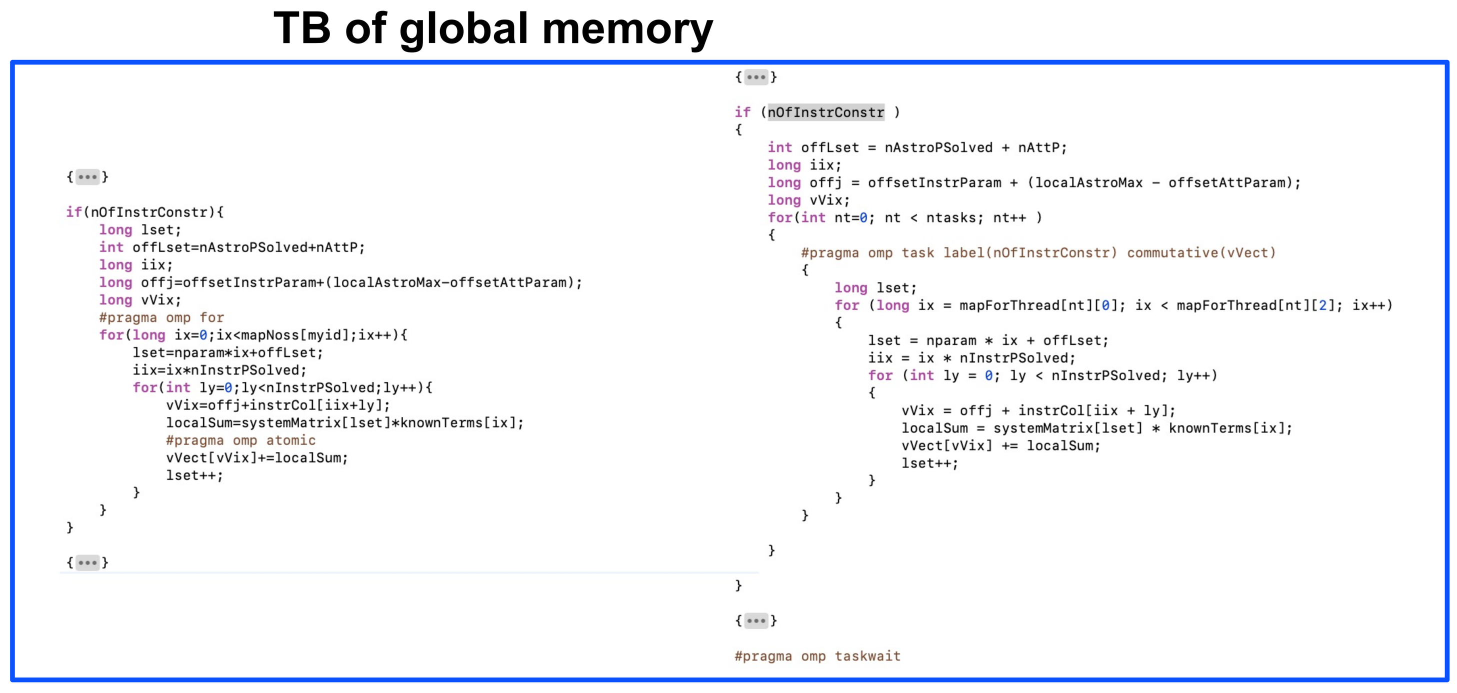
→ $7.2*10^{10}$ equations with 24 parameters for each obs.

→ $3*10^{12}$ coeff.

→ 15 TBytes Ram

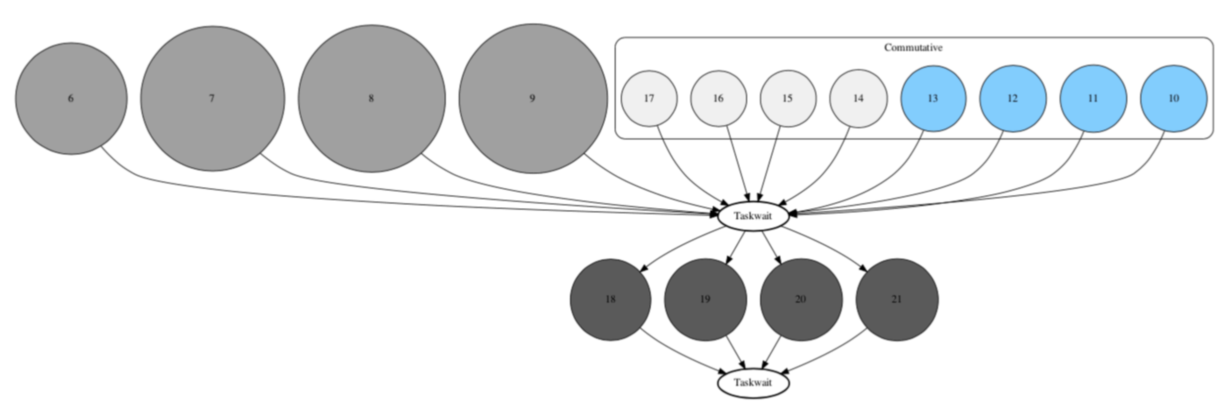
EuroEXA core ap can be executed autonomously with simulated parameters from few GB of up to hundreds TB of global memory

The developed demonstrator (core application) is a subset of complete AVU-GSR application to compute GaiaGsr system solution employing OmpSs which is a modified version of the OpenMP programming model. While OpenMP has a fork-join model, OmpSs defines a thread-pool model where all the threads exist from the beginning of the execution. Among these threads, only the master thread starts executing the user code while the other threads (workers) remain ready to execute the task when available. The master thread can generate work for the other threads by means of the regular OpenMP work-sharing or task constructs.
The structure of the core application was redesigned to use data-dependencies be tween the different tasks of the program,



Poster presenter

Ugo Becciani is a senior Research Astronomer at the INAF Astrophysical Observatory of Catania. He has been PI and Co-I of several research projects on: computational astrophysical topics, supercomputing and parallel computing, visual analytics, data exploration, virtual observatory.



Graphical representation of AVU-GSR running with MPI + OmpSs workers

The research leading to these results has received funding from the European Commissions Horizon 2020 under the FETHPC program (call H2020-FETHPC-2016, n. 754337)

EURO EXA