LightAMR data structure for ultra compact AMR data volume

Loic Strafella*[1]

[1]Institut de Recherches sur les lois Fondamentales de lÚnivers – Commissariat à l'énergie atomique et aux énergies alternatives : DRF/IRFU/DEDIP/LILAS – France

Résumé

The evolution of parallel I/O library as well as new concepts such as 'in transit' and 'in situ' visualization
and analysis have been identified as key technologies to circumvent I/O bottleneck and huge data production
by future exascale applications. Nevertheless, data structure and data format can also be improved for both
reducing I/O volume, data storage and improving data interoperability between data producer and data consumer.
In this talk, we propose to present a very lightweight and purpose-specific post-processing data model
added to the RAMSES code which significantly reduce data volume production by several order of magnitude saving
terabytes of disk storage. Results of benchmark of this lightAMR data format will be presented as well as
lossless and lossy compression algorithms allowing us to reduce the data volume of an use case dataset:
Extreme-Horizon (cosmological simulation), going from 2.8 Terabytes of data down to less than 170 Gigabytes.
Finally, we will present the first result of a large scale simulation of isolated galaxy with a base resolution of
_˜1pc in 90% of the mass of gas. This multiscale simulation to study the turbulent cascade integrates star formation
(Kretschmer M. and Teyssier R., 2020), cooling/heating (Hennebelle et al, 2015) and stellar feedback will produce
resonable amount of data, despite requiring almost one hundred thousand cores, for analysis thanks to the
lightAMR data model and a parallel I/O library. For consistency with the presentation of a new post-processing
data structure, the data analysis strategy will be discussed.

*Intervenant