

LA ASTROFÍSICA COMPUTACIONAL EN ESPAÑA:

En este artículo pretendemos dar una visión global de la situación en España en relación con la Astrofísica/Cosmología computacional, entendida ésta como el área de investigación donde la herramienta principal es la realización y análisis de simulaciones numéricas de objetos astrofísicos individuales, o del Universo en su conjunto. No hemos pretendido hacer una revisión exhaustiva de los resultados científicos derivados de las simulaciones, sino más bien exponer las líneas principales de investigación, basándonos en el uso por parte de los grupos españoles de las infraestructuras de supercomputación de altas prestaciones (HPC) accesibles a nivel estatal y europeo. Para ello hemos recopilado a partir de datos públicos, los tiempos de computación concedidos por los diferentes Comités de Acceso a proyectos liderados por investigadores afiliados a cualquier universidad o instituto de investigación científica española.



Gustavo Yepes

Universidad Autónoma de Madrid
gustavo.yepes@uam.es

Claudio Dalla Vecchia

Instituto de Astrofísica de Canarias
caius@iac.es

INTRODUCCIÓN

La Astrofísica es una ciencia puramente observacional, donde el objeto de estudio no puede ser manipulado. Por tanto, es difícil poder realizar experimentos para poner a prueba las predicciones de las teorías físicas que intentan explicar los mecanismos responsables de la formación y evolución de los diferentes sistemas astrofísicos, incluido el Universo en su conjunto.

Gracias al enorme avance de los sistemas de computación, en Astrofísica, como muchas otras ramas de la ciencia que comparten con ella la imposibilidad de realizar experimentos, es ahora factible construir laboratorios virtuales donde se pueden testar las diferentes teorías.

El campo de aplicación de la simulación numérica en Astrofísica es enorme, precisamente por la razón anteriormente expuesta. Abarca desde la simulación del movimiento de objetos en el Sistema Solar, la dinámica de la atmósfera solar y de los planetas, la formación y evolución de estrellas en nubes moleculares, y su relación con el medio interestelar, la dinámica de sistemas binarios de estrellas de neutrones y agujeros negros que producen ondas gravitacionales, los procesos de formación y evolución de galaxias y cúmulos de galaxias, y la formación y evolución de la estructura a gran escala en el Universo.

Todos estos ejemplos astrofísicos requieren de la modelización de diferentes procesos físicos que actúan simultáneamente entre las diversas componentes del sistema que se pretende estudiar: fenómenos hidrodinámicos de la componente gaseosa, o magneto hidrodinámicos en el caso de plasmas, fenómenos radiativos debido a las transiciones atómicas, reacciones nucleares en interiores estelares, interacción entre materia bariónica y radiación, efectos debido a la presencia de materia y energía oscuras a grandes escalas, radiación de ondas gravitacionales por coalescencia de estrellas o agujeros negros, etc.

Cada uno de estos fenómenos ocurre en escalas espacio-temporales que pueden variar en varios ordenes de magnitud. Es por ello que, al igual que ocurre en otras áreas de investigación con la misma problemática (por ejemplo, en meteorología o climatología), hay que realizar aproximaciones para tener en cuenta aquellos efectos debido a los fenómenos que ocurren en escalas por debajo de la resolución

máxima que se puede alcanzar en la simulación y que depende sobre todo de las capacidades computacionales del sistema informático donde se vaya a realizar. A mayor capacidad computacional, mejores resoluciones espacio-temporales se pueden conseguir y, por tanto, menos aproximaciones a la física del sistema se necesitan, con lo que los resultados son más realistas. Sin embargo, con mejores resoluciones más fenómenos físicos, que antes se aproximaban en un conjunto macroscópico, se tienen que modelar en detalle. En este sentido, la Astrofísica computacional ha sido una de las áreas de investigación que ha necesitado siempre de los sistemas de supercomputación más potentes en cada época. En consecuencia, dada esta fuerte interrelación entre el acceso a los grandes recursos computacionales y la generación de simulaciones realistas de diferentes problemas astrofísicos, el desarrollo de grupos de investigación competitivos está condicionado a la disponibilidad de acceso a estas infraestructuras.

En España, la infraestructura de supercomputación era bastante limitada hasta la instalación por IBM del superordenador *MareNostrum* en Barcelona en el año 2004. Previo a esa fecha, los pocos grupos que se dedicaban a la simulación astrofísica debían ingeniárselas para conseguir acceso a recursos de computación fundamentalmente de centros extranjeros, mediante colaboraciones con otros grupos que sí que tenían posibilidad de pedir tiempo de computación en los centros de datos de países como EE.UU., Alemania, Francia, Reino Unido, o Italia. En España había ciertas infraestructuras de computación en Universidades, centros de investigación, y algunas empresas de alta tecnología, si bien no existía una política de acceso bien definida. Había que llamar a muchas puertas para conseguir acceso a ellas, muchas veces sin éxito. Por contraposición, si nos fijamos en la Astrofísica observacional en España, ésta tuvo un desarrollo vertiginoso debido a la instalación de grandes infraestructuras de telescopios de Canarias, Calar Alto, etc., que permitió disponer de tiempo de observación garantizado para desarrollar proyecto de investigación liderados por grupos españoles.

La situación para los astrofísicos computacionales españoles cambió radicalmente en 2004 con la creación de un centro español de supercomputación y la instalación por parte de IBM del mayor supercomputador europeo de aquella época.

LA INSTALACIÓN DE MARENOSTRUM Y LA CREACIÓN DE LA RES

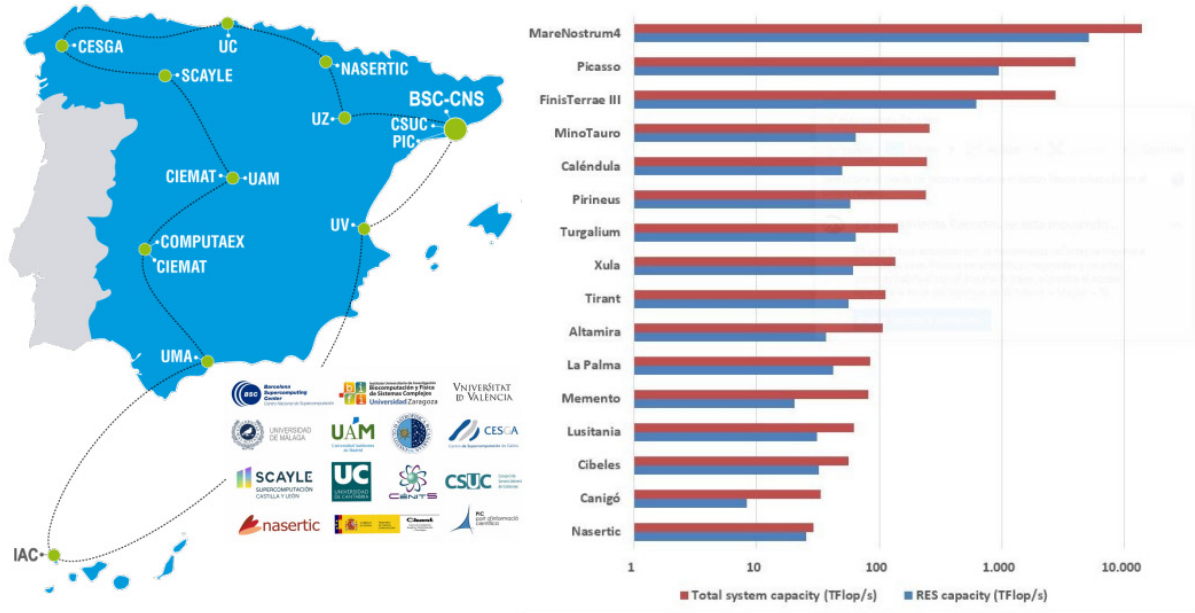
El precursor del *Barcelona Supercomputing Center-Centro Nacional de Supercomputación* (BSC-CNS) fue el Centro de Paralelismo de Barcelona (CEPBA), un centro asociado a la Universitat Politècnica de Catalunya (UPC) que desde 1991 prestaba servicios a los distintos departamentos de la universidad.

En el año 2000 el CEPBA firmó un acuerdo con IBM de cuatro años de duración para crear un instituto de investigación, que se llamó CIRI y, en 2004, el acuerdo entre IBM y el CEPBA es ampliado por iniciativa del Ministerio de Educación y Ciencia, la Generalitat de Catalunya y la UPC creando el BSC-CNS. En abril del año siguiente inician sus actividades con el supercomputador *MareNostrum*, el de más capacidad de cálculo de toda Europa en aquel momento y cuarto a nivel mundial.

En el año 2006 se produce una actualización del *MareNostrum* que dobla su capacidad y continúa ostentando el primer puesto de Europa. En el mismo año se amplió el supercomputador *MAGERIT* perteneciente al CeSViMa (*Centro de Supercomputación y Visualización de Madrid*) y dependiente de la Universidad Politécnica de Madrid, conformando los dos primeros nodos de la *Red Española de Supercomputación* (RES), bajo la coordinación del BSC (<https://www.res.es>).

A lo largo del año 2007 entran en funcionamiento los distintos nodos de la RES que se crean a partir de la actualización del supercomputador *MareNostrum*, incrementando los recursos de cómputo ofertados por la RES. Al día de hoy, la RES está compuesta de 15 nodos de 14 instituciones, sumando una capacidad de cálculo de más de 20 petaflops, previo a la puesta en producción de *MareNostrum-5*, del que hablaremos más adelante

Todos los nodos de la RES tienen la característica que son sistemas basados en procesadores INTEL, o AMD, multinúcleos y unidos por una red de alta velocidad y baja latencia. En este sentido, cada uno de los nodos RES pueden ser utilizados de forma casi transparente por una gran variedad de códigos numéricos con la condición de que estén paralelizados, usando MPI o bien MPI + OpenMP.



Total RES capacity 20 PFLOPS

Figura 1. Composición actual de la Red Española de Supercomputación.

En septiembre de 2011 se añade al BSC, *MINOTAURO*, el primer superordenador basado en aceleradores gráficos (GPUs) de NVIDIA fabricado por Bull y en noviembre de ese año el BSC alcanza un acuerdo con Intel por el cual establecen en conjunto un laboratorio I+D sobre los problemas de escalabilidad y programación que puedan surgir en supercomputadores a la escala exaflop (1000 petaflops o 1.000.000 teraflops).

El BSC-CNS ha llegado también a un acuerdo con NVIDIA para construir un superordenador basado en arquitectura ARM de bajo consumo energético con aceleradores NVIDIA Tesla, bajo el proyecto llamado *Mont-Blanc*.

El BSC gestiona el superordenador *MareNostrum* desde su instalación en la capilla de la Universidad Politécnica de Cataluña en 2004. Hasta el momento se han instalado cinco versiones, la penúltima de ellas, *MareNostrum-4*, entró en funcionamiento en junio de 2017 y la última versión, *MareNostrum-5*,

ha entrado en producción recientemente, en marzo de 2024. Las versiones anteriores de *MareNostrum* no han desaparecido. Siguen funcionando de forma fraccionada en diversos nodos de la RES, dando servicio a las necesidades de los usuarios locales de los nodos y contribuyendo a la oferta de horas de supercomputación de la RES en su conjunto.

En la actualidad, la RES es una de las Infraestructuras Científica-Técnica Singular (ICTS) en España, y la única relacionada con la computación de altas prestaciones.

El acceso a los recursos de la RES se realiza a través de una convocatoria abierta todo el año, con tres resoluciones anuales. Los grupos de investigación de cualquier disciplina científica que realicen su trabajo en universidades y centros públicos en España pueden solicitar recursos a la RES, si bien el acceso está abierto a solicitantes de centros de otros países. Para facilitar la selección de las propuestas que obtiene tiempo de computación, existe

un Comité de Acceso compuesto por cuatro miembros que ocupan el cargo por un periodo máximo de cuatro años. Este Comité de Acceso está asesorado de dos *Paneles de Expertos*:

- El *Panel Técnico*. Los nodos de la RES nombran a los miembros de este panel, que es el responsable de realizar la evaluación de las solicitudes desde el punto de vista técnico.
- El *Panel Científico*. Está formado por científicos reconocidos a nivel internacional y externos a la RES que se divide en seis áreas científicas:
 - Astronomía, Espacio y Ciencias de la Tierra (AECT)
 - Biología y Ciencias de la Vida (BCV)
 - Física (FI)
 - Ingeniería y Matemáticas (IM)
 - Química Homogénea (QH)
 - Química Heterogénea y de Sólidos (QHS)

Cada solicitante tiene que elegir el área científica que se mejor se ajuste su propuesta. Todas las solicitudes de cada área son evaluadas, al menos, por dos revisores del área correspondiente. Por cada área se nombra uno o dos coordinadores que son los encargados de distribuir las solicitudes entre el panel de revisores de cada área y que se encargan de compilar los informes y puntuaciones obtenidas por cada solicitud y elaboran la lista priorizada de solicitudes que es luego enviada al Comité de Acceso para decidir el listado final de actividades aceptadas y las horas de cálculo asignadas a cada una de ella, según las horas de cómputo disponibles.

En el ámbito que nos concierne de Astronomía, compartimos un mismo panel con Espacio y Ciencias de la Tierra. La distribución de recursos entre las distintas áreas científicas de la RES depende de la calidad de las solicitudes enviadas. En este sentido, tenemos que presumir que en el área de AECT, a pesar que el número de investigadores es considerablemente menor que en otras áreas como Química, Física, o Biología, tenemos un éxito considerable en la consecución de recursos.

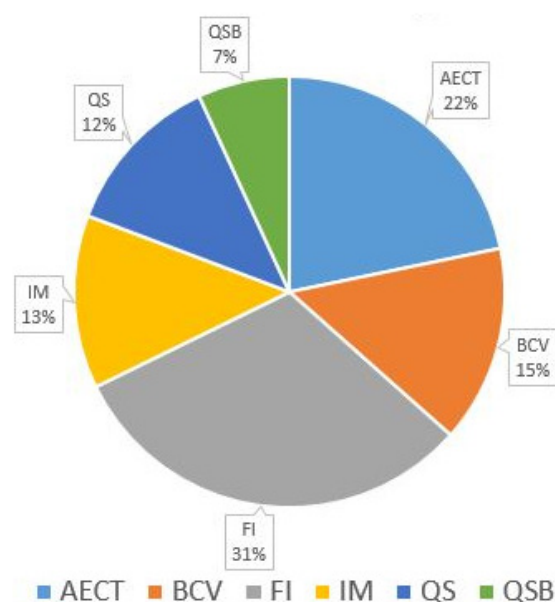


Figura 2. Distribución de horas asignadas en 2020 a las diferentes disciplinas científicas.

En la Figura 2, tomada de la página web de la RES, se representa la distribución de los recursos por áreas para el año 2020. Como se puede ver, el área de AECT consiguió casi un tercio de todos los recursos de la RES, a pesar de que 2020 fue el año de la pandemia de COVID y que se priorizaban los proyectos de Biología relacionados con la misma. En general, esta tendencia se ha mantenido más o menos constante a lo largo de los años.

Nos centraremos ahora en analizar en detalle el uso de los recursos de la RES por parte de la comunidad de Astrofísica computacional en España. Para ellos hemos recopilado los datos sobre los proyectos que han conseguido recursos de la RES dentro del periodo comprendido entre 2015 y 2023 y que estaban relacionados con algún tema de Astrofísica/Cosmología computacional y cuyo investigador principal desarrollaba su labor investigadora en cualquier institución científica española.

De esta base de datos hemos realizado diferentes estudios estadísticos. En primer lugar, hemos querido identificar cuáles son los temas científicos donde se han

concedido más horas de supercomputación (ver Figura 3). Seguidamente, queríamos saber cómo se han distribuido esos recursos entre los diferentes nodos de la RES (ver Figura 4) y finalmente hemos representado también la asignación de recursos por institución del investigador principal de los proyectos (Figura 5).

Distribución de recursos de RES por temas científicos

En primer lugar, para simplificar las cosas, hemos clasificado los proyectos en los siguientes temas científicos:

- **Cosmología:**
 - Simulaciones N-cuerpos de gran volumen para estudio de estructura a larga escala. Aplicaciones para surveys de galaxias.
- **Galaxias:**
 - Simulaciones de formación de galaxias a diferentes épocas cosmológicas en general y modelado de la formación de la Vía Láctea y el Universo Local.
- **Medio Interestelar**
 - Simulaciones del medio interestelar, turbulencia, evolución química.
- **Física Estelar**
 - Simulaciones de formación estelar en nubes moleculares y modelización de últimos estadios de evolución estelar.
- **Ondas Gravitacionales**
 - Modelado de la emisión de ondas gravitacionales en coalescencia de estrellas de neutrones o agujeros negros de cara a su uso por colaboraciones internacionales LIGO /VIRGO.
- **Astrofísica Relativista**
 - Simulaciones magneto-hidrodinámicas (MHD) de fluidos relativistas y en entornos de alto campo gravitatorio (agujeros negros, estrellas de neutrones).
- **Física Solar**
 - Simulaciones magneto-hidrodinámicas de plasmas en la atmósfera solar.
- **Gaia**
 - Reducción de datos del satélite *Gaia*.
- **Planetas**
 - Simulaciones de atmósferas planetarias (Marte, Jupiter, Saturno).
- **Espacio**
 - Simulaciones de turbulencia en fluidos.

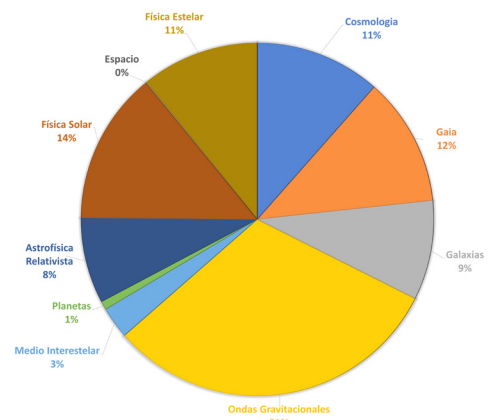


Figura 3. Distribución del tiempo de la RES por áreas (2015-2023).

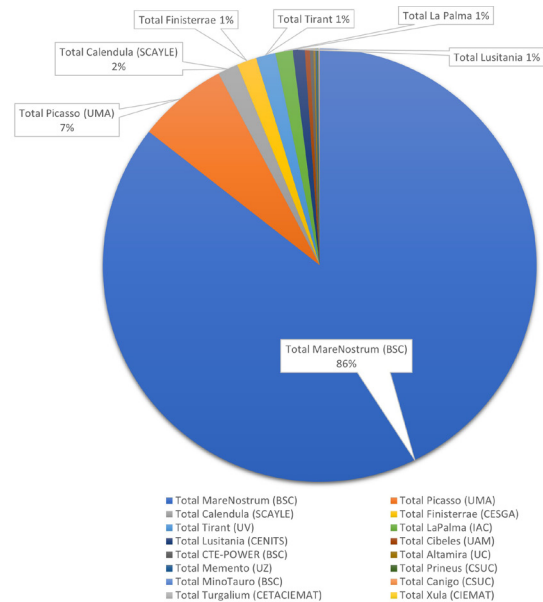


Figura 4. Distribución de tiempo de máquinas por áreas (2015-2023).

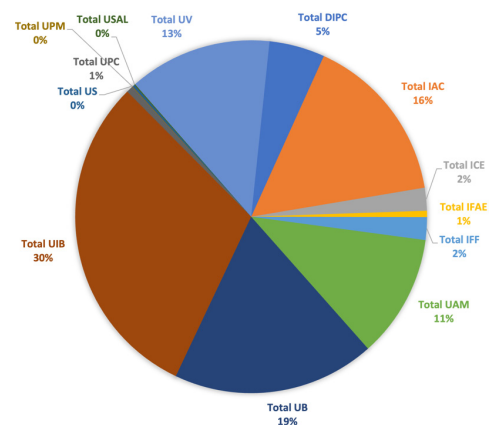


Figura 5. Distribución de tiempo por institución (2015-2023).

En total, dentro del periodo analizado, el número de horas computacionales asignados a los proyectos de Astrofísica fueron **565 millones**. La fracción de horas asignadas a los proyectos englobados en los temas científicos arriba indicados están representados en el diagrama de la Figura 3. Como se puede comprobar, el campo que domina en el consumo de recursos computacionales (más de un tercio) es el de la simulación de la generación de ondas gravitacionales a partir de fenómenos de coalescencia de agujeros negros y/o estrellas de neutrones. Es evidente que el descubrimiento de la primera señal de onda gravitacional por las colaboraciones LIGO /VIRGO, anunciado el 11 de septiembre de 2016, supuso una enorme necesidad de disponer de *templates* de producción de ondas gravitatorias durante el proceso de coalescencia de objetos estelares compactos. Para ello es necesario realizar simulaciones detalladas de este proceso descrito por las ecuaciones de Einstein de la Relatividad General. El grupo de Sacha Husa de la UIB ha venido desarrollando este tipo de simulaciones en la RES desde 2015 y los resultados obtenidos son fundamentales para la identificación del origen de los distintos eventos detectados por la colaboración LIGO.

En segundo lugar, con un 14% del tiempo asignado se encuentran los proyectos relacionados con simulaciones magneto-hidrodinámicas de plasmas en la atmósfera solar. En este tema, los grupos del IAC con Fernando Moreno Insertis, Elena Khomenko y Javier Trujillo han sido los más activos.

A continuación, hay una distribución más o menos similar de recursos de HPC entre Cosmología (12%), Gaia (12%), Física Estelar (11%) y Galaxias (9%).

En el tema de *Cosmología*, hemos recopilado todos aquellos proyectos que utilizan grandes recursos computacionales para simular la formación de estructuras en grandes volúmenes cosmológicos, principalmente, mediante métodos de N-cuerpos en diferentes modelos cosmológicos, aunque también hay simulaciones hidrodinámicas de formación de galaxias y de tipo zoom que se focalizan en regiones alrededor de cúmulos de galaxias. Una parte considerable de los resultados numéricos derivados de los proyectos RES en este campo han sido utilizados exhaustivamente por la comunidad internacional, tanto

«Gracias a la creación de estructuras como RES y PRACE, que agrupan recursos computacionales de diferente capacidad, los investigadores que requerimos de acceso a estos sistemas podemos elegir aquel que más de adecue al problema que queremos tratar.»

«La comunidad de Astrofísica/Cosmología computacional en España ha alcanzado una posición muy competitiva, gracias en parte al haber dispuesto de acceso a grandes recursos computacionales, tanto a nivel estatal, a través de la RES, como europeo por medio de los consorcios PRACE y EuroHPC.»

por investigadores y grupos individuales, como por las grandes colaboraciones que se encargan de desarrollar los cartografiados galácticos (BOSS, eBOSS, DES, DESI, Euclid, JPAS, etc). Los grupos más activos en este campo son los de la UAM (G. Yepes), UB (P. Fosalba), DIPC (R. Angulo) o IAA (F Prada).

En cuanto al tema *Gaia*, hay que hacer una mención especial, ya que los proyectos relacionados con este campo son exclusivamente dedicados al análisis de los datos provenientes del satélite *Gaia* por parte del grupo de la Universidad de Barcelona que colabora en esta misión espacial. Es sin duda el grupo más longevo en solicitar recursos de la RES desde su fundación, primero liderado por el recordado profesor Jordi Torra, y posteriormente por Xavier Luri. Sin duda, el enorme éxito de la misión *Gaia* y su excelente base de datos tienen mucho que agradecer a los recursos computacionales garantizados por la RES a este proyecto.

En cuanto al tema de *Física Estelar*, aquí hemos incluido aquellos proyectos que tienen que ver tanto con la simulación de procesos de formación estelar en nubes moleculares, como con el modelado numérico de los últimos estadios del colapso estelar previo a la explosión como supernovas. En este campo los grupos más activos con diferencia han sido los dirigidos por Paolo Padoan de la Universidad de Barcelona y Martin Obergaulinger y Miguel Ángel Aloy de la Universidad de Valencia.

En cuanto al tema de *Galaxias*, hay una variedad de grupos que han obtenido recursos para realizar simulaciones hidrodinámicas de formación de galaxias a nivel cosmológico, o bien focalizadas en la dinámica de la Vía Láctea y el Universo Local. Los grupos que más recursos RES han utilizado para estos temas son los del IAC (C. Dalla Vecchia, I. Alonso), Universidad de Sevilla (J. Oñorbe), UAM (D. Ceverino, W. Cui) y UB (C. Laporte).

El resto de temas, *Planetas* y *Medio Interestelar*, han tenido un uso más moderado de los recursos de RES, con menos del 5%. En el caso de Planetas, ha habido únicamente cuatro proyectos aprobados en el periodo analizado, todos ellos relacionados con simulaciones de la atmósfera de Júpiter (E. García-Melendo), Saturno (M. Soria) y extrasolares (D. Vigano). En Me-

dio Interestelar, los proyectos que han utilizado recursos de RES están directamente relacionados con la simulación mediante métodos de química cuántica de procesos que generan moléculas complejas en el medio interestelar. El investigador Octavio Ronceiro del CSIC es el investigador con mayor número de proyectos aprobados en esta línea.

Una conclusión que se desprende del análisis anterior por temas científicos es el gran uso de la infraestructura de supercomputación de la RES por parte de la comunidad astrofísica española en prácticamente todas áreas de investigación, desde las más grandes escalas cosmológicas hasta la modelización de las atmósferas planetarias. Esto es debido en gran medida a las posibilidades que ofrece la RES de acceder a diferente tipo de hardware en base a las necesidades computacionales del problema a resolver. Esto permite que proyectos con diferentes necesidades de HPC puedan ser ejecutados en el nodo RES más adaptado a sus necesidades, optimizando así los recursos asignados. No siempre el nodo con mayor capacidad de cálculo es el más útil para llevar a cabo un proyecto que requiere de un número moderado de procesadores o necesita mucha memoria por procesador. En general en HPC, la distribución de trabajos a ejecutar se comporta cómo las funciones de masa o luminosidad de galaxias o estrellas. Esto es, cuantos menos procesadores necesita una simulación, mayor número de ejecuciones se requieren. A mayor necesidad de recursos, menos número de ejecuciones. En este sentido, es necesario priorizar aquellos trabajos que requieren de acceso a los supercomputadores con mayor capacidad frente a los que necesitan menos recursos que pueden ser ejecutados en máquinas más pequeñas.

Para darnos una idea de cómo los investigadores utilizan el hardware gestionado por la RES, hemos hecho también un análisis sobre la utilización de las diferentes máquinas RES por parte de todos los proyectos del área de Astrofísica y Cosmología computacional.

Distribución de tiempo de proyectos de Astrofísica en nodos de la RES

En la Figura 4 representamos la fracción de tiempo por máquina de la RES. Como es evidente, *MareNostrum* del BSC, con el 86% del total, es la infraestructura más solicitada por los proyectos de Astrofísica dada su alta capacidad de computación y el que mayor número de

horas computacionales aporta al total de recursos de la RES. Por detrás están el resto de nodos con máquinas menos potentes, pero de arquitecturas similares a *MareNostrum*, donde la mayoría de los códigos numéricos utilizados pueden ser ejecutados de forma similar en modo de paralelismo híbrido (MPI + OpenMP). En este sentido es hacer notar la relativa poca ocupación que tiene el superordenador Finisferre del CESGA a pesar de ser único en la RES en cuanto a la alta capacidad de memoria RAM por núcleo que dispone en los nodos de cálculo, y que son de especial interés para muchos códigos de simulación, donde este factor es muy importante.

Distribución de tiempo de proyectos de Astrofísica por Institución

Por último, en la Figura 5 representamos la fracción de horas computacionales asignadas por la RES en función de la institución española donde el IP de cada proyecto tiene su afiliación.

Aquí se puede ver cómo hay una institución que destaca en cuanto al tiempo obtenido y es la Universidad de las Islas Baleares, con casi un tercio de los recursos. Esto es así por el hecho de que el grupo que se dedica a simulaciones de *templates* de ondas gravitacionales, liderado por Sasha Husa, está ubicado en esta Universidad. Como se ve en el gráfico de la Figura 4, este tema de investigación ha sido también el que más recursos ha utilizado, dada su importancia para las colaboraciones de LIGO/VIRGO. A continuación, le sigue la Universidad de Barcelona con un 19% de recursos, fundamentalmente debido a que allí está localizado el grupo que participa en el satélite Gaia y que se ha encargado en gran medida del análisis de datos del satélite en *MareNostrum*. Estas dos instituciones han usado prácticamente el 50% del total de horas asignadas por RES a proyectos de Astrofísica en el periodo analizado. El otro 50% restante se divide principalmente entre tres instituciones: El IAC con 16%, la Universidad de Valencia con el 13% y la Universidad Autónoma de Madrid con el 11%. El 10% restantes se lo distribuyen entre institutos públicos de investigación como el Donostia International Physics Center (5%) y varios institutos del CSIC: el Instituto de Ciencias Espaciales (ICE, 2%) el Instituto de Física Fundamental (IFF, 2%), Instituto de Física de Altas Energías (IFAE, 1%). En porcentajes menores del 1% aparecen también las Universidades Politécnicas de Madrid y de Cataluña, así como la Universidad de Sevilla y la de Salamanca.

Una conclusión que se desprende de este gráfico es la gran diferencia entre los recursos utilizados por grupos de investigación en Universidades (~75%) frente a aquellos que se encuentran realizando su labor en institutos de investigación no universitarios (~25%). Esto nos lleva a concluir que la Astrofísica computacional en España se ha venido realizando sobre todo en departamentos universitarios donde los recursos financieros son mucho menores que en los institutos de investigación con lo que se ha priorizado la simulación numérica, mucho menos costosa si se dispone de acceso a las infraestructuras de supercomputación como RES.

Por último, también es interesante fijarse en los temas científicos prioritarios en aquellas instituciones con más de un 10% de cuota de tiempo en RES. Como ya hemos mencionado, la UIB se focaliza casi exclusivamente en el tema de Ondas gravitacionales, mientras que la UB esta fundamentalmente involucrada en el análisis de datos de Gaia. Así mismo, en el IAC domina históricamente los proyectos de Física Solar. Mientras que la UAM tiene también una larga tradición en el tema de las simulaciones cosmológicas y de formación de galaxias. Por último, la Universidad de Valencia domina claramente el tema de Astrofísica Relativista y Física Estelar. En cuanto al resto de Universidades y Centros de investigación, los temas científicos están más dispersos.

PRACE COMO MOTOR DE LA SUPERCOMPUTACIÓN EN EUROPA

El "*Partnership for Advanced Computing in Europe*" (PRACE, <https://prace-ri.eu>) es una organización europea muy similar a RES tanto en sus objetivos como en la fecha de creación. La motivación detrás de esta iniciativa fue la de promover el acceso a la supercomputación de alto rendimiento para la investigación científica y la innovación tecnológica en Europa. Para ellos, los grandes centros de supercomputación europeos decidieron aportar una fracción de sus recursos de computación al consorcio PRACE para que fueran distribuidos de forma centralizada mediante un Comité de Acceso único que se encargaría de evaluar los proyectos enviados por investigadores de cualquier lugar del mundo.

A continuación, se presenta una descripción de sus orígenes, evolución e impacto hasta la fecha.

Orígenes y fundación del consorcio PRACE

La idea de PRACE surgió en 2006 como una respuesta al creciente reconocimiento de la importancia de la supercomputación en la ciencia y la innovación. El proyecto se inició bajo la dirección de la Comisión Europea, que identificó la necesidad de una infraestructura de supercomputación paneuropea. En 2008, PRACE se estableció como un proyecto financiado por la Comisión Europea. El consorcio original incluía a cuatro países: Francia, Alemania, Italia y España, que alojaban los mayores superordenadores europeos gestionados por centros de cálculo líderes en HPC como GENCI (Francia), GCS (Alemania), CINECA (Italia) y BSC (España). En 2010, se estableció formalmente la "*PRACE Research Infrastructure*" como una entidad legal sin fines de lucro con sede en Bruselas. Se creó una red de supercomputadoras de alto rendimiento, permitiendo a investigadores europeos acceder a recursos punteros. Durante esta etapa, se desarrollaron además programas de formación y se establecieron vínculos con la industria. PRACE amplió la lista de miembros hasta tener más de 25 países europeos dentro del consorcio, con un conjunto variado de infraestructuras de supercomputación tanto a nivel de arquitecturas como de potencia de cálculo, dividiéndose entre máquinas tipo Tier-0 y Tier-1.

En 2017, PRACE lanzó su segunda fase, que supuso una mejora significativa en la infraestructura y los servicios ofrecidos. El presupuesto de PRACE 2 ascendió a 400 millones de euros, y se fortaleció la colaboración con socios industriales y académicos.

Desde 2010, PRACE ha proporcionado acceso a superordenadores a miles de proyectos científicos, cubriendo campos como la física, la biología, la ingeniería y las ciencias medioambientales. Los proyectos apoyados por PRACE han contribuido significativamente al desarrollo de nuevos medicamentos, la comprensión de los cambios climáticos y el diseño de materiales innovadores. PRACE ha organizado cientos de talleres y programas de formación para investigadores, promoviendo la excelencia en el uso de la supercomputación. Así mismo, ha fomentado la cooperación con empresas, permitiendo que la industria europea aproveche las capacidades de la supercomputación.

En 2021 PRACE entró en su tercera fase con un enfoque renovado en la preparación para la llegada de la computación de exaescala. Hay un esfuerzo concertado por mejorar la eficiencia energética de los centros de supercomputación y aumentar la inclusión de las empresas pequeñas y medianas. PRACE forma parte integral de la estrategia europea para la supercomputación en la próxima década, en conjunto con la iniciativa EuroHPC de la que hablaremos más adelante.

Se espera que la organización siga desempeñando un papel fundamental en el desarrollo de sistemas de supercomputación exaescala y la promoción de la soberanía tecnológica europea.

En la Figura 6 representamos los superordenadores tipo Tier-0 que estaban integrados en PRACE en la última época, previa a la creación de EuroHPC. La potencia de cálculo acumulada es de aproximadamente 196 PFLOPS. Es decir, no llega ni al 20% de

Figura 6. Superordenadores tipo Tier-0 integrados en PRACE durante la última época previa a la creación de EuroHPC.



Country	Site (Partner)	HPC System	RPeak (TF/s)
France	CEA (GENCI)	Joliot-Curie KNL	2 300
France	CEA (GENCI)	Joliot-Curie Rome	12 000
France	CEA (GENCI)	Joliot-Curie SKL	6 600
Germany	JUELICH (FZJ)	JUWELS Booster	71 000
Germany	JUELICH (FZJ)	JUWELS Cluster	9 900
Germany	LRZ (GCS)	SuperMUC-NG	26 900
Germany	HLRS (GCS)	HAWK	25 200
Italy	CINECA (CINECA)	Marconi100	20 000
Spain	BSC (BSC)	MareNostrum 4	11 150
Switzerland	CSCS (ETH)	Piz Daint	15 988.0

un EXAFLOP, y muy similar a la potencia pico que va a tener *MareNostrum-5*. No obstante, comparado con los recursos a nivel nacional proporcionados por RES, es casi un factor 10 superior. En cuanto a arquitecturas de los superordenadores, salvo el caso del ordenador CRAY *Piz Daint*, en Suiza, que tiene una arquitectura híbrida con CPU multinúcleo + GPUs como aceleradores de cálculo, todos los demás son muy homogéneos con sistema multinúcleo AMD o Intel e interconectados por una red de baja latencia, muy similar al resto de computadoras de menor capacidad a nivel nacional (por ejemplo, RES en España). Esto quiere decir que para la mayoría de los grupos de investigación que usan códigos numéricos basados en paralelismo híbrido OpenMP + MPI, el acceso a estos sistemas Tier-0 de PRACE sería relativamente sencillo. El único requisito es por tanto el interés del caso científico del proyecto y la necesidad de disponer de una capacidad de computación superior a la que podrían disponer a nivel nacional.

En resumen, gracias a la creación de estructuras como RES y PRACE, que agrupan recursos computacionales de diferente capacidad, los investigadores que requerimos de acceso a estos sistemas podemos elegir aquel que más de adecue al problema científico que queremos tratar. En general, la puesta a punto del código se puede realizar en sistemas informáticos más modestos, accesibles a nivel del grupo de investigación, departamental, universitario o centro de investigación. Este tipo de infraestructuras serían de tipo Tier-2, por debajo del PFLOP de potencia, pero con un acceso de tiempo mucho más flexible. Una vez optimizado y testado el código de simulación, se puede pedir tiempo en régimen competitivo, bien a recursos a nivel nacional con máquinas tipo Tier-1, del orden del PFLOPS, pero con mayores restricciones de tiempos de ejecución. Si la simulación lo necesitara, todavía se podría acceder a las infraestructuras de tipo Tier-0, de decenas o centenas de PFLOPS a nivel europeo gracias a PRACE. Obviamente, cuanto mayores son los recursos que se necesitan, mayor será la competencia sobre ellos y más complicado será conseguir que se aprueben las propuestas de tiempo de cálculo.

Puesto que en este artículo estamos interesados en ofrecer una visión de la astrofísica computacional española, no queremos dejar pasar la ocasión

de realizar también un análisis sobre el grado de utilización de PRACE por parte de los investigadores españoles.

Como en el caso de RES, hemos recopilado todos los proyectos liderados por un investigador que tenía afiliación en una institución española y que han conseguido tiempo de computación en PRACE. Nos hemos centrado exclusivamente en analizar cómo se ha distribuido el tiempo total asignado en función del área científica, la institución de origen del IP y el tipo de máquina donde se ha asignado el tiempo de cálculo. Queremos dejar claro que hay muchas más propuestas en PRACE donde investigadores de centros españoles actúan como colaboradores. Hemos decidido no tomarlas en cuenta en nuestro análisis para poder identificar con claridad las áreas donde los investigadores españoles han tenido una posición de liderazgo.

El resultado de nuestro análisis se puede ver en las Figuras 7, 8 y 9. En la Figura 7 representamos los temas científicos de proyectos que han conseguido tiempo en PRACE. Como puede verse, representan una distribución bastante parecida a lo que vimos en RES, salvo el caso especial de Gaia que no es una actividad específica de HPC. Al igual que en RES, son los proyectos de ondas gravitacionales de la UIB los que se llevan la fracción mayor del tiempo en PRACE. A continuación, le siguen los temas relacionados con física estelar, cosmología, Astrofísica Relativista y Física Solar. Si tenemos en cuenta lo dicho anteriormente en el análisis de tiempos concedidos en RES, estos temas científicos están directamente relacionados con los mismos grupos de investigación que han liderado también las propuestas a nivel nacional.

En la Figura 8 se representa la distribución de tiempos por institución de referencia del IP. La UIB es la que mayor fracción de tiempo ha conseguido, gracias a los proyectos de Sasha Husa, en relación con las simulaciones de radiación de ondas gravitacionales durante la coalescencia de objetos estelares densos, seguida en este caso por la UB, con Paolo Padoan liderando proyectos para simulaciones de formación estelar en el ISM. A continuación, se encuentra la UAM, donde uno de los autores de este informe ha liderado varios proyectos de simulaciones cosmológicas en grandes volúmenes cosmológicos y de partículas, tales como *Multidark*, (<https://www.cosmosim.org>), *The Curie*

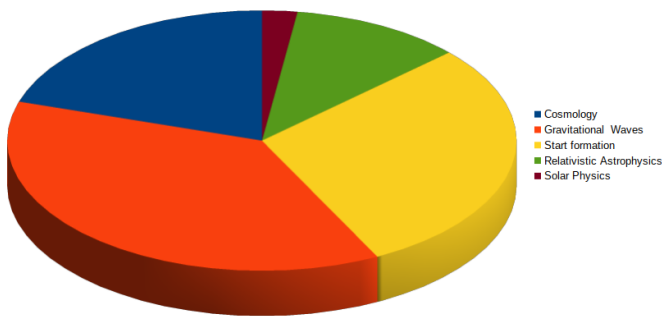


Figura 7. Distribución de tiempo de PRACE por área científica.

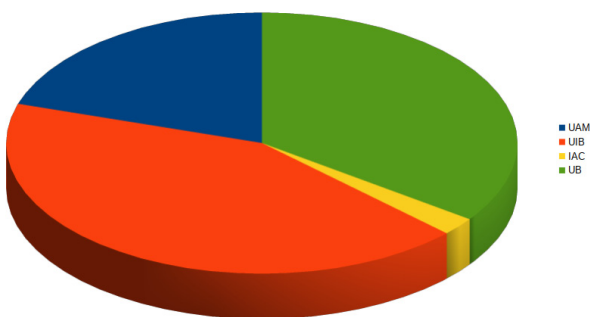


Figura 8. Distribución de tiempo concedido en PRACE por institución.

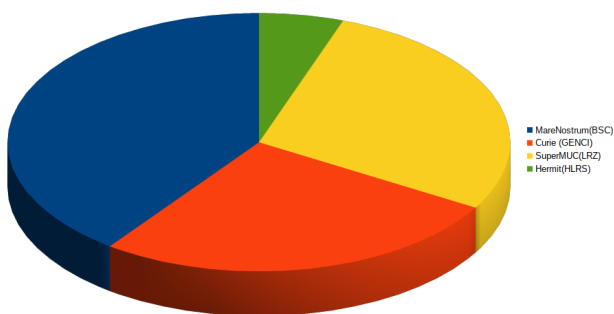


Figura 9. Distribución de tiempo concedido en máquinas PRACE.

Universe (<https://curiehz.ft.uam.es/>) o *UNITSIMS* (<https://unitsims.ft.uam.es/>). Finalmente, también el IAC ha conseguido tiempo de PRACE para proyectos de Física Solar liderados por Elena Khomenko.

En la Figura 8 representamos la distribución de tiempo PRACE por máquina. Como es lógico imaginar, la mayoría los proyectos liderados por un IP de España fueron asignados a *MareNostrum*, ya que esos grupos tenían experiencia previa en esta máquina gracias a sus proyectos de RES. Sin embargo, hay dos máquinas similares a *MareNostrum* que también han sido ampliamente utilizadas por los grupos españoles, con porcentajes similares, *SuperMUC* en LRZ (Alemania) y *Curie* en GENCI (Francia).

EUROHPC, EL CAMINO EUROPEO HACIA LA EXAESCALA

EuroHPC (<https://eurohpc-ju.europa.eu> *European High-Performance Computing Joint Undertaking*) es una iniciativa impulsada por la Unión Europea y sus Estados miembros con el objetivo de situar a Europa en la vanguardia de la computación de alto rendimiento (HPC) a nivel mundial. Este consorcio estratégico, establecido en 2018, se enfoca en desarrollar, adquirir y operar sistemas de supercomputación exaescala, es decir con una potencia de cálculo superior al EXAFLOP = 1000 PFLOPS.

Entre sus principales metas está la construcción y puesta en funcionamiento de varias supercomputadoras de exaescala antes de mediados de la década de 2020, así como la promoción de un ecosistema europeo sólido de tecnologías y servicios relacionados con la HPC. Esto implica no solo adquirir y operar estas potentes máquinas, sino también fomentar el desarrollo de hardware y software europeo, así como capacitar a una nueva generación de expertos en supercomputación.

Al igual que PRACE, EuroHPC trabaja bajo un enfoque colaborativo, involucrando a gobiernos nacionales, instituciones académicas, centros de investigación y empresas tecnológicas. A través de financiamiento conjunto y estrategias compartidas, este consorcio busca optimizar los recursos y acelerar el progreso en el campo de la HPC. Además, promueve la accesibilidad y el uso eficiente de estas infraestructuras entre investigadores, científicos y empresas de toda Europa, facilitando el intercambio de conocimientos y la solución de problemas trans fronterizos.



Figura 10. Infraestructuras actualmente gestionadas por EuroHPC.

Country	Site (Partner)	HPC System	RPeak (PFlop/s)
Finland	CSC	LUMI	550
Italy	CINECA	LEONARDO	324.4
Spain	BSC-CNS	MareNostrum5	314
Slovenia	IZUM	VEGA	10
Luxemburg	LuxProvide	MELUXINA	18.3
Czech Rep.	IT4	KAROLINA	15.7
Portugal	MACC	DEUCALION	10
Bulgaria	SofiaTech Park	DISCOVER	5.94
Total			1248.34

En resumen, EuroHPC representa un compromiso firme de Europa hacia la excelencia en computación de alto rendimiento, buscando no solo mantenerse al día con las principales potencias tecnológicas globales (EE.UU, China y Japón), sino también liderar el avance de soluciones basadas en supercomputación para los desafíos más apremiantes de nuestra época. A través de su trabajo, EuroHPC contribuye al fortalecimiento de la investigación científica, el desarrollo industrial y la competitividad de Europa en el escenario global.

En julio de 2019, la *EuroHPC Joint Undertaking* seleccionó al Barcelona Supercomputing Center como una de las entidades para albergar un superordenador pre-exaescala de la red de supercomputadores de alta capacidad promovida por la Comisión Europea. *MareNostrum-5* está cofinanciado por EuroHPC y el consorcio español

formado por el Ministerio de Universidades, Ciencia e Innovación, la Generalitat de Catalunya y la UPC y ha entrado recientemente en producción para dar servicio, tanto a proyectos de la RES como de EuroHPC y tiene una capacidad total de cálculo de más de 300 PFLOPS con dos particiones, una con nodos de Intel CPU multinúcleo, igual que en versiones anteriores de *MareNostrum*, y otra con aceleradores gráficos NVIDIA de última generación. Es decir, *MareNostrum-5* tendrá más potencia que toda la infraestructura gestionada por PRACE en el pasado.

En la Figura 10 se pueden ver la infraestructura de HPC gestionada por EuroHPC en la actualidad. Como se puede apreciar, hay tres máquinas que sobresalen por encima de las demás: *LUMI* (550 PFLOPS), *LEONARDO* (324 PFLOPS) y *MareNostrum-5* (313 PFLOPS). Estos tres sistemas son los

pre-exaescala co-financiados por la Unión Europea. En una segunda fase, habrá dos superordenadores con potencia de exaescala instalados en Alemania y Francia. El resto de máquinas de EuroHPC son de potencia similar a las máquinas PRACE de años anteriores. Sin embargo, la característica principal que comparten todas ellas es la de tener una arquitectura mixta, con nodos basados en CPU clásica y nodos con CPU + GPU. Sería imposible poder aumentar la potencia de cálculo a las escalas de EXAFLOPS sin recurrir a los aceleradores gráficos. De hecho, se puede ver esta tendencia si se revisa la famosa lista Top500 (<https://www.top500.org>) que compila los superordenadores de mayor potencia en el mundo. En ella, hay solo dos superordenadores instalados en EE.UU. que traspasan la barrera del EXAFLOP. *Frontier* tiene una arquitectura combinada de una CPU AMD Epyc + 4 GPUs AMD Indistinct por nodo. *Aurora*, es el otro superordenador que combina CPU y GPUs, en este caso fabricadas ambas por Intel.

DESAFÍOS Y RETOS DE LA ASTROFÍSICA COMPUTACIONAL EN LOS PRÓXIMOS AÑOS

La Astrofísica computacional es una ciencia cuyo desarrollo va parejo al incremento de la potencia de cálculo disponible en cada momento de la historia. A mayor potencia de procesamiento, mejores datos simulados se pueden obtener, ya que los objetos a simular pueden ser modelizados con un mayor número de elementos computacionales, bien sean partículas o mallas. Esto lleva asociado a su vez un desarrollo teórico para modelizar los procesos físicos que ocurren a escalas que antes no se resolvían. Así mismo, es necesario también realizar un enorme esfuerzo para diseñar los códigos numéricos que van a resolver las ecuaciones que describen la física de los objetos de estudio de forma que puedan utilizar la potencia máxima posible de cálculo de la infraestructura donde se van a ejecutar.

Esto ha implicado el cambio de paradigma a nivel de programación varias veces en el pasado. Los pioneros en la simulación numérica en Astrofísica todavía recordarán cuando los códigos numéricos se programaban en el lenguaje FORTRAN 66, donde no existía ni siquiera el concepto de bucle de operaciones. En esa época la velocidad del código era directamente proporcional a la velocidad del (único) procesador donde se ejecutaba. Como se cumplía eficientemen-

te con la Ley de Moore (la potencia de los procesadores se doblaba cada 18 meses aproximadamente), el mismo código podía ser utilizado durante años aplicándolo a problemas cada vez más complejos. Cuando la tecnología cambió y se introdujeron los chips con múltiples líneas de ejecución simultánea (multinúcleo), se tuvieron que modificar los códigos para adaptarlos a esta tecnología. El proceso no fue excesivamente complicado, y bastaba con evitar las dependencias de las operaciones, fundamentalmente dentro de los bucles, para que el compilador se encargara de paralelizarlos y que se ejecutaran las instrucciones de forma simultánea en varios núcleos a la vez. Simplemente había que aprender varias directivas de compilación del estándar OpenMP, que había que poner como comentarios en los bucles a paralelizar. En general, un astrofísico, generalmente estudiante de doctorado, o postdoc, podía aprenderlo con relativa facilidad y optimizar los códigos para sacarles un gran rendimiento en los sistemas multinúcleo.

Posteriormente llegaron los superordenadores basados en la unión de múltiples nodos multinúcleo interconectados por una red de alta velocidad, donde el acceso a los datos en memoria ya no está disponible por todos los procesadores, sino que había que distribuir los datos entre la memoria de los diferentes nodos y tener que enviar y recibir los datos que no estaban locales mediante mensajes a otros nodos. La ventaja es que se podían usar una gran cantidad de procesadores en un mismo problema, pero a costa de tener que reescribir los códigos de simulación, utilizando librerías como MPI que facilitaban la labor del tráfico de datos entre los nodos.

En la actualidad, la mayoría de los códigos utilizados en la simulación astrofísica en los diferentes temas de los que hemos hablado en este artículo están bastante bien adaptados para ser eficientes en arquitecturas de múltiples nodos donde cada nodo tiene uno o varios chips multinúcleo y una memoria por núcleo relativamente limitada, del orden de los 1-2 Gbytes.

Ahora nos viene el siguiente cambio de paradigma. Como hemos comentado anteriormente, las infraestructuras de HPC del presente y futuro inmediato van a añadir un elemento adicional, los aceleradores gráficos. Si queremos seguir progresando en

la calidad y verosimilitud de nuestras simulaciones numéricas, es necesario realizar un gran esfuerzo para adaptar de nuevo los códigos más habitualmente utilizados en Astrofísica a este nuevo paradigma. Sin embargo, la situación no es sencilla. La complejidad técnica se va haciendo cada vez mayor y se necesita la ayuda de personal altamente cualificado con formación en ciencia computacional para ayudar en este proceso. Ya no es tan sencillo el que un estudiante de doctorado con formación astrofísica sea capaz de llevar a cabo esta tarea. Nos enfrentamos a un desafío enorme que va a requerir de la inversión en recursos humanos especializados, no fáciles de conseguir, dado la alta competitividad actual para contratar a estos especialistas a raíz de la explosión de la inteligencia artificial en todos los ámbitos de nuestra sociedad. En este sentido, la Unión Europea a través de EuroHPC es consciente de este problema que afecta a todas las ramas de ciencia que tienen a la supercomputación como una herramienta fundamental en su desarrollo.

A principios del año 2023 se crearon 10 Centros de Excelencia en Europa (<https://www.hpccoe.eu/about/>) para dar apoyo a diferentes disciplinas científicas en la transición de sus códigos de simulación hacia la era Exascale. En lo que concierne al ámbito de Astrofísica se creó el Centro de Excelencia *SPACE* (*Scalable Parallel Astrophysical Codes for Exascale*, <https://www.space-coe.eu>), localizado en Irlanda con un presupuesto de 8 millones de euros por un periodo de 48 meses. Varios de los códigos numéricos más usados en la comunidad de Astrofísica/Cosmología computacional (GADGET, RAMSES, PLUTO, IPI3D, BHAC, FIL), están recibiendo apoyo técnico de *SPACE* para adaptarlos a la arquitectura de los superordenadores gestionados por EuroHPC. Uno de los requisitos exigidos para que estos códigos tengan apoyo por parte de *SPACE* es que estén disponibles públicamente para toda la comunidad.

CONCLUSIONES

En este artículo hemos intentado presentar una visión cuantitativa y cualitativa a nivel español de una de las áreas de la Astrofísica más importantes en la actualidad, que tiene que ver con la producción de simulaciones numéricas que actúan como laboratorios virtuales donde comparar y comprender la inmensa colección de datos observacionales que

obtenemos del Universo en todo el rango espectral electromagnético y de nuevos mensajeros como las ondas gravitacionales.

La conclusión que derivamos de nuestro análisis es que la comunidad de Astrofísica/Cosmología computacional en España ha alcanzado una posición muy competitiva, gracias en parte al haber dispuesto de acceso a grandes recursos computacionales, tanto a nivel estatal, a través de la RES, como europeo por medio de los consorcios PRACE y EuroHPC.

A nivel estatal, la constitución de la red de nodos de la RES ha impulsado la Astrofísica computacional en distintos centros de investigación y universidades. En algunos centros (sobre todo departamentos universitarios), la línea de investigación teórica/numérica tenía una larga tradición y se ha aprovechado de la infraestructura proporcionada por la RES para el fortalecimiento de esos grupos. Sin embargo, otros centros que no tenían esa vocación han decidido crear nuevos grupos de investigación dedicados a la Astrofísica computacional gracias a las posibilidades que ofrecía la RES, tanto de acceso a recursos locales como estatales. Disponer de instalaciones de supercomputación locales facilita el desarrollo de códigos y el diseño de simulaciones que posteriormente se ejecutan en los grandes superordenadores europeos.

Para el futuro inmediato, uno de los grandes retos que se plantea la comunidad para continuar con el desarrollo de simulaciones más realistas es la adaptación de los códigos numéricos actuales a las nuevas infraestructuras de supercomputación exaescala basadas en hardware híbrido con CPU + aceleradores GPUs. Por otro lado, el desarrollo exponencial de la inteligencia artificial (IA) también va a suponer una enorme revolución dentro del campo de la simulación astrofísica. Las ecuaciones que describen los fenómenos físicos son habitualmente resueltas mediante métodos numéricos muy complejos que cuestan mucho tiempo de cálculo. Los algoritmos de IA pueden ser entrenados a partir de resultados de simulaciones para que aprendan las relaciones entre las diferentes variables del sistema a simular. Estos algoritmos entrenados pueden actuar como emuladores de las ecuaciones físicas, acelerando enormemente los tiempos de cálculo y

permitiendo incrementar la resolución espacio-temporal y la complejidad de los modelos físicos de las simulaciones. Esto no es una predicción de futuro. Existe ya mucha literatura al respecto donde este tipo de aplicaciones en el ámbito de la Astrofísica es ya una realidad, pero un análisis más en profundidad de la revolución que la IA va a representar tanto en los campos de la simulación como en el análisis de los datos astronómicos merecería de otro artículo dedicado exclusivamente a este tema.

Para finalizar este artículo, y a modo de resumen, presentamos en la Figura 11 una visualización de resultados de algunas de las simulaciones numéricas realizadas con proyectos de RES y PRACE liderados por grupos españoles, que abarcan objetos en un rango de escalas *astronómicas*: desde kilómetros de la superficie solar hasta escalas de mega parsecs de una simulación cosmológica.

Figura 11. Resultados de simulaciones numéricas con proyectos RES y PRACE abarcando diferentes escalas, de megaparsecs (arriba) hasta kilómetros (abajo).

