

10 AÑOS

Supercomputación Castilla y León
Anuario · Informe de actividad **2018**

Supercomputación Castilla y León

Anuario · Informe de actividad

2018



universidad
de León



Junta de
Castilla y León

Anuario · Informe de actividad · 2018
Supercomputación de Castilla y León, SCAYLE
Depósito Legal LE-244-2016

Fundación Centro de Supercomputación de Castilla y León
Edificio CRAI-TIC, Campus de Vegazana s/n • Universidad de León • 24071 León (España)
Teléfono: (+34) 987 293 160
www.scayle.es

Otras marcas y nombres son propiedad de sus respectivos propietarios.
Other brands and name are property of their respective owners.

Caja España – Obra social
Premio Medio Ambiente Caja España 2009



I Edición enerTIC Awards 2013
"Innovación y Reconocimiento a las mejores prácticas"
Proyecto Ganador Categoría: Smart IT Infrastructure



El Mundo Diario de Castilla y León
Premios Innovadores 2016
Premio Innovador al Mejor Proyecto de León



Índice de contenidos

Presentación Dirección General	pág.7
Composición del Patronato	pág.9
Consejo Asesor	pág.11
SCAYLE	
Descripción	pág.15
Descripción técnica	pág.17
Cartera de servicios	pág.19
SCAYLE, Supercomputación y RES	pág.21
SCAYLE: ICTS distribuida "Red Española de Supercomputación"	pág.23
RedCAYLE	pág.25
Servicios TIC para las administraciones públicas	pág.27
Proyectos I+D+i	
Colaborador	pág.31
Colaborador en proyectos I+D+i a través de la RES	pág.63
Publicaciones científicas	pág.65
Formación	
Cursos	pág.75
Otras colaboraciones en materia de formación de SCAYLE	pág.81
Datos de SCAYLE	
10º aniversario SCAYLE	pág.85
Datos técnicos	pág.87
Datos económicos	pág.89
Protocolos y convenios de colaboración	pág.91
Visitas a SCAYLE	pág.93
Actividades de promoción, difusión y comunicación	pág.95

presentación

Dirección General

El año 2018 ha sido un año de celebración en SCAYLE, el de nuestro décimo aniversario, el del paso a doble dígito. Pueden parecer pocos años comparados con los de la ciudad que acoge al centro, o con los de las de las instituciones a las que sirve, por ejemplo, los 800 años que celebró también la Universidad de Salamanca en 2018; pero diez años son muchos en el campo de la informática, que ha visto grandes cambios en ese periodo.

Un ejemplo de esos cambios ha sido Caléndula, el cluster de cálculo científico de SCAYLE. Durante estos diez años la potencia de cálculo de Caléndula ha pasado de los 29 TeraFLOPs de su inauguración en 2009, a los 226 de la última renovación tecnológica realizada en 2018 con la incorporación de más 2.700 nuevos cores y 8 tarjetas GPU de altas prestaciones. De igual forma, la capacidad de almacenamiento ha alcanzado el Petabyte y la granja de virtualización albergaba a finales de 2018 más de 500 servidores de diversas instituciones.

Pero no sólo se han producido mejoras tecnológicas durante estos diez años, también han aumentado las competencias de SCAYLE más allá del cálculo científico. Merece la pena destacar el encargo de gestión de la Red Regional de Ciencia y Tecnología (RedCAYLE). Durante 2018 se han seguido incorporando instituciones a ella, conectando ya a la mayoría de los centros de I+D de la Comunidad de Castilla y León. Una incorporación especialmente relevante a RedCAYLE durante 2018 ha sido la de la red de colegios de la Junta de Castilla y León, que supondrá a medio plazo el mayor número de usuarios de RedCAYLE, cuando finalice el proyecto de "Escuelas Conectadas" que dotará de accesos de banda ancha a todos los centros escolares de la comunidad.

El décimo aniversario sirvió también para actualizar la imagen de SCAYLE, con un nuevo logo que intenta representar los objetivos y valores de SCAYLE: el apoyo al crecimiento a partir del apoyo a la investigación representado por las dos barras verticales, que corresponden con la vista lateral de Caléndula con un colorido inspirado en los mapas de calor que genera la máquina.

El año 2018 también se ha producido un reconocimiento muy relevante para SCAYLE: el Consejo de Política Científica, Tecnológica y de Innovación, presidido por el ministro de Ciencia, Innovación y Universidades, acordó el de 6 noviembre incluir a Caléndula en el mapa de Infraestructuras Científico Técnicas Singulares como parte de la Red Española de Supercomputación. Ha sido la culminación de un largo proceso que avala la calidad de los servicios que presta SCAYLE a la comunidad científica nacional.

Durante el 2018 se han empezado a producir los primeros resultados en la apuesta por el incremento en la participación en proyectos de investigación y desarrollo. Se han presentado solicitudes a diferentes programas europeos, con distintos socios de la comunidad autónoma. Por ejemplo, arrancó el proyecto Cross-Forest, en colaboración con la Escuela de Ingenieros de Montes de Palencia (Universidad de Valladolid), el grupo Tragsa y otros socios internacionales, en la que SCAYLE espera colaborar en la mejora de la silvicultura a través de la simulación del crecimiento de nuestros bosques. Es un buen ejemplo de las aplicaciones de la supercomputación a campos que a priori parecen poco relacionados. Igualmente colaboramos con grupos de las distintas universidades como son el grupo del Departamento de

Matemática Aplicada de la Universidad de Salamanca, con la Universidad de Burgos en el diseño de nuevos materiales, o con el grupo de mejora genética animal de la Universidad de León, por citar algunos de entre las múltiples áreas en las que trabajamos.

Finalmente, aprovechando el décimo aniversario, se han intensificado las tareas de difusión de las actividades de SCAYLE. Así, se realizaron presentaciones en diversas ciudades de Castilla y León para exponer las oportunidades que ofrece la supercomputación a los grupos de investigación y a las empresas.

En resumen, creo que 2018 ha sido un gran año para SCAYLE, como muestra este anuario y estoy seguro que el empuje de todos los que trabajamos en la Fundación de Supercomputación de Castilla y León hará que 2019 sea aún mejor.

Vicente Matellán Olivera
Director General de SCAYLE

composición del **patronato**

El Patronato de la Fundación del Centro de Supercomputación de Castilla y León está compuesto por:

Presidente

D. Juan Carlos Suárez-Quñones y Fernández

*Consejero de Fomento y Medio Ambiente
Junta de Castilla y León*

Vicepresidente

D. Juan Francisco García Marín

Rector Magnífico de la Universidad de León

Vocales

D. Juan Casado Canales

*Secretario General de la Consejería de Educación
Junta de Castilla y León*

D^a. M^a Victoria Seco Fernandez

*Directora General de Telecomunicaciones, Consejería de Fomento y Medio Ambiente
Junta de Castilla y León*

D^a. Susana García Dacal

*Directora General de Infraestructuras y Tecnologías de la Información de la Gerencia Regional de Salud, Consejería de de Sanidad
Junta de Castilla y León*

D. Ricardo González Mantero

*Director General de Energía y Minas, Consejería de Economía y Hacienda
Junta de Castilla y León*

D. Carlos G. Polanco de la Puente

*Vicerrector de Investigación
Universidad de León*

D. Adolfo Rodríguez de Soto

*Vicerrector de Gestión de Recursos e Infraestructuras
Universidad de León*

D. Luis Panizo Alonso

*Profesor Titular de Universidad
Universidad de León*

Secretario

D. Jesús Ignacio Sanz Valdivieso

*Coordinador de Servicios de la Dirección General de Telecomunicaciones
Consejería de Fomento y Medio Ambiente
Junta de Castilla y León*

Vicesecretario

D. Pablo Carlos Vicente Villafáfila

*Jefe del Servicio de Medios Audiovisuales y Régimen Jurídico de la Dirección
General de Telecomunicaciones
Consejería de Fomento y Medio Ambiente
Junta de Castilla y León*

El Consejo Asesor de la Fundación del Centro de Supercomputación de Castilla y León está compuesto por:

Presidente

D. Juan Francisco García Marín

*Rector Magnífico de la Universidad de León
Vicepresidente del Patronato de la Fundación Centro de Supercomputación de Castilla y León*

Vocales

D. Manuel Pérez Mateos

Rector Magnífico de la Universidad de Burgos

D. Ricardo Rivero Ortega

Rector Magnífico de la Universidad de Salamanca

D. Antonio Largo Cabrerizo

Rector Magnífico de la Universidad de Valladolid

Secretario

D. Jesús Ignacio Sanz Valdivieso

*Coordinador de Servicios de la Dirección General de Telecomunicaciones
Consejería de Fomento y Medio Ambiente
Junta de Castilla y León*



universidad
de León



UNIVERSIDAD
DE BURGOS



UNIVERSIDAD
DE SALAMANCA
CAMPUS DE EXCELENCIA INTERNACIONAL



Universidad de Valladolid

SCAYLE

Descripción

Descripción técnica

Cartera de servicios

SCAYLE, Supercomputación y RES

SCAYLE: ICTS distribuida "Red Española de Supercomputación"

RedCAYLE

Servicios TIC para las administraciones públicas

descripción

El **Centro de Supercomputación de Castilla y León (SCAYLE)** es una Organización no lucrativa, perteneciente al Sector Público de Castilla y León, que tiene como actividad principal, la mejora de las tareas de investigación en las Universidades, los centros de I+D+i y las empresas de Castilla y León, promoviendo y desarrollando acciones de innovación en el mundo de la Sociedad del Conocimiento, el área del cálculo intensivo, las comunicaciones y los servicios avanzados, contribuyendo mediante el perfeccionamiento tecnológico al desarrollo económico de la Comunidad y a la mejora de la competitividad de las empresas.

Valores

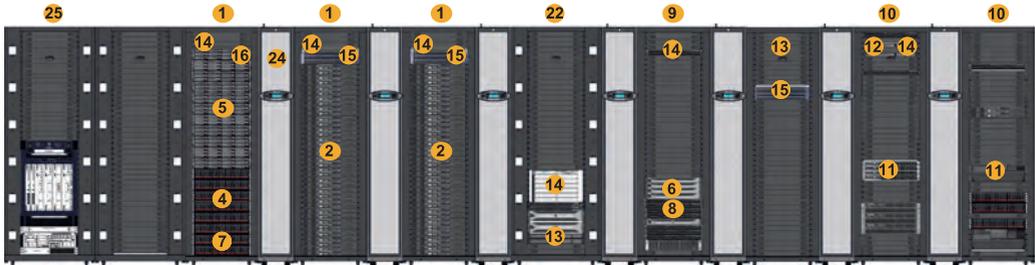
- **Compromiso social** contribuyendo mediante el perfeccionamiento tecnológico al desarrollo económico de la Comunidad y a la mejora de la competitividad de las empresas.
- Desarrollo de líneas estratégicas fundamentadas en el **establecimiento de redes de colaboración permanente** que garanticen el uso eficiente de recursos tecnológicos compartidos y el diseño de nuevos modelos organizativos.
- **Actualización constante** contribuyendo de forma eficiente al desarrollo de la ciencia y a la transferencia de la tecnología.
- **Adopción del paradigma ecológico** formalizado con la integración de la gestión ambiental en el sistema de gestión, principalmente basado en la dinamización de modelos de gestión de recursos tecnológicos con criterios de máxima eficiencia energética.
- **Ciberseguridad** garantizando la integridad y confidencialidad de la información de las instituciones a las que se prestan servicios.
- **Participación destacada en la construcción de la ventaja competitiva** adoctrinada por la Estrategia Regional de Investigación Científica, Desarrollo Tecnológico e Innovación de Castilla y León.
- **Profesionalidad** gracias a un personal competente, con rigor profesional y comprometido con el servicio, procurando la estabilidad y desarrollo profesional de los trabajadores.

Visión

Liderar en Castilla y León la gestión y el uso de las Tecnologías de la Información, la supercomputación, las redes de comunicaciones avanzadas para apoyar y mejorar la calidad y la eficiencia de los sistemas de investigación, los sistemas públicos de tratamiento de la información y la productividad y competitividad del tejido empresarial compatibilizando las necesidades de los servicios con los objetivos de desarrollo de los empleados.

Misión

Gestionar infraestructuras basadas en las Tecnologías de la Información y las Comunicaciones para dar servicio a Organismos Públicos de Investigación, Empresas e Instituciones, impulsando la cooperación científica y el uso de infraestructuras comunes, **actuando como centro de servicios de computación y comunicaciones avanzadas** en Castilla y León y promoviendo el desarrollo de la ciencia y la transferencia de tecnología.



Caléndula. Racks, fila delantera. © SCAYLE. Plataforma Tecnológica 2018.

Cluster MPI 1

El clúster para proceso paralelo está compuesto por 371:

- 152 servidores IBM DX360 M4 en formato twin. Cada uno de estos nodos cuenta con dos procesadores Intel Xeon E5-2670 de 8 cores, 32GB de RAM e Interfaz Infiniband FDR10. 2
- 6 servidores HP Proliant SL270s Gen8. Cuentan cada uno con dos procesadores Intel Xeon E5-2670 de 10 cores, 128GB de RAM e interfaz Infiniband FDR. 3
- 30 servidores Supermicro SS 6028TR-HTFR con 2 procesadores Intel Xeon E5-2630V3 con 8 cores, 32GB de RAM e interfaz infinid FDR. 4
- 84 servidores Supermicro SS 2028TP-DC1 FR con 2 procesadores Intel Xeon E5-2630V3 de 8 cores, 32 GB de RAM e interfaz infiniband FDR. 5
- 1 servidor Dell R715 con tarjeta coprocesadora Intel Xeon Phi 5110P. 6
- 1 servidor Supermicro con 2 procesadores Intel Xeon E5-2695 v4 con 18 cores, 384GB de RAM e interfaz infiniband FDR 7

Este nodo integra 8 tarjetas coprocesadoras Nvidia Tesla V100 SXM2.

Sistemas Auxiliares 9

Los sistemas auxiliares son responsables de tareas de gestión interna de SCAYLE y del control y monitorización del resto de la infraestructura. Incluye servidor de hora, monitorización, gestión de despliegue, de colas, etc. Se trata de entre 8 y 12 servidores de diferentes modelos. Entre ellos hay HP Proliant DL160, DELL/EMC R610, DELL/EMC R710 y BULL Novascale. 8

Granja de Virtualización 10

Está basada en un clúster VMWare VSphere 6.5 y contiene máquinas virtuales tanto de gestión interna de SCAYLE como de clientes.

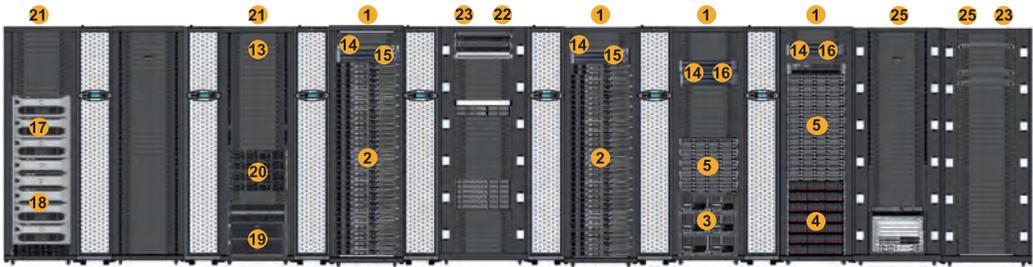
Está compuesto por 15 servidores de diferentes fabricantes, entre los que están Huawei, Dell, Supermicro y HP. Todos ellos con 4 interfaces 10Gbe, 256 GB de RAM y entre 12 y 16 cores por servidor. 11

El servidor de copias de seguridad de esta granja es un HP Proliant DL1 60 g5 con dos procesadores Intel Xeon E5472 con 4 cores (teniendo un total de 8), 16 GB RAM. Con el software Veeam Backup instalado. 12

Redes de Interconexión Internas

En SCAYLE existen físicamente tres redes a las que se conectan todos los equipos:

- Red 10G: el backbone de esta red está formado por dos conmutadores Force1 0 S4810 en configuración de alta disponibilidad (racks 6 y 14) y dos conmutadores Juniper EX4600 (rack 4) en configuración de alta disponibilidad. 13
- Red Gigabit: esta está distribuida en múltiples armarios en una configuración ToR (Top of Rack) con conmutadores de múltiples fabricantes de 1U de altura y un chasis Procurve 5412zl. 14
- Red Infiniband: el núcleo del cálculo intensivo es la red Infiniband. Existe un fabric de tecnología FDR10 15 formado por conmutadores SX6036 y un segundo fabric de tecnología FDR basado en el mismo tipo de conmutador. 16



Caléndula. Racks, fila trasera. © SCAYLE, Plataforma Tecnológica 2018.

Comunicaciones exteriores

Existe una doble conexión a Internet: la conexión a la Red de Ciencia y Tecnología de Castilla y León (RedCayle), que a su vez está conectada a RedIris y un acceso con un proveedor comercial. Cada acceso dispone de dos cortafuegos en alta disponibilidad. La imagen muestra donde se ubican los sistemas de comunicaciones 22 y las fibras ópticas de entrada 23.

Alta eficiencia

El datacenter de SCAYLE está preparado para alta densidad (hasta 44KW/rack) y alta eficiencia (PUE=1,25). Para ello se utilizan enfriadoras de agua con free cooling y tecnología HACS (pasillo caliente cerrado) con intercambiadores agua/aire APC InRow RC 24 ubicados entre cada dos armarios. Esta configuración es simétrica en la fila delantera y trasera.

Red Regional

SCAYLE aloja en su CPD un Nodo de la red óptica de alta capacidad de Castilla y León (RedCAYLE).

Para ello cuenta con varios armarios 25 dedicados a equipos DWDM de conmutación óptica y equipos de routing. Actualmente está instalado un equipo INFINERA DTN como nodo integrante del anillo global de la Red Regional, cada uno de estos chasis puede alcanzar una capacidad máxima de 400 Gb/s. Por otro lado, existe también un equipo de menor capacidad INFINERA ATN que conecta con un nodo final en el Campus de Ponferrada.

Almacenamiento 21

El almacenamiento se basa en una Equallogic PS6XXX y un sistema de ficheros Lustre.

La Equallogic está formada por 9 cabinas (Tres PS6100E 17 y cinco PS6100X 18 en local y una PS6610 en remoto). Disponen de doble controladora con memoria no volátil. En cuanto a capacidad de almacenamiento, cada PS6100E tiene un total de 48TB distribuidos en 24 discos NL-SAS a 7200rpm de 2TB. Cada PS6100X tiene un total de 14TB distribuidos en 24 discos SAS a 10000 rpm de 600GB. La cabina PS6610, usada para backup, dispone de 84 discos de 6TB NL-SAS a 7200rpm.

El almacenamiento Lustre está formado por dos cabinas Eternus DX100 S3 19 y cuatro servidores Primergy RX2530 20, con una capacidad total neta de 512,6 TB. Cada una de las cabinas tiene doble controladora con 1 6GB de memoria caché (8GB por controladora). La primera de ellas, usada para los metadatos, tiene 1 1 discos SAS de 900GB a 1 0000 rpm, mientras que la segunda, destinada a los datos, tiene 1 21 discos NL-SAS de 6TB cada uno a 7200rpm. Cada uno de los servidores tiene a bordo dos procesadores Intel Xeon E5-2630 con 1 6 cores cada uno, y 64GB de RAM .

cartera de servicios

Cálculo intensivo (HPC, High Performance Computing)

La vocación fundamental del Centro es la prestación de servicios de cálculo intensivo, tanto en programas que demandan proceso paralelo masivo (bajo estándar MPI, *Message Passing Interface*), como en aquellos que necesitan gran capacidad en ejecuciones secuenciales o para los que requieran en su ejecución disponer de alta capacidad de memoria RAM. Pero la orientación de SCAYLE no es sólo poner las infraestructuras a disposición de los usuarios para ejecución de programas en el clásico esquema FIFO (First In-First Out).

Caléndula dispone de un datacenter innovador y eficiente que permite el cumplimiento de SLA's (Service Level Agreement-Acuerdo de Nivel de Servicio) exigentes, por lo que el Centro se define como un socio tecnológico idóneo en proyectos HPC donde las condiciones de contorno requieran cumplimientos de plazos estrictos, alta seguridad, etc.

SCAYLE facilita a los usuarios la capacidad de cálculo y de almacenamiento necesario y el software de aplicación. El Centro dispone de un equipo de expertos multidisciplinar para asesorar a los grupos de usuarios en la mejor solución para sus necesidades.

Desarrollo de Proyectos

SCAYLE participa en proyectos de investigación en diferentes campos (astrofísica, meteorología, biotecnología, etc.), aportando conocimiento sobre soluciones innovadoras de cálculo. Esta modalidad de trabajo permite la incorporación de nuevas metodologías basadas en modelización, simulación y uso intensivo de datos en proyectos de investigación de diferentes áreas científicas y tecnológicas, contribuyendo a la mejora de las tareas desarrolladas en los mismos. Esto supone un valor añadido para los usuarios, favoreciéndose el alcance de resultados más precisos en un menor intervalo de tiempo y optimizando los costes.



Servicios TIC para las administraciones públicas

Desde sus comienzos, el Centro ha hecho una apuesta decidida por el *Cloud Computing* y sus tecnologías y metodologías asociadas, sobre la base de que el *Cloud* supone una forma eficiente de utilización de infraestructuras TI. Por ello el centro participa en proyectos de desarrollo de tecnología *Cloud*, y ofrece servicios basados en ella. Para ello se ha creado el concepto de *Datacenter Virtual (Datacenter-as-a-Service)*, y presta servicios para diversas organizaciones.

SCAYLE puede llevar a cabo la prestación de servicios de *Cloud Computing* a la Junta de Castilla y León actuando como *medio propio instrumental* y *servicio técnico* de la misma, no incurriendo en conflictos de competencia:

- Servicio de *Backup*: realización de copias de seguridad de los centros usuarios y almacenamiento de las mismas en el Centro. Los centros usuarios obtienen un doble beneficio:
 - Bajo coste debido a las economías de escala que ofrece una infraestructura como la de SCAYLE.
 - Cumplimiento de la norma: alojar la copia de seguridad fuera de las dependencias donde reside el dato (requisito legal).
- Servicio de *DataCenter Virtual*: se mantiene un entorno virtual completo (servidores, redes, etc.) que puede complementar la demanda de los centros usuarios y acabar reemplazando centros de proceso de datos:
 - Reemplazar CPDs de pequeña entidad y, por tanto, poco rentables.
 - Complementar la capacidad de CPDs con picos de demanda.
- Servicios de *Big Data*: la capacidad de proceso de los sistemas de SCAYLE permiten ofrecer servicios donde son necesarias metodologías *Big Data*.

Operación y Gestión de la Red Regional de Ciencia y Tecnología (RedCAYLE)

El primer propósito de la red es la prestación de servicios de conectividad y telecomunicaciones, pero además es el medio idóneo para canalizar servicios de valor añadido:

- Servicios de Operación de Red:
 - Gestión de las incidencias de la red.
 - Control de las métricas.
 - Gestión de los equipos de la red regional.
 - Control del inventario y versiones.
- Servicios de Valor Añadido:
 - Consolidación de servidores.
 - Optimización de licencias de SW.
 - Centros de respaldo.
 - Copias de seguridad remotas.
 - Administración electrónica.
 - Gestión de bibliotecas.
 - Servicios de aulas virtuales.

Supercomputación y RES

Supercomputación

Hoy en día, para poder hacer frente a los complejos cálculos científicos o técnicos con los que los investigadores trabajan, se hace imprescindible el uso de equipos informáticos extremadamente potentes: los superordenadores.

Los superordenadores se componen de centenares o miles de potentes ordenadores, ordenadores que son convencionales en cuanto a sus componentes y diseño. Sin embargo, disponen de una tecnología que facilita el fraccionamiento de un problema complejo en tareas más sencillas que sean asumibles por los ordenadores que forman el superordenador, estas pequeñas tareas pueden ser calculadas en distintas partes del superordenador para que finalmente, y trabajando de forma coordinada, sean capaces de llegar a la solución buscada.

El FLOP (FLOating Point operations en inglés), operaciones de coma flotante en castellano, es la unidad que mide el rendimiento de los superordenadores, y de forma sencilla se puede interpretar que una operación de coma flotante, sea una suma, resta, multiplicación o división realizada cada segundo.

Aplicaciones como las usadas en la predicción meteorológica, o las utilizadas en la modelización del comportamiento de los fluidos como pueden ser al aire o el agua en el diseño de coches o barcos, son aplicaciones tradicionalmente empleadas en supercomputación y que ya son conocidas por el usuario de a pie. Pero hay otros usos de los superordenadores aún muy desconocidos en la sociedad, un claro ejemplo de ello es el crecimiento del uso de la supercomputación en todas las investigaciones relacionadas con las ciencias de la vida, los estudios en genética o el desarrollo de nuevos fármacos, que demandan grandes potencias de cálculo, son un claro ejemplo de la necesidad del empleo de este tipo de infraestructuras tecnológicas para lograr un avance en la investigación, y que en un futuro no muy lejano se convertirá en una herramienta de trabajo habitual y necesaria.

Otro uso de la supercomputación en rápido crecimiento es la ejecución de trabajos que hacen uso de técnicas y métodos de la Inteligencia Artificial. Así términos como Machine Learning (ML) y Deep Learning (DL) (un subcampo dentro de Machine Learning) se han incorporado al lenguaje habitual entre el personal que mantiene y administra supercomputadores.

Estas nuevas técnicas también han obligado a la incorporación de hardware específico, especialmente diseñado para llevar a cabo las tareas necesarias, como son las GPUs (Graphics Processing Units). En sus últimas versiones, estas tarjetas gráficas integran la electrónica para trabajar de forma óptima con algunas de las estructuras de datos usadas en ML y DL.



Renovación tecnológica 2018 de equipos cedidos por el Centro Nacional de Supercomputación para cálculo de usuarios de la RES©. ©SCAYLE.

SCAYLE: ICTS distribuida

"Red Española de Supercomputación"

En el año 2015 SCAYLE-Supercomputación Castilla y León se incorporó como miembro de la Red Española de Supercomputación RES. La RES está reconocida por el Ministerio de Economía, Industria y Competitividad como una Infraestructura Científico y Técnico Singular distribuida-ICTS distribuida. Desde el pasado 19 de noviembre de 2018, el Consejo de Política Científica, Tecnológica y de Innovación (CPCTI) ha aprobado el nuevo mapa de ICTS en el que ya se ha incorporado la RES como ICTS distribuida, con los nuevos nodos que conforman la RES entre los que está SCAYLE. A su vez, la ICTS distribuida "Red Española de Supercomputación-RES" ha quedado integrada en la nueva Red de ICTS "Red de e-Ciencia" que se constituirá a lo largo del 2019.

La RES es una red que interconecta a 11 instituciones y 12 superordenadores distribuidos por todo el territorio nacional y que colaboran juntos para proporcionar recursos de supercomputación en los diferentes proyectos de investigación a los que estén vinculados.

Dentro de la RES, SCAYLE, a través del superordenador Caléndula, dedica parte de su sistema de cálculo para proporcionar horas de cálculo a los investigadores que participan de las convocatorias de la RES, de la misma forma, a dichos investigadores también se les proporciona soporte por parte del personal técnico de SCAYLE, para conseguir que los cálculos que realizan sean lo más optimizados posible.

Desde la incorporación de SCAYLE a las distintas convocatorias públicas publicadas por la RES, se han asignado aproximadamente 25,5 millones de horas de procesador a los diferentes proyectos asignados y han hecho uso de nuestro sistema de cálculo más de 30 grupos investigadores.

Entre los proyectos asignados a Caléndula podemos mencionar proyectos de investigación sobre nuevas técnicas de radioterapia, cálculos sobre tectónica de placas, investigación sobre nuevos materiales o el análisis de datos provenientes de diferentes instrumentos de observación del Universo.



Mapa de Infraestructuras Científicas y Técnicas Singulares (ICTS). © RES.

La Fundación Centro de Supercomputación de Castilla y León asumió las funciones de "Centro de Gestión y Operación de RedCAYLE (Red de Ciencia y Tecnología de Castilla y León)" en virtud del convenio con la Administración General de la Comunidad de Castilla y León de 10 de junio de 2016.

El convenio por el que SCAYLE asume las funciones de Centro de Gestión y operación de RedCAYLE detallan funciones tales como:

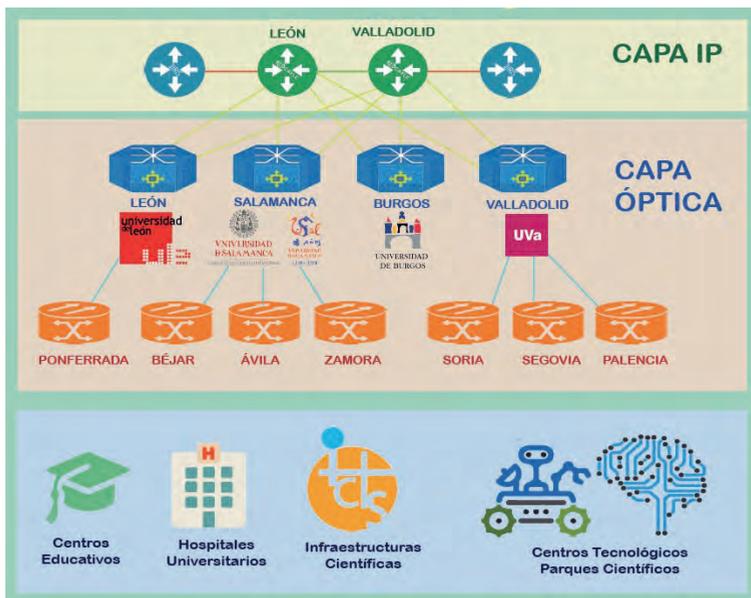
- operación diaria de la red,
- configuración y mantenimiento de circuitos,
- red de fibras y equipos,
- gestión de las incidencias en la red de fibras,
- monitorización de la red y control de métricas,
- atención de usuarios,
- incorporación de nuevos servicios a la red,
- altas de nuevos nodos.

El ámbito territorial de RedCAYLE es la Comunidad Autónoma de Castilla y León. En este último año se ha avanzado en proyectos de fibra urbana para facilitar la conexión a nuevas entidades locales, de este modo se pretende que en los siguientes años se amplíe de forma estratégica la capilaridad de la huella de fibra actual.

Durante el 2018 se ha continuado con la migración de los diferentes afiliados a RedIRIS en la Comunidad de Castilla y León, así como la conexión a Internet de la red de colegios de Castilla y León.



Código 14-3R-C001



REDCAYLE: Diagrama de Capas. © SCAYLE.

El Centro de Supercomputación de Castilla y León actúa como oficina técnica y centro de coordinación de la RedCAYLE

A finales de 2015 la Red se había desplegado en su totalidad con los 11 campus universitarios públicos de Castilla y León conectados a ella. Estas interconexiones suponen un gran salto cualitativo respecto a lo que disponen actualmente, pero apenas suponen una fracción de la capacidad posible de la Red. La Red está técnicamente diseñada para poder soportar sobre ella redes de alta capacidad superpuestas y separadas. Este diseño técnico facilita otros usos: permite incorporar los más de 1.000 centros de enseñanza primaria y secundaria existentes en la Comunidad mediante un proyecto puesto en marcha por Red.es, y de esta forma disponer de una red de interconexión de muy alta capacidad para centros docentes universitarios y no universitarios. De la misma manera, la propia administración autonómica puede rentabilizar la Red al tener la posibilidad de utilizarse los enlaces entre capitales para los usos corporativos de la Junta de Castilla y León.

El Centro de Supercomputación de Castilla y León ha sido el responsable del diseño y despliegue de la red Regional de Castilla y León, y lo es también de su operación y explotación. Asimismo, SCAYLE es el interlocutor con RedIRIS para las tareas necesarias de coordinación entre redes, participa en el Grupo de Trabajo de Redes Regionales de RedIRIS, en el que tiene un papel muy relevante, y es LIR (Local Internet Registry) de RIPE, la organización que controla el espacio de direcciones de Internet en Europa, Oriente Medio y Asia Central.

Para la puesta en marcha de la red, el centro adquirió los correspondientes equipos, que están basados en tecnología de multiplexación densa en longitud de onda (Dense Wavelength Division Multiplexing, DWDM). Gracias a la Red Regional se ha multiplicado por diez la velocidad de conexión entre los campus universitarios periféricos y centrales, de esta forma las conexiones de los campus de Zamora, Ávila y Béjar con el de Salamanca; las de Palencia, Segovia y Soria con el de Valladolid y la conexión de Ponferrada con León se realizan actualmente a través de circuitos de 10Gbs. A finales de 2016 se instalaron dos Routers de altas prestaciones, uno en León y otro en Valladolid, que absorberán todo el intercambio de tráfico IP entre los centros conectados a la Red Regional y el resto de redes académicas nacionales e internacionales. La instalación de estos Routers supone también un salto de calidad del servicio, ya que actualmente sólo existe uno para dar servicio a Castilla y León y, en caso de fallo, podría quedar incomunicada la región.

Durante 2018 se ha continuado con la extensión de la red que ha llegado al Hospital Universitario de León y con la migración de las universidades y centros de investigación de la comunidad.



Categorización del tráfico IP de RedCAYLE. © SCAYLE.

Servicios TIC para las administraciones públicas

El centro de datos entendido como entidad física y aislada que proporciona servicios físicos a terceros, ha dado paso a otro modelo. Ahora estamos en la era de la virtualización y de las infraestructuras distribuidas geográficamente.

Los centros de datos son el núcleo central de soporte físico de los servicios ofrecidos por Internet usando líneas de comunicaciones de alta velocidad. Desde el hospedaje de equipos, comercio electrónico, redes sociales y en general conceptos recientemente puestos en práctica a raíz del modelo de computación en la nube o modelo cloud donde todo se ofrece como servicio.

La evolución del modelo cloud es continua y superando los tres paradigmas originales de prestación de servicios: SaaS (software como servicio), PaaS (plataforma como servicio) e IaaS (infraestructura como servicio); han surgido muchos más como HaaS (hardware como servicio), NaaS (red como servicio) o DCaaS (datacenter como servicio). Tal ha sido el crecimiento que se ha dado en llamar XaaS (Anything as a service) para definir a la prestación de cualquier servicio del ecosistema cloud.

En definitiva, el cloud, apoyado en técnicas de virtualización, nos permite la optimización y consolidación de recursos IT y mejora la redistribución de éstos en función de las necesidades de los servicios que lo requieran.

Se están llevando a la nube los servicios asociados al nicho de negocio de SCAYLE para su virtualización, hecho que permite despreocuparse de la infraestructura física, dejando esa gestión a entidades que poseen la tecnología, y que proporcionan el mismo servicio a costes más bajos. Sería un error dar salida a toda esta demanda únicamente construyendo centros de datos más grandes y potentes, sin tener en consideración otros aspectos que ayuden a llevar a cabo estos nuevos retos.

La creación y gestión de plataformas globales para la virtualización de centros de datos en el ámbito de las administraciones públicas, proporciona la posibilidad de tener escalabilidad no limitada, movilidad de la plataforma completa y gestión operacional centralizada que facilita el mantenimiento, control y monitorización de la misma. Estos objetivos marcados, suponen una reducción de costes significativa comparados con una implementación tradicional de la misma infraestructura a nivel local, así como dotar de capacidad móvil al conjunto de la plataforma hacia cualquier otra infraestructura física.



PROYECTOS I+D+i

Colaborador

Defectos topológicos en teorías de campos clásicas y cuánticas

Detección de muestras de malware pertenecientes a APTs mediante aprendizaje basado en Machine Learning

SMART-ID, Identificación inteligente de objetos y personas para incrementar la seguridad de los autobuses

Computación paralela de medidas de centralidad de grafos

Diseño y caracterización de líquidos porosos para la captura y el almacenamiento de CO₂

SEÑALES, Integración de datos obtenidos mediante sensores remotos para el desarrollo de herramientas predictivas y de seguimiento a gran escala ante el cambio global en el sector forestal

Nuevos escenarios de la óptica de pulsos de Femto y Attosegundos

Nuevas tecnologías de secuenciación (NGS) para el estudio de los micovirus en *Fusarium circinatum*

Análisis de múltiples pathways

AMIGA6, gas in and around galaxies. Preparation for SKA science and contribution to the design of the SKA data flow

Carácter multifactorial de los polifenoles: oportunidad para el desarrollo de herramientas terapéuticas frente a obesidad y enfermedades infecciosas

Ajuste, Validación e Implantación del Modelo Físico PHFFS de Simulación de Incendios Forestales

Cirugía Oncológica: modelo de gestión y formalización del conocimiento (casos de estudio cáncer colorrectal - CCR)

Colaboraciones en proyectos I+D+i a través de la RES

Publicaciones científicas

Defectos topológicos en teorías de campos clásicas y cuánticas

Objetivo del proyecto

Este proyecto se enmarca en el estudio del scattering de soluciones de tipo solitón o kink (defectos topológicos) que aparecen en las ecuaciones en derivadas parciales de tipo Klein-Gordon no lineales asociadas a teorías de campos escalares. Estas soluciones pueden ser interpretadas como partículas extensas en el marco físico. El choque de dos de estas soluciones constituye un proceso complejo que deriva en interacciones internas de los defectos topológicos poco conocidas y que dependen de forma crítica de la velocidad de impacto. Existen rangos de la velocidad de colisión en los que los kinks: (1) se ven forzados a formar un estado ligado cuasi-estable, (2) se aniquilan entre sí, (3) chocan y rebotan alejándose, (4) entran en resonancia rebotando un número finito de veces antes de escapar. La distribución de las ventanas de velocidades iniciales donde aparece resonancia sigue una estructura fractal.

El objetivo de este proyecto es estudiar estos fenómenos en distintos modelos de teorías de dos o más campos escalares. Se estudia además el caso en el que los campos están confinados en una esfera (o en otra variedad riemanniana), lo que permite describir modelos asociados a cadenas de espines. También resulta de interés en este proyecto analizar los procesos de scattering en modelos que involucran campos vectoriales, en los que los defectos topológicos a tratar son vórtices o monopolos.

Periodo de ejecución

Marzo del año 2018 a diciembre del 2020.

Financiación del proyecto

Ministerio de Economía y Competitividad, MINECO, www.mineco.es, a cargo del proyecto "Nuevos retos en Sistemas Dinámicos Supersimétricos y Superintegrables".

Participantes del proyecto

Universidad de Salamanca, www.usal.es

Instituto Universitario de Física Fundamental y Matemáticas (IUFFyM), www.usal.es/iuffym

UIC Física Matemática de Castilla y León

Líder del proyecto

El Grupo de Física Matemática de la Universidad de Salamanca, <http://campus.usal.es/~mpg/>. Es un grupo de investigación reconocido de la Universidad de Salamanca que a su vez está integrado en la Unidad de Investigación Consolidada de Física Matemática de la Junta de Castilla y León (UIC 011 MathPhys-CyL) desde 2015.

El grupo está constituido por cuatro investigadores senior de la Universidad de Salamanca y uno de la Universidad de Oviedo.

Las líneas de investigación de este grupo involucran el estudio de defectos topológicos en Teorías de Campos Clásicas y Cuánticas, Mecánica cuántica supersimétrica, sistemas dinámicos integrables, etc.



Justificación del proyecto

El estudio de las soluciones de tipo defecto topológico que aparecen en ecuaciones no lineales es un ámbito de investigación muy activo desde hace ya varias décadas, dado que permite explicar algunos fenómenos físicos que no tenían parangón desde la perspectiva de las teorías lineales.

La presencia de fenómenos de superconductividad y superfluidez en Materia Condensada, las propiedades de algunos polímeros unidimensionales como el polioxietileno con amplias aplicaciones biomédicas y biotecnológicas, la evolución del universo temprano a través de la formación de paredes de dominio en Cosmología, etc., son ejemplos particulares del marco general aludido previamente.

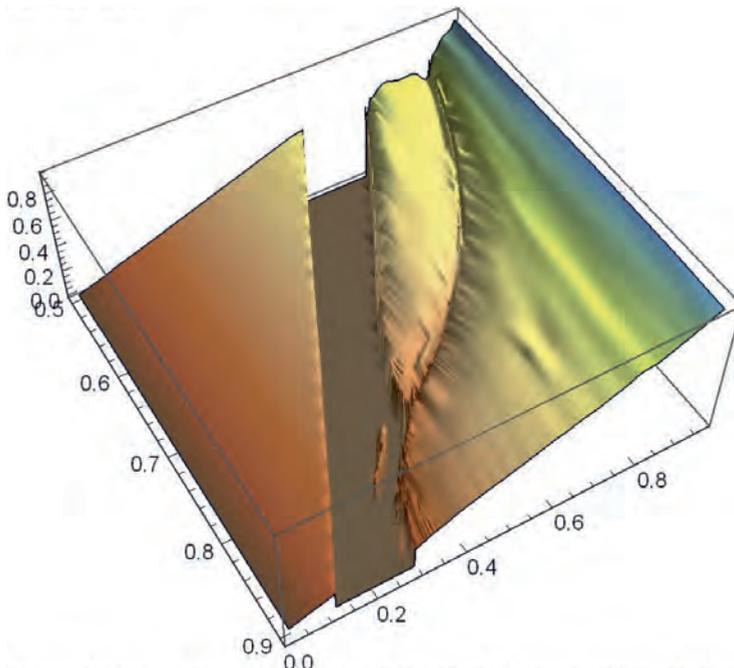
En cada uno de dichos casos la presencia de estructuras topológicas que interactúan entre sí es un escenario muy probable en la realidad. Es, por tanto, imprescindible comprender los posibles eventos de scattering que pueden darse, y que pueden alterar la naturaleza de los fenómenos explicados por este tipo de soluciones no lineales.

El análisis de la dependencia de dichos procesos de choque con respecto a la velocidad de impacto de los defectos topológicos es también un aspecto esencial de este tipo de estudios.

Funciones de SCAYLE

El estudio de la evolución de soluciones de tipo soliton o kink en las ecuaciones de Klein-Gordon no lineales exige la ejecución de algoritmos que requieren un alto nivel de cálculo computacional para tener una fotografía de alta resolución de los posibles escenarios en el proceso de scattering de defectos topológicos. Los fenómenos de resonancia implican la presencia de una dependencia caótica de las velocidades de choque (introduciendo una distribución de las velocidades finales que involucran patrones fractales). Este comportamiento exige el lanzamiento de millones de simulaciones en el choque de kinks para tener una visión precisa del comportamiento del scattering de estos objetos.

Es en este punto donde el supercomputador de SCAYLE representa un papel esencial en la investigación, permitiendo realizar estas simulaciones de forma paralela (aprovechando el gran número de procesadores con altas prestaciones de memoria de las que dispone Caléndula) en un tiempo de ejecución razonable. Gracias a SCAYLE se pueden obtener resultados sobre el comportamiento de defectos topológicos que de otra forma no sería viable.



Representación gráfica de la velocidad final de los kinks resultantes en la colisión de defectos topológicos en los modelos MSTB. La velocidad final es dibujada como función de la velocidad de impacto y de la constante de acoplamiento del modelo de campos.

Detección de **muestras de malware** pertenecientes a APTs mediante **aprendizaje** basado en **Machine Learning**

Objetivo del proyecto

El objetivo de este estudio es el desarrollo de una herramienta que permita la detección de muestras de malware perteneciente a APTs mediante el análisis estático y dinámico de muestras anteriores utilizando técnicas de Machine Learning.

Periodo de ejecución

15 de marzo de **2018** al 17 de diciembre de **2019**.

Participantes del proyecto

INCIBE, www.incibe.es

SCAYLE Supercomputación Castilla y León, www.scayle.es

Financiación del proyecto

El proyecto se enmarca dentro de la adenda 18 del convenio de colaboración entre la universidad de León e INCIBE ("Desarrollo del equipo de investigación avanzada en ciberseguridad").

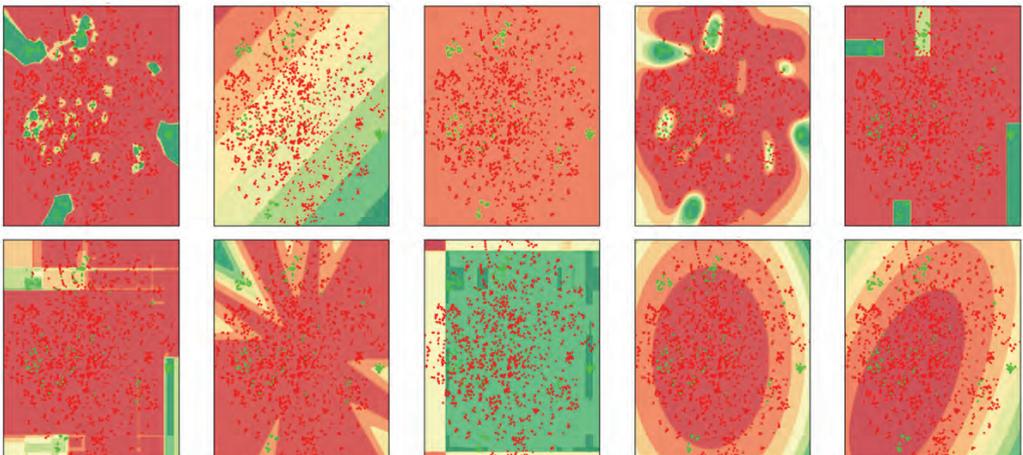
Funciones de SCAYLE

Scayle ha aportado el apoyo técnico y logístico para la realización de la investigación en sus

instalaciones. Los cálculos en el supercomputador han permitido agilizar el proyecto, gracias a la capacidad de cómputo de Caléndula. SCAYLE también ha prestado soporte a los procesos de evaluación del rendimiento, así como proporcionado el espacio de almacenamiento para los datos de entrenamiento.

En una segunda fase el objetivo del proyecto es, entre otros, poder discernir, dentro del tráfico habitual de una organización si se vislumbra tráfico perteneciente a un ataque APT. Para ello es necesario disponer de una fuente de tráfico estable.

La RedCayle, Red Iris de Castilla y León, que gestiona Scayle es el medio perfecto para simular el tráfico que tendría una organización. El análisis del tráfico generado por el malware de una APT dentro del flujo de tráfico de RedCayle, es analizado en busca de patrones que permitan detectarlo en una organización real.



Detección de grupos de APTs utilizando diferentes técnicas de Machine Learning.

Justificación del proyecto

Una APT (del inglés Advanced Persistent Threat) es un tipo de ataque basado en malware, a menudo orquestado con un objetivo específico, dirigido a penetrar la seguridad informática de una entidad específica, ya sea privada o pública. Una APT generalmente fija sus objetivos en organizaciones o naciones por motivos de negocios o políticos. Este tipo de ataques involucra generalmente técnicas muy sofisticadas que utilizan malware para explotar vulnerabilidades, a menudo de día cero, en los sistemas internos de una organización.

A menudo, cuando las entidades son atacadas de forma exitosa mediante un ataque APT, no son conscientes de dicho ataque hasta pasado un tiempo muy elevado, que puede llegar a ser un plazo de años, tiempo durante el cual han perdido su ventaja corporativa, así como los datos de usuarios y clientes, lo que implica una pérdida de imagen y credibilidad de la que a menudo no pueden recuperarse.

Por otra parte, muy comúnmente las casas antimalware son incapaces de detectar si una muestra de malware pertenece o no a un ataque APT, puesto que se trata de muestras de malware normales que se camuflan entre los millones de muestras que estas casas antimalware reciben a diario.

La detección temprana de un ataque APT ayudaría a una entidad a evitar dichas pérdidas de dinero y credibilidad. Sin embargo, la detección temprana de ataques APT continúa siendo, a día de hoy, una tarea muy compleja.

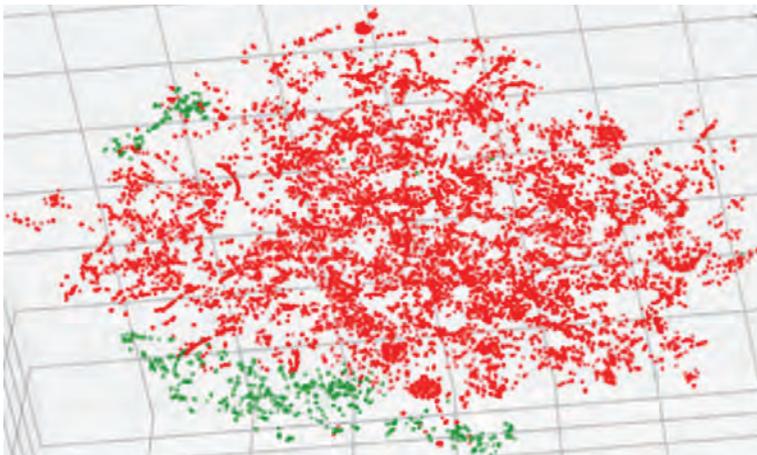
El presente proyecto pretende desarrollar un modelo inteligente, mediante técnicas de Machine Learning que permita detectar si una muestra de malware es susceptible de pertenecer a un ataque APT. Para ello se apoya por una parte en el análisis estático y dinámico de las muestras de malware y por otra en el análisis de tráfico de red generado por las APTs, de manera que la herramienta desarrollada sea capaz de analizar el tráfico de red circulante y permita determinar si una entidad está siendo infectada por un APT de una forma más proactiva.

Líder del proyecto

El proyecto está liderado por el profesor **Adolfo Rodríguez de Soto** y el grupo de trabajo formado en la actualidad por el investigador **Luis Martín Liras**.

Adolfo Rodríguez de Soto es Profesor Titular de Ciencias de la Computación e Inteligencia Artificial en el Departamento de Ingenierías Mecánica, Informática y Aeroespacial de la Universidad de León. Ha impartido docencia en la Universidad de León desde 1990 en diversas asignaturas, principalmente en el área de Estadística y de la Programación. Es autor y coautor de 3 libros docentes y ha participado en varios proyectos en este ámbito.

Su labor investigadora se enmarca en el área de la Inteligencia Artificial, siendo autor y coautor de más de 40 publicaciones tanto nacionales como internacionales y en revistas internacionales de alto impacto y ha participado en más de 35 congresos nacionales e internacionales.



Nube de puntos representando las muestras de malware de APT(en verde) y las muestras de malware lanzadas, donde se aprecia que es posible filtrar unas sobre otras

SMART-ID, Identificación inteligente de objetos y personas para incrementar la seguridad de los autobuses

Objetivo del proyecto

El objetivo del proyecto es desarrollar una solución tecnológica para incrementar la seguridad del transporte urbano, gracias a la capacidad que ésta tendrá de identificar objetos sospechosos o personas peligrosas dentro de un autobús. Se instalarán en el autobús sistemas autónomos embarcados que reportarán todas las incidencias detectadas al sistema central, considerándose incidencias a las personas no autorizadas en el puesto del conductor, objetos sospechosos y personas consideradas peligrosas por estar fichadas por las Fuerzas y Cuerpos de Seguridad del Estado. Desde el sistema central será posible supervisar las incidencias registradas, lo que traerá consigo la confirmación de la existencia de una situación de peligro que podrá requerir la activación de protocolos de seguridad.

Periodo de ejecución

Desde **2018** al **2019**.

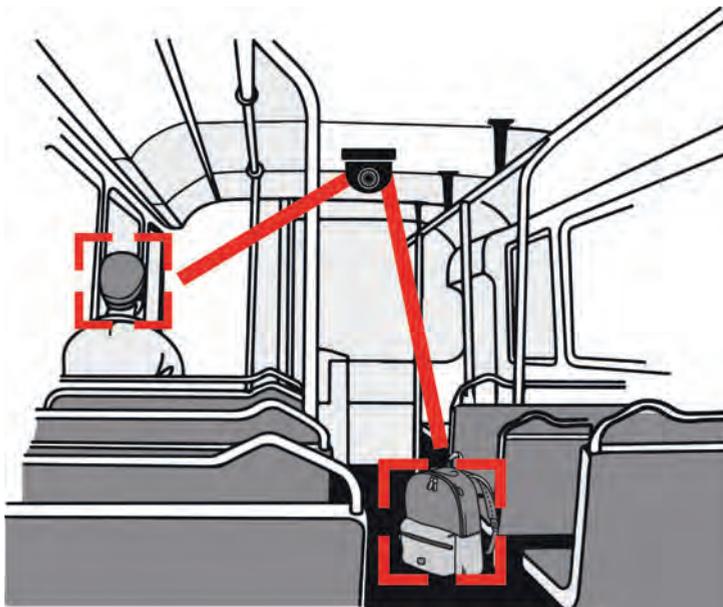
Participantes del proyecto

Proconsi, www.proconsi.com

ALSA Tánger, www.alsa.es

Financiación del proyecto

CDTI - Proyecto de Investigación y Desarrollo. Proyecto cofinanciado por el Fondo Europeo de Desarrollo Regional (FEDER), a través del Programa Operativo Plurirregional de Crecimiento Inteligente 2014-2020.



Smart-Id identificando objetos y personas en el interior de un autobús

Justificación del proyecto

El proyecto busca acercar la última tecnología disponible al sector del transporte, consiguiendo con ello incrementar la seguridad en los autobuses y, con ello, en las ciudades.

Funciones de SCAYLE

Es imprescindible el uso de Caléndula por la alta capacidad de procesamiento de datos necesaria para el entrenamiento de las redes neuronales que se emplean en el proyecto.

Los más de 40000 CUDA cores de sus multi-GPUs reduce el tiempo de una tarea que en nuestros sistemas puede llevar una semana a unas pocas horas.

Esto facilita enormemente la prueba de diferentes librerías (DarkFlow, Darknet, etc) y configuraciones de las mismas; cambios frecuentes en este proyecto con una carga importante de I+D y que sin la reducción de tiempo proporcionada por Caléndula harían inviable la entrega en la fecha planificada.



Código IDI-20180057

Líder del proyecto

Proconsi, www.proconsi.es, es una empresa de Tecnologías de la Información y la Comunicación especializada en el desarrollo e integración de soluciones informáticas para todo tipo de empresas. Más de dos décadas de experiencia avalan a una compañía tan flexible como responsable. Cuenta con un equipo multidisciplinar de más de 100 profesionales cualificados, expertos y comprometidos con un único objetivo: hallar la solución tecnológica exacta para cada cliente. PROCONSI es especialista en la creación y el desarrollo de software de gestión, consultoría tecnológica, dirección y gestión de proyectos I+D+i basados en TIC, soporte técnico, aplicaciones móviles y fomento de tendencias en nuevas tecnologías, como el cloud computing.

PROCONSI es un líder tecnológico que se encuentra en pleno proceso de expansión. Con un crecimiento anual sostenido mantenido en los últimos años y un incremento continuo en Investigación, Desarrollo e Innovación (I+D+i), la compañía afianza su posición nacional y abre las puertas a nuevas fronteras en Europa, Latinoamérica y África. Como consecuencia, la plantilla de trabajadores se amplía continuamente con el fin de enriquecer la investigación, pulir los proyectos de innovación y dar salida a las herramientas tecnológicas más precisas del mercado TIC. De esta forma los clientes de PROCONSI logran convertir sus negocios en empresas altamente competitivas y rentables dentro de su propio sector.

Computación paralela de medidas de centralidad de grafos

Objetivo del proyecto

Implementación y estudio de rendimiento de algoritmos de cómputo de centralidad de grado, cercanía e intermediación. Paralelización a nivel de medida y de grafo/subgrafos.

Periodo de ejecución

Mayo de 2018 a diciembre de 2018.

Participantes del proyecto

RIASC - Universidad de León, www.riasc.unileon.es

INCIBE, Instituto Nacional de Ciberseguridad, www.incibe.es

Financiación del proyecto

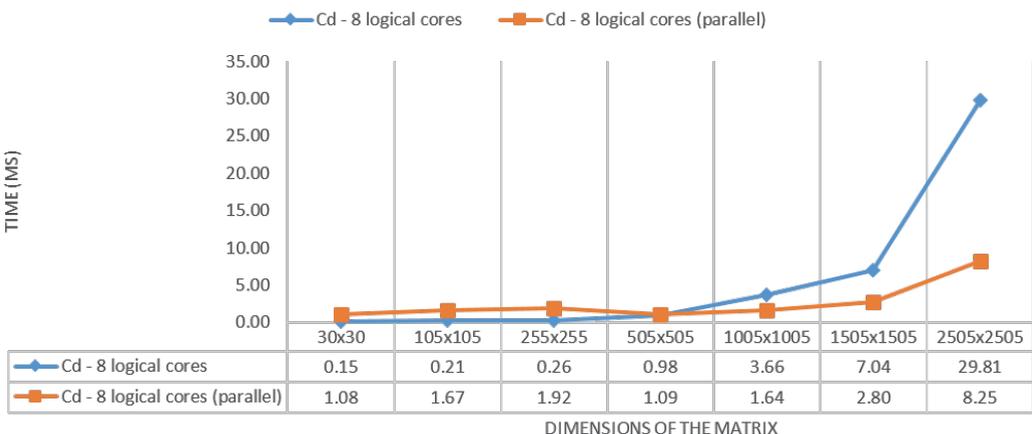
Estudio enmarcado dentro de la adenda 18 del convenio marco con INCIBE - Instituto Nacional de Ciberseguridad.

Justificación del proyecto

Los diferentes indicadores de centralidad de redes y grafos permiten determinar los nodos más importantes de las mismas de acuerdo a diferentes criterios. Su cómputo conlleva cierta complejidad, lo que le hace costoso para redes de mediano y gran tamaño. La paralelización a distintos niveles de estos algoritmos debería mejorar el rendimiento de las implementaciones existentes, lo que permitiría trabajar en la práctica con redes de mayor tamaño.

Multitud de disciplinas que trabajan con redes o sus representaciones, tanto tecnológicas como biosanitarias, socioeconómicas, etc. Disciplinas que se verían, por tanto, beneficiadas con esta aportación.

CD MEASURE-LEVEL PARALLELIZATION



Paralelización a nivel de medida, dado un grafo o red, se calcula la centralidad de grado de todos sus nodos de forma paralela

Funciones de SCAYLE

El Superordenador Caléndula se ha utilizado en la fase de evaluación de rendimiento de los algoritmos paralelizados utilizando memoria compartida (uso de POSIX Threads), confirmando los resultados previos obtenidos en plataformas de menor potencia de cálculo, y permitiendo calcular indicadores para redes de mayor complejidad (mayor número de nodos y/o enlaces).

Dependiendo de la medida de centralidad, se han llegado a obtener algoritmos hasta n veces más rápidos que la versión secuencial de los mismos, siendo n el número de procesadores lógicos de la plataforma de ejecución.

En un futuro próximo, se utilizará Caléndula para realizar pruebas de paralización específicas de entornos HPC (*High Performance Computing*) usando memoria distribuida mediante MPI (*Message Passing Interface*).



INSTITUTO NACIONAL DE CIBERSEGURIDAD



Adenda 18 - "Acuerdo de colaboración para el desarrollo del equipo de investigación avanzada en ciberseguridad y estudios especializados"

Líder del proyecto

La adenda 18 está liderada por el Dr. Miguel Carriegos, director del Instituto de Ciencias Aplicadas a la Ciberseguridad, además de profesor titular de la Universidad de León. Nacido en León, se licenció en Matemáticas en la Universidad de Valladolid en 1994, centro en el que obtuvo en 1999 el grado de Doctor. Su vinculación con la ULE se remonta a 1996. En su trayectoria en Vegazana, Carriegos fue Subdirector de la Escuela de Ingenierías entre los años 2003 y 2007, y Director del Departamento de Matemáticas desde 2008 hasta 2016. Es Director del Instituto de Ciencias Aplicadas a la Ciberseguridad (RIASC) desde 2014.

El estudio e implementación de los algoritmos paralelos corre a cargo de Juan Felipe García Sierra, Doctor Ingeniero en Informática y Coordinador de Especialidad del RIASC. De 2008 a 2016 colaboró con el Grupo de Investigación de Robótica de esta misma Universidad. Durante los últimos cuatro años antes de unirse a RIASC en 2014, trabajó simultáneamente como profesor asociado y desarrollador de software en la empresa Indra. También ha trabajado como ayudante de investigación en la Universidad Nacional de Irlanda, Galway. Sus intereses de investigación incluyen la robótica, el control de la atención, la visión artificial y por computador, la biometría, el desarrollo de software seguro y la vigilancia.

Diseño y caracterización de líquidos porosos para la captura y el almacenamiento de CO₂

Objetivo del proyecto

Se basa en el estudio bottom-up de los líquidos porosos para la separación de gases; el estudio químico físico de las propiedades micro y macroscópicas de los sistemas líquidos con porosidad permanente, y en el estudio de la capacidad de dichos sistemas para separación de gas.

Periodo de ejecución

1 de marzo de 2017 al 28 de febrero de 2021.

Financiación del proyecto

Convocatoria de propuestas H2020-MSCA-RISE-2016(MSCA-RISE).

Entidades Participantes del proyecto.

Università di Sassari, www.uniss.it

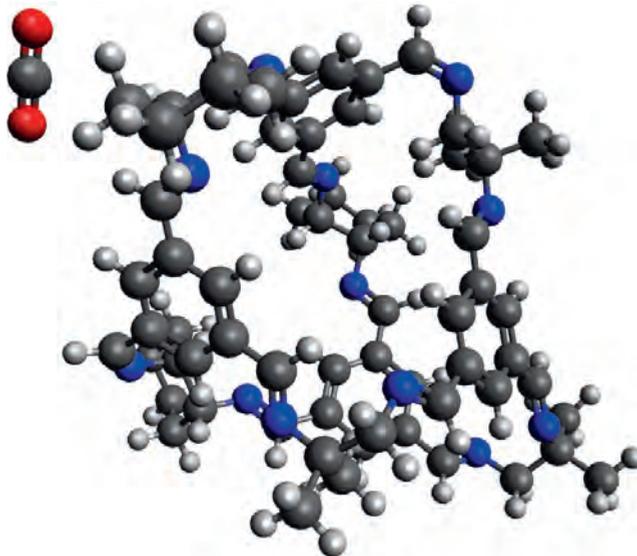
Universidad de Burgos, www.ubu.es

Helmholtz-Zentrum Geesthacht Zentrum für Material und Küstenforschung GmbH, www.hzg.de

Monolithos katalitis kai anakiklosi etaireia periorismenis efhinis Vlachos Nicholas,

Universidad de Chile, www.uchile.cl

Comisión Nacional de Energía Atómica, www.cnea.gov.ar/es



Interacción entre una molécula porosa con el gas CO₂.

Justificación del proyecto

En el proyecto CO2MPRISE se propone buscar nuevos sistemas para la captura y el almacenamiento de CO₂, desde el diseño y la caracterización de nuevas estructuras moleculares hasta la implementación a una escala más grande y el análisis de sus propiedades macroscópicas. En concreto, el proyecto de investigación se enfoca en la búsqueda de líquidos porosos (sistemas formados para un solvente y una molécula porosa que, por el tamaño del agujero, no deja entrar el solvente, manteniendo así la estructura líquida). Estas estructuras pueden almacenar CO₂ y funcionar, así, como sistemas de separación y almacenamiento de los gases. Dichos sistemas pueden encontrar aplicaciones en varios ramos industriales, debido a su potencial capacidad de separar CO₂ a partir del mismo flujo de gas (durante el proceso) y no, como en el caso de materiales porosos sólidos, cuando el proceso se haya acabado.

Funciones de SCAYLE

El acceso a la infraestructura de SACYLE dota de los recursos computacionales necesarios para poder realizar los cálculos DFT y MD.

Líder del proyecto

UNIVERSIDAD DE BURGOS, www.ubu.es.

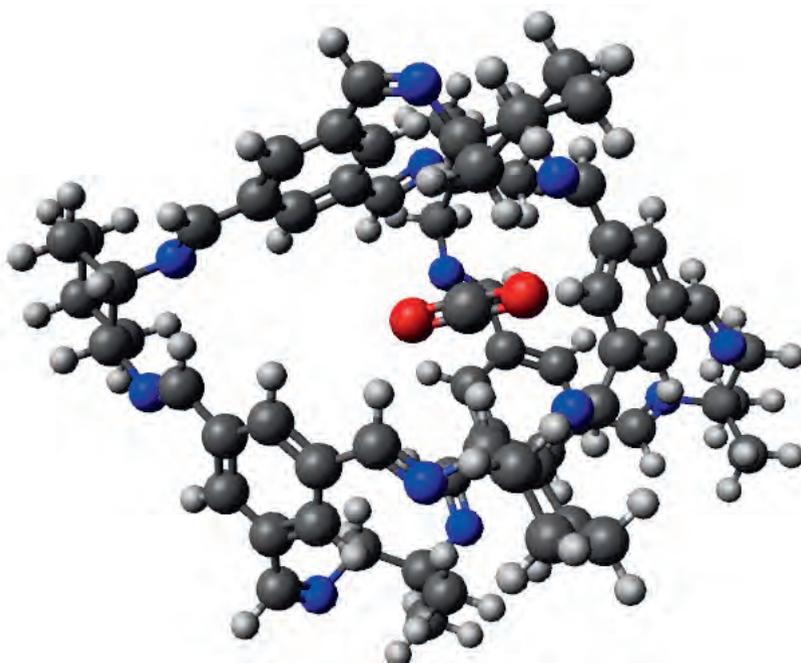
El grupo de investigación está formado por los profesores Santiago Aparicio (IP) y Rafael Alcalde, junto con los estudiantes de doctorado Claudia Pecoraro, Loukia Maritsa, Alberto Gutiérrez, Cesar Herrera y la investigadora postdoctoral María A. Monge.

El grupo lidera la parte computacional del proyecto CO2MPRISE y se ocupa de buscar estructuras para la captura y el almacenamiento de CO₂ y estudiar sus propiedades termodinámicas, a través de cálculos mecanocuánticos, y de analizar sus propiedades dinámicas a través de simulaciones MD, MC, KMC.

El objetivo que se persigue en el grupo de investigación es el de encontrar estructuras con una alta capacidad de almacenamiento de CO₂.



RISE-2016-CO2MPRISE-734873



Interacción entre una molécula porosa con el gas CO₂. Uno de los objetivos del estudio es el comprobar si las moléculas porosas pueden accoger el CO₂ y almacenarlo.

SEÑALES, Integración de **datos obtenidos** mediante sensores remotos **para el** desarrollo de **herramientas** **predictivas** y de **seguimiento** a gran escala ante el **cambio global** en el **sector forestal**

Objetivo del proyecto

El principal objetivo es el desarrollo de una herramienta que use sensores remotos para reconstruir y monitorear el impacto de la defoliación por helada tardía en los hayedos de Castilla y León, y su calibrado mediante datos terrestres basados en anillos de crecimiento.

Periodo de ejecución

Del año **2017** al año **2019**.

Participantes del proyecto

Universidad de Valladolid, www.uva.es

Föra forest technologies, www.fora.es/es/

Financiación del proyecto

ORDEN EDU/986/2017, de 8 de noviembre, por la que se resuelve la convocatoria de subvenciones del programa de apoyo a proyectos de investigación cofinanciadas por el Fondo Europeo de Desarrollo Regional.

Justificación del proyecto

El aumento de las temperaturas asociado al cambio global está adelantando la fenología de las plantas. Esto puede tener efectos favorables, ampliando su período vegetativo, especialmente en las especies caducifolias, que pueden desarrollar antes las hojas.

Sin embargo, simultáneamente con el aumento de la temperatura, se está produciendo un incremento en la variabilidad en dichas temperaturas. Esto podría generar un aumento de los fenómenos de heladas tardías que dañarían el aparato fotosintético de las especies caducifolias durante su brotación. Su impacto podría ser especialmente crítico en los límites altitudinales superiores de las especies donde el riesgo de estos fenómenos es mayor de modo natural, pero también para muchos cultivos agrícolas.

Dentro del proyecto SEÑALES se evalúa este fenómeno en los hayedos, formaciones caducifolias de gran valor ecológico y que en Castilla y León se encuentran cerca de su límite sur de distribución europea.



Código VA026P17

Funciones de SCAYLE

Caléndula ha permitido la reducción del tiempo de pre-procesamiento de imágenes provenientes del satélite Sentinel 2.

En este sentido, Caléndula ha sido empleado principalmente para la corrección atmosférica de los niveles digitales de las imágenes mediante software de la Agencia Espacial Europea (ESA). El uso de array jobs ha permitido procesar múltiples imágenes y al mismo tiempo reducir los tiempos de procesamiento en casi un 90%.

Líder del proyecto

El proyecto está liderado por D. José Miguel Olano Mendoza, profesor del área de Botánica de la Universidad de Valladolid, que desarrolla su trabajo en la Escuela de Ingeniería de la Industria Forestal, Agronómica y de la Bioenergía - EiFAB del Campus de Soria, ingenieriasoria.blogs.uva.es/, dentro del Instituto Universitario de Investigación en Gestión Forestal Sostenible - iuFOR, sostenible.palencia.uva.es/, en Soria.

Su investigación se centra en el impacto del cambio global sobre las formaciones vegetales, con especial énfasis en los sistemas gipsícolas, la alta montaña y las formaciones de Juniperus.

Información detallada de sus publicaciones puede encontrarse en https://www.researchgate.net/profile/Jose_Olano.



Impacto de la helada sobre un hayedo en el Macizo del Moncayo a finales de abril de 2017. Los árboles sin hojas perdieron sus hojas en brotación por una helada, tuvieron que brotar de nuevo y la foliación completa no ocurrió hasta finales de junio (más de un mes más tarde de lo habitual). En 2018 hubo una helada de menor intensidad que afectó a alguno de los árboles a mayor altitud. Estos efectos pueden ser identificados con imágenes satelitales.

Nuevos escenarios de la óptica de pulsos de Femto y Attosegundos

Objetivo del proyecto

Los pulsos láser ultracortos, con duración de unos cuantos femtosegundos (10^{-15} s) han supuesto una auténtica revolución tecnológica en la física de fenómenos ultrarrápidos. Hoy en día, el desarrollo de nuevas tecnologías de anclaje de modos y postcompresión permiten la generación de pulsos en el infrarrojo (IR) de pocos ciclos de duración (unos cuantos femtosegundos), perfectamente coherentes y con parámetros controlados. El desarrollo de esta tecnología láser, unido a la aparición de nuevas técnicas de generación de armónicos, permiten generar pulsos con radiación en el ultravioleta lejano (XUV) o, incluso, en los rayos X blandos, coherentes y con duraciones de unas decenas de attosegundos (10^{-18} s). Estos pulsos láser de attosegundo constituyen las herramientas más precisas para observar y controlar los procesos más rápidos que ocurren en la naturaleza, como por ejemplo el movimiento de electrones dentro de sistemas atómicos o moleculares. Comparado con otros escenarios en la frontera de la física, la investigación en este campo no requiere de grandes instalaciones, lo que fomenta la diversidad y competencia entre laboratorios, permitiendo un rápido desarrollo a partir de la interacción constante entre teoría, experimentos y desarrollos tecnológicos.

En este proyecto se propone tres objetivos fundamentales para la investigación en los próximos tres años. En primer lugar, se trata de afianzar y extender los modelos de simulación de la generación de armónicos y su propagación, y de la generación de pulsos de attosegundo. El constante desarrollo de estos métodos permite la colaboración con grupos experimentales de vanguardia en la propuesta de nuevos esquemas y técnicas que les sitúan en la frontera del conocimiento. El continuo desarrollo de éstas estimula a introducir nuevos escenarios, como por ejemplo, gases a alta presión, haces con polarizaciones arbitrarias, esquemas de doble haz, etc. Entre ellos cabe resaltar la generación de armónicos de orden alto con vórtices de luz, tema del que son pioneros en la aproximación teórica. El segundo objetivo se centra en la descripción de la propagación no lineal del campo infrarrojo que genera los armónicos y los pulsos de attosegundo. El estudio de la variación de estos pulsos de femtosegundo durante la propagación es fundamental para describir situaciones realistas de generación de armónicos y, por tanto, es un objetivo muy interrelacionado con el primero. Exploraran también nuevos escenarios de propagación basados en fibras de cristal fotónico, especialmente "twisted fibres", como las que están siendo desarrolladas actualmente en el Instituto Max Planck (Erlangen). Finalmente, plantearan un tercer objetivo que representa una exploración de recientes escenarios cuya aplicación a la generación de armónicos y pulsos de attosegundo está todavía muy poco investigada. En este sentido, las propiedades ópticas inusuales del grafeno y de las resonancias plasmónicas en puntas nanométricas ofrecen un apasionante punto de ruptura con respecto a las aplicaciones tradicionales de los pulsos ultracortos.

Financiación del proyecto

Programa Estatal de Fomento de la Investigación Científica y Técnica de Excelencia, Subprograma Estatal de Generación de Conocimiento, en el marco del Plan Estatal de Investigación Científica y Técnica y de Innovación 2013-2016. Convocatoria 2016, www.idi.mineco.gob.es

Participantes del proyecto

Grupo ALF-USAL (UIC016), Universidad de Salamanca, <https://laser.usal.es>, www.usal.es

Ministerio de Economía, industria y competitividad, www.mineco.gob.es

Periodo de ejecución

Del año 2016 al 2019.

Justificación del proyecto

En las últimas décadas, el desarrollo de la ciencia y de la tecnología de los pulsos láser ultracortos ha sido espectacular. El grupo desarrolla su actividad en este ámbito, contando con un equipo teórico en la vanguardia para las simulaciones de los procesos de interacción láser-materia, que ofrece apoyo teórico a grupos experimentales, tanto del ámbito nacional como internacional. La colaboración con los grupos experimentales es estratégica puesto que amplifica el impacto de nuestros trabajos. Gracias a ésta han tenido el privilegio de formar parte de los equipos de investigación que lideran el desarrollo de las fuentes coherentes de XUV y rayos X, estando en la primera propuesta de método para la generación de rayos X intensos basados en armónicos, proponiendo nuevos métodos para el control de la polarización y demostrando nuevas vías para la generación de armónicos y su detección. Todos estos resultados publicados en revistas de alto impacto, algunas de ellas, como Science, evaluadas en los ranking internacionales de universidades, lo que redundan en la mejora de la posición de Castilla y León en ellos.

La investigación que proponen permitirá, por un lado, seguir posicionando sus modelos y métodos entre los más eficientes en el panorama internacional, y por tanto seguir contribuyendo de forma colegiada con grupos experimentales a la ampliación de las fronteras actuales de la femto- y atto-física. Por otro lado, exploran nuevos escenarios desde el punto de vista teórico, con el fin de elaborar propuestas originales que signifiquen puntos de ruptura con los métodos actuales y, por tanto, con un fuerte potencial innovador.

Funciones de SCAYLE

La forma más precisa de calcular armónicos de orden alto (cuyas siglas en inglés son HHG), es mediante la resolución numérica de la ecuación de Schrödinger. Sin embargo, es bien conocido que para obtener soluciones de esta ecuación, las cuales recrean las interacciones luz láser-materia que dan lugar a la generación de armónicos de alta frecuencia, se necesitan ingentes cantidades de memoria. Dichos recursos hoy en día sólo están disponibles en supercomputadores, donde muchas máquinas comunicadas entre sí cooperan para la realización de una tarea. En este caso, el tamaño de los cálculos viene dado por la longitud de onda de la luz láser utilizada en la interacción: a mayor longitud de onda, mayor la memoria requerida para el cálculo. Por esta razón, el cálculo de interacciones con láseres de longitudes de onda en el infrarrojo medio (entorno a unas dos micras), es

especialmente problemático, porque los recursos computacionales necesarios llegan incluso al límite de capacidad de los actuales supercomputadores, convirtiendo esta tarea en irrealizable de forma práctica.

Por ello se hace indispensable utilizar modelos simplificados que permitan reducir el coste computacional de este tipo de cálculos. Para este fin nos servimos de la SFA+, una extensión de la aproximación de campo fuerte (SFA son sus siglas en inglés) en su versión convencional. De esta manera conseguimos reducir en horas, incluso minutos, el tiempo de cálculo que de otra manera, utilizando la ecuación de Schrödinger, nos podría llevar semanas.

Una vez que obtenemos el espectro de armónicos para un sólo átomo, pasamos a calcular la propagación de estos armónicos en un medio (por ejemplo un gas) sirviéndonos de las ecuaciones del electromagnetismo de Maxwell. Considerando cada átomo como una fuente de radiación, podemos sumar coherentemente las contribuciones de todos los átomos para finalmente obtener la radiación total.

Este esquema de resolución del problema nos permite plantear estrategias de distribución y paralelización del cálculo en supercomputadores, en la que cada procesador pueda calcular independientemente el espectro de armónicos de un átomo en particular, para posteriormente comunicar todas las contribuciones locales y obtener el resultado final. Es aquí donde el uso intensivo de los recursos de SCAYLE es crucial para la obtención de los cálculos, y por lo tanto para la continuidad de la investigación.



Código FIS2016-75652-P

Líder del proyecto

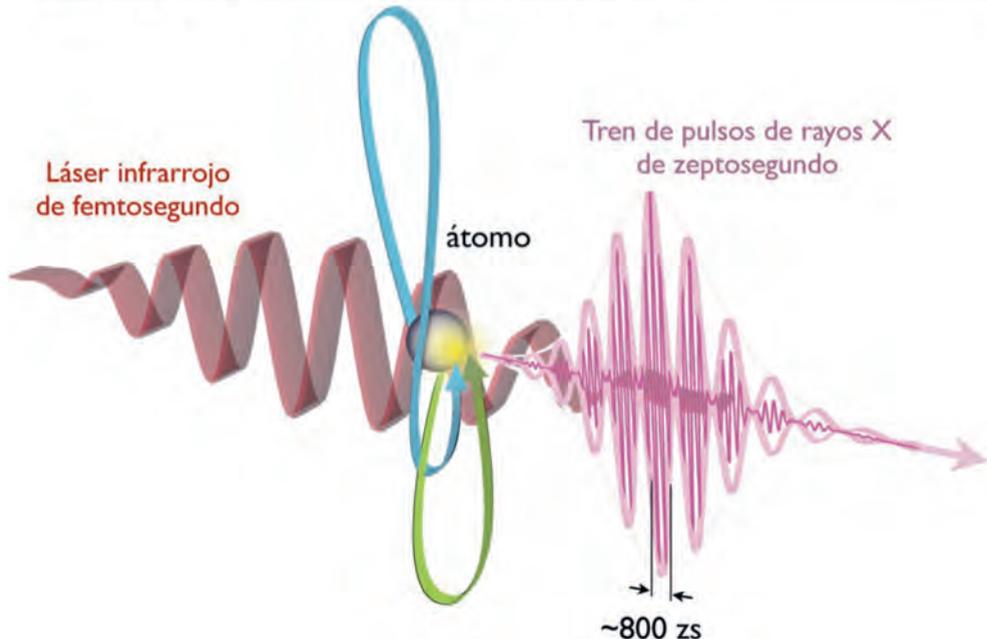
GRUPO ALF-USAL (UIC016), <https://laser.usal.es>

Los miembros del grupo ALF-USAL (UIC016) llevan desarrollando sus investigaciones en torno a los láseres de potencia de pulso corto desde 1992. Desde esa fecha la Universidad de Salamanca ha desarrollado un papel de creciente relevancia en el ámbito internacional. El grupo está compuesto de 7 investigadores sénior, cada uno reconocido con dos o más sexenios de investigación continuada. La participación en el grupo de profesores de varios departamentos (Física Aplicada e Ingeniería Mecánica) introduce un carácter interdisciplinar que permite abordar su ámbito de investigación desde las diferentes perspectivas de la I+D+i (fundamental, aplicada y tecnológica), sin descuidar la transferencia de tecnología y servicios a otros centros y empresas. Por otro lado, como profesores de universidad ponen un esfuerzo especial en que su tarea científica revierta en la formación de los jóvenes, a todos los niveles de educación universitaria (grado, máster y doctorado), así como en la captación de talento pre-universitario y en la divulgación en general.

Entre sus líneas de investigación, orientadas siempre al estudio de los láseres de pulso ultracorto y alta potencia, se encuentran:

- Propagación no lineal de pulsos ultracortos (filamentación y postcompresión).
- Técnicas de caracterización de pulsos ultracortos.
- Procesos paramétricos con pulsos ultracortos.
- Generación de armónicos de orden elevado.
- Ablación de materiales con pulsos ultracortos.
- Fabricación de dispositivos fotónicos mediante escritura láser.
- Microprocesado láser 3D.
- Fotónica aplicada a la biomedicina (transiluminación).

Colaboran con numerosos grupos del panorama internacional, resaltando el Prof. Feng Chen (State Key Laboratory on Crystal Materials, Shandong University, China), que lidera uno de los grupos líderes mundiales en fotónica integrada en cristales no lineales, o el grupo del Prof. Helder Crespo (Universidade do Porto), reconocido por su método de caracterización d-Scan.



Esquema de generación de pulsos de attosegundo y zeptosegundo.

Nuevas **tecnologías** de **secuenciación** (NGS) para el estudio de los **micovirus** en *Fusarium circinatum*

Objetivo del proyecto

El objetivo principal del proyecto es utilizar las nuevas tecnologías de secuenciación para:

- *Identificar las moléculas víricas existentes en los aislamientos de *F. circinatum*,*
- *estudiar el efecto de los micovirus sobre *F. circinatum*, y*
- *estudiar el efecto de la inoculación de *F. circinatum* con y sin micovirus sobre plántulas de pino.*

La consecución de estos objetivos clarificará las posibilidades de esta estrategia para el control la enfermedad del Chancro Resinoso del Pino en España.

Periodo de ejecución

Desde el año **2016** al **2019**.

Financiación del proyecto

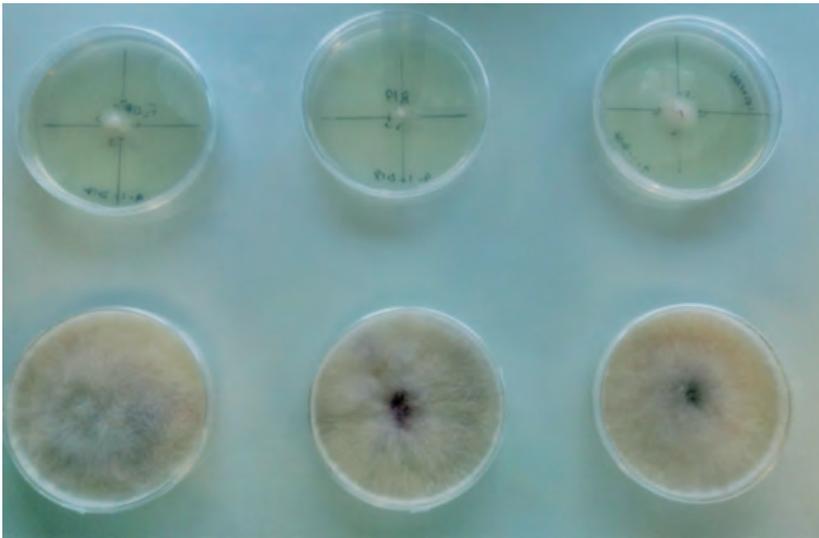
Convocatorias 2015, Proyectos EXCELENCIA y Proyectos RETOS, Dirección General de Investigación Científica y Técnica, Subdirección General de Proyectos de Investigación, MINECO, www.mineco.gob.es

Entidades Participantes del proyecto.

Universidad de Valladolid, España, www.uva.es

Natural Resources Institute, Luke, Finlandia, www.luke.fi/en

Forestry and Agricultural Biotechnology Institute, FABI, Sudáfrica, www.fabinet.up.ac.za



Cepas del hongo *Fusarium circinatum* creciendo en medio de cultivo para la posterior extracción de ARN/E.
Jordán Muñoz Adalia.

Justificación del proyecto

Fusarium circinatum Nirenberg & O'Donnell (teleomorfo = *Gibberella circinata*) es un hongo ascomiceto causante del Chancro Resinoso del Pino (PPC, de sus iniciales en inglés), una enfermedad que causa importantes daños y pérdidas en bosques y plantaciones de pinos, siendo además considerado el patógeno más importante de viveros forestales en varios países de todo el mundo.

La enfermedad del PPC llegó a Europa en el año 2004, cuando España notificó la presencia del patógeno en varios viveros de pinos en Asturias, y en plantaciones de Cantabria. Desde entonces, diversos estudios científicos han evaluado la diversidad genética de las poblaciones del hongo, su patogenicidad, y el papel de algunas especies de insectos como posibles vectores de la enfermedad. Recientemente, el grupo de investigación ha detectado la presencia de diferentes moléculas de RNA de origen vírico en algunos aislamientos de la población española de *F. circinatum*, tres cepas que pertenecen a dos nuevas especies del género *Mitovirus*, *Fusarium circinatum mitovirus 1* (FcMV1), FcMV2- 1 y FcMV2-2. La detección ha sido frecuente en la población española del patógeno, donde un 32% de los aislamientos estaban infectados por al menos una de las tres cepas víricas detectadas. Estas moléculas de RNA podrían desencadenar procesos hipovirulentos en el hongo hospedante, como ocurre en *Cryphonectria parasitica*, el agente causal del chancro del castaño, que es controlado eficazmente con este método en el monte. Sin embargo, las primeras pruebas de inoculación llevadas a cabo parecen descartar este efecto, lo que nos obliga a la búsqueda de nuevos micovirus. Las técnicas de detección de virus rutinariamente utilizadas, como la reacción en cadena de la polimerasa con transcriptasa inversa (RT-PCR) requieren de un conocimiento previo del virus para ser detectado, y por tanto no son aconsejables para la detección de micovirus desconocidos en hongos.

La utilización de la metagenómica, apoyada en las nuevas tecnologías de secuenciación (NGS, de sus iniciales en inglés), posibilita determinar toda la comunidad vírica de un organismo, incluyendo micovirus.

Funciones de SCAYLE

La capacidad de supercomputación de Caléndula ha sido requerida para el manejo de las secuencias procedentes de secuenciación masiva de ARN de cadena corta.

Así mismo, Caléndula ha servido para el ensamblaje de estas lecturas para su posterior análisis.

Líder del proyecto

Julio Javier Diez Casero es Catedrático de Patología Forestal, de la Escuela de Ingenierías Agrarias, de la Universidad de Valladolid. Lleva más de 20 años ejerciendo como profesor en materia de enfermedades forestales. Realizó su doctorado en la Universidad Politécnica de Madrid y la Licenciatura de Biología en la Universidad de Salamanca. Ha publicado más de 130 artículos en revistas científicas y participado en numerosos congresos nacionales e internacionales. Ha dirigido 15 tesis doctorales y más de 60 tesis de máster y proyectos fin de carrera. Ha coordinado numerosos proyectos técnicos y de investigación, nacionales e internacionales y actualmente lidera la acción COST FP1406 *Pine pitch canker - strategies for management of Gibberella circinata in greenhouses and forests (PINESTRENGTH)*.



AGL2015-69370-R

Análisis de múltiples pathways

Objetivo del proyecto

El objetivo general del estudio MCC-Spain es evaluar los factores ambientales y genéticos asociados con los tumores a estudio: cáncer colorrectal, de mama, de estómago, de próstata y las leucemias linfáticas crónicas. Los objetivos específicos son:

- *Evaluar el riesgo de cáncer de cada localización tumoral en relación a exposiciones ambientales y ocupacionales incluyendo contaminantes del agua potable (arsénico, nitratos, cromo, subproductos de cloración), disruptores hormonales y otros contaminantes orgánicos persistentes, así como la disrupción del ritmo circadiano a raíz de los trabajos con turno de noche, y otros factores.*
- *Evaluar el riesgo de los cánceres en relación al consumo de determinados fármacos, incluyendo estatinas y analgésicos.*
- *Evaluar el riesgo de cáncer relacionado con diversos factores como estilos de vida, nutrición y actividad física, historia médica y familiar y otros factores.*
- *Evaluar el riesgo de cáncer de mama y próstata en relación a factores hormonales, infecciones y fenotipos permanentes relacionados con exposiciones ambientales en fases precoces de la vida.*
- *Validar la evaluación de la exposición a agentes químicos ambientales mediante modelos de exposición utilizando biomarcadores de exposición, información individual y medidas de exposición ambiental.*
- *Evaluar, en una primera fase, una serie limitada de genes tanto en relación a efectos principales como en relación a su interacción con factores ambientales.*
- *Almacenar suficiente material biológico para su utilización en el futuro en estudios GWAS (Genome Wide Association Studies) y otros.*



El equipo multidisciplinar que realiza la investigación del MCC-Spain durante la reunión que celebró en León.

Periodo de ejecución

Del año **2016** al **2019**.

Financiación del proyecto

El estudio MCC-Spain está financiado por:

- CIBERESP, www.ciberesp.es
- Financiación del Fondo de Investigación Sanitaria (FIS), <https://portalfis.isciii.es>
- Otras agencias de financiación en determinadas áreas.

Participantes del proyecto

CREAL - ISGlobal, Instituto de Salud Global, www.creal.cat/es

Instituto de Salud Carlos III, www.isciii.es

CIBERESP - Consorcio de Investigación Biomédica en Red de España, www.ciberesp.es

Justificación del proyecto

Investigar la influencia de factores ambientales y su interacción con factores genéticos en tumores frecuentes o con características epidemiológicas peculiares en nuestro país, en las que los factores ambientales implicados no son suficientemente conocidos

Funciones de SCAYLE

La necesidad de utilización de Caléndula es debida a la necesidad de una precisión de potencia estadística y el uso continuado de grandes ordenadores, permitiendo de esta manera la obtención de resultados de una manera más rápida y eficaz.

Líder del proyecto

CREAL - ISGLOBAL, INSTITUTO DE SALUD GLOBAL, www.creal.cat/es

CREAL es un centro público de investigación que abarca una amplia gama de exposiciones ambientales como el agua, la radiación o la contaminación del aire, y efectos en la salud tanto salud infantil, cáncer como enfermedades respiratorias.

INSTITUTO DE SALUD CARLOS III, www.isciii.es

El instituto Carlos III es el principal Organismo Público de Investigación (OPI), que financia, gestiona y ejecuta la investigación biomédica en España. Con una trayectoria de más de 20 años de investigación en ciencias de la vida y de la salud y prestación de servicios de referencia, es además el organismo gestor de la Acción Estratégica en Salud (AES) en el marco del Plan Nacional de I+D+I.



El último proyecto concedido de Fondo de Investigaciones Sanitarias: PI15/01032.

Supervivencia de cáncer colorrectal en el proyecto MCC-Spain: Modelos de predicción que integran datos genéticos y clínico-epidemiológicos.

AMIGA6, gas in and around galaxies. Preparation for SKA science and contribution to the design of the SKA data flow

Objetivo del proyecto

AMIGA6 se basa en los resultados de proyectos anteriores de AMIGA y es parte del trabajo preparatorio que AMIGA realiza para la explotación científica de SKA. Los ciclos de vida HI, tanto en galaxias aisladas como en grupos densos, todavía son poco conocidos, ya que sus densidades bajas de columna solo son alcanzables por el SKA.

En AMIGA6 nuestros principales objetivos son: refinar modelos de acreción de gas frío utilizando galaxias aisladas, y analizar el papel de la eliminación de marea HI en el SF suprimido en los Grupos Compactos de Hickson.

Como preparación para el desafío que supondrá la explotación de datos de SKA, complementan la ciencia fundamental con la investigación aplicada en 3 paquetes de trabajo de SKA, contribuyendo así a los consorcios de Big-Data SKA.

AMIGA6 está liderado por el coordinador de la participación española en el SKA, y reúne a todos los grupos españoles involucrados en el flujo de datos del SKA: señalización y transporte de datos (SaDT), procesador de señal central (CSP) y procesador de datos científicos (SDP).



Primer plano de las matrices de apertura de baja frecuencia de SKA y platos ASKAP en Australia. © Organización SKA.

Participantes del proyecto

Subproyecto 1 (coordinador)

- Instituto de Astrofísica de Andalucía, www.iaa.csic.es
- Supercomputación Castilla y León, www.scayle.es

Subproyecto 2:

- Universidad Politécnica de Madrid, www.upm.es

Subproyecto 3:

- Univeridad de Granada, www.ugr.es
- Universidad Politécnica de Valencia, www.upv.es
- Real Instituto y Observatorio de la Armada, www.armada.mde.es/roa

Financiación del proyecto

Ministerio de Economía y Competitividad. Convocatoria 2015 - Proyectos I+D+I - Programa Estatal de Investigación, Desarrollo e Innovación Orientada a los Retos de la Sociedad, MINECO, www.mineco.es

Justificación del proyecto

AMIGA6 está liderado por la coordinadora de la participación de España en SKA, y reúne a los grupos españoles participantes en el flujo de datos, diseñado por Consorcios Internacionales Big Data SKA: Signal and Data Transport (SaDT), Central Signal Processor (CSP) y Science Data Processor (SDP, donde está incluido SCAYLE). AMIGA6 contribuye a una distribución eficiente de datos/metadatos entre la comunidad, colaborando en la transferencia de conocimiento del CSP al SDP en eficiencia energética..

Funciones de SCAYLE

Con el conocimiento de SCAYLE en algoritmos de análisis de datos, la monitorización y el control del consumo de energía en los grandes sistemas de HPC, se pretende desarrollar aplicaciones para identificar y proponer los modelos predictivos más apropiados y la optimización dinámica estrategias para el SDP, ya que asociado a recursos de HPC se logran grandes reducciones en el consumo de energía.

Periodo de ejecución

1 de enero de **2016** al 31 de diciembre de **2018**.

Líder del proyecto

El proyecto está liderado por la Dra. Dña Lourdes Verdes-Montenegro Atalaya.

Sus principales temas de investigación científica en los últimos años han sido el estudio del medio interestelar en la evolución de las galaxias, analizando dos ambientes extremos: galaxias aisladas y grupos compactos, enfocados en el estudio ISM de Grupos Compactos de Hickson en diferentes etapas evolutivas, con un énfasis especial en grupos fuertemente deficientes en HI.

Actualmente, es la única astrónoma española experta en el estudio de galaxias cercanas mediante observaciones de gases atómicos.

Su trabajo en grupos compactos resolvió las aparentes contradicciones sobre su formación estelar, demostrando que no está mejorado (Verdes-Montenegro et al 1998; 75 citas). Comprender este resultado sigue siendo tema de estudio de grupos internacionales a través de programas en, por ejemplo, XMM, Chandra, Herschel, Spitzer, CARMA, ALMA, utilizando como referencia la secuencia evolutiva propuesta en Verdes-Montenegro et al (2001; 191 citas).

Además es miembro de la colaboración internacional AMIGA (Análisis del medio interestelar en galaxias aisladas, <http://amiga.iaa.es>, Verdes-Montenegro et al 2005, 82 citas) involucrando desde 2003 a más de 15 grupos de investigación en el exterior. También realizó un esfuerzo especial para complementar la investigación básica con actividades destinadas a fortalecer la participación de España en proyectos instrumentales.



AYA2015-65973-C3-1-R

Carácter multifactorial de los polifenoles: oportunidad para el desarrollo de herramientas terapéuticas frente a obesidad y enfermedades infecciosas

Objetivo del proyecto

Uno de los objetivos de este proyecto es hacer un cribado "in silico" de posibles moduladores de la actividad biológica de las proteínas AMPK, PPARalpha, beta y gamma, y de la proteína de unión a la penicilina (PBP2a). Adicionalmente se quieren cribar los compuestos polifenólicos frente a distintas polimerasas de virus humanos.

Periodo de ejecución

Del año 2016 al 2018.

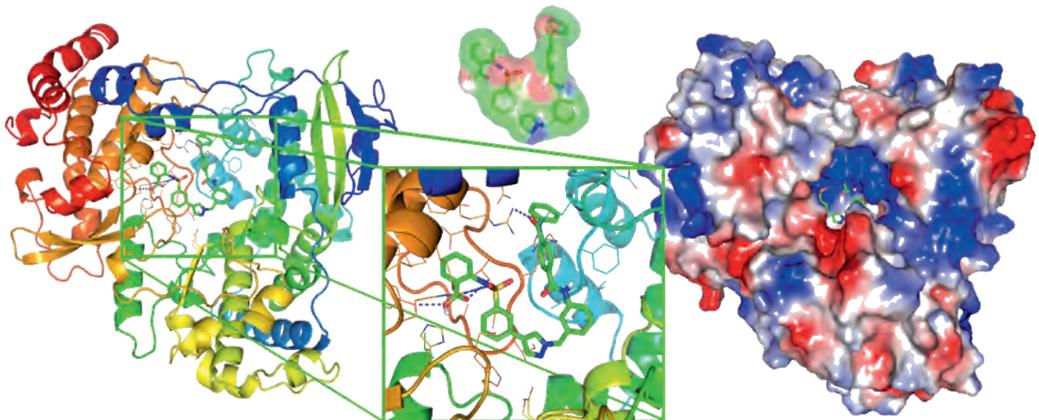
Financiación del proyecto

Financiación propia del Grupo de Investigación.

Participantes del proyecto

Instituto de Biología Molecular y Celular,
<http://ibmc.umh.es>

Departamento de Física y Arquitectura de
Computadores de la Universidad Miguel Hernandez
de Elche,
www.umh.es/contenido/pas/:uor_33/datos_es.html



Ejemplo de resultados de experimentos de acoplamiento molecular de un inhibidor de la RNA-polimerasa dependiente de RNA del virus del Dengue.

Justificación del proyecto

Las plantas generan una gran diversidad de compuestos que presentan una acción farmacológica múltiple alcanzando una gran variedad de dianas moleculares. Entre ellos, los compuestos polifenólicos han demostrado propiedades potenciales en la prevención y el tratamiento de enfermedades no transmisibles y también de enfermedades infecciosas. En el presente proyecto se plantean dos aproximaciones multidisciplinares para la búsqueda de herramientas terapéuticas a base de mezclas de compuestos polifenólicos dirigidos a dos áreas de máximo interés en la salud humana: la prevención y/o tratamiento de patologías asociadas a la obesidad y el tratamiento de infecciones nosocomiales. Para el desarrollo de ingredientes enfocados a la obesidad, se recurrirá al cribado "*in silico*" de compuestos polifenólicos de librerías virtuales que sean moduladores de las dianas moleculares AMPK y/o PPAR que luego serán validados en ensayos celulares, bien utilizando compuestos puros o extractos enriquecidos, buscando finalmente una actividad sinérgica entre los mejores candidatos. Por otra parte, para el desarrollo de ingredientes antimicrobianos contra bacterias resistentes (*Staphylococcus aureus* resistente a meticilina, se partirá de datos previos obtenidos por el grupo de investigación sobre el efecto de compuestos polifenólicos puros conocidos y extractos polifenólicos en bacterias no resistentes que se ensayarán en un panel de 100 cepas resistentes hospitalarias. Se realizará un estudio "*in silico*" de la interacción molecular de estos compuestos y dos dianas moleculares: la membrana de *S. aureus* y la proteína de unión a penicilina 2a (PBP2a), responsable de la resistencia en cepas estafilocócicas. En el mismo sentido, se quieren cribar los compuestos polifenólicos frente a distintas polimerasas de virus humanos..

Funciones de SCAYLE

La infraestructura de HPC de SCAYLE ha permitido realizar el cribado virtual de librerías de unos 300.000 compuestos sobre sitios activos de RNA polimerasas dependientes de RNA de diversos virus humanos. Se ha ejecutado software no comercial y programas en lenguaje Python desarrollados por los autores del proyecto.

Líder del proyecto

El IBMC, Instituto de Biología Molecular y Celular, <http://ibmc.umh.es>, se constituye como Instituto Universitario de Investigación de la Universidad Miguel Hernández de Elche en 2002. Su mayor éxito ha sido la comercialización de productos innovadores generados por los proyectos de investigación en los campos de nutraceuticos, cosmecéticos y biotecnología, mediante el desarrollo de dos líneas principales de investigación: "Diseño Molecular y Celular", destacando estudios en Reconocimiento molecular e ingeniería y biofísica de proteínas, Relaciones estructura-función en proteínas de membrana, y Plataforma de modelado molecular. La otra línea es "Diagnóstico y Terapia Moleculares" investigando en materias como son la Moléculas bioactivas, Estrategias antivirales y Oncología molecular y celular.

Además también ofrecen servicios como la Plataforma de cribado biológico de alto rendimiento (HTS), Cultivos celulares eucariotas y procariotas, Análisis celular, desarrollo de métodos de cribado de alto rendimiento, Análisis microcalorimétrico, espectroscopia molecular, Centrifugación y ultra centrifugación, Análisis y fraccionamiento cromatográfico, Procesamiento de ácidos nucleicos, Adquisición y tratamiento de imágenes, Laboratorio de zumos y planta piloto anexa (Campus de Orihuela), Ultracongelación y criopreservación de muestras biológicas, Bioinformática, Agua, lavado y esterilización.

Tienen proyectos activos financiados por convocatorias europeas y nacionales. También reciben financiación privada de empresas con las que colaboran, afianzando el compromiso de transferencia de tecnología que tiene, facilitando la traslación de la ciencia básica al mundo productivo y clínico, hecho que les lleva a conseguir sus éxitos en la comercialización de los productos desarrollados.

Ajuste, Validación e Implantación del Modelo Físico PhFFS de Simulación de Incendios Forestales

Objetivo del proyecto

El propósito principal de este proyecto de investigación es mejorar, validar y ajustar el modelo físico de propagación de incendios forestales PhFFS (Physical Fast Fire Simulator) desarrollado por el grupo de investigación en Simulación Numérica y Cálculo Científico SINUNCC, de la Universidad de Salamanca, así como transferir la tecnología resultante a los sectores públicos y privados dedicados a la prevención, seguridad y gestión de incendios forestales.

El proyecto se concreta a través de los siguientes objetivos y actividades:

- *Mejora del modelo PhFFS mediante la incorporación de la fase gaseosa y el efecto de la radiación solar sobre el combustible.*
- *Mejora de eficiencia de las técnicas de asimilación de datos que permiten la incorporación de datos reales de forma interactiva.*
- *Identificación de las características de los combustibles forestales relacionadas con los parámetros del modelo PhFFS.*
- *Recopilación y análisis de casos reales de incendios forestales para su posterior uso en la validación del modelo PhFFS.*
- *Simulación mediante el modelo PhFFS mejorado tanto de casos experimentales como reales de incendios para la puesta en práctica del modelo, el ajuste de sus parámetros y la asimilación de datos a través de casos prácticos.*
- *Comparación del modelo PhFFS mejorado con modelos de propagación de incendios de tipo semi-empírico.*
- *Integración del modelo PhFFS mejorado en sistemas de información geográfica y visualización a través de plataforma on-line.*

Periodo de ejecución

21 de marzo de 2016 al 30 de junio de 2018.

Participantes del proyecto

Grupo de Investigación de Simulación Numérica y Cálculo Científico de la Universidad de Salamanca. Reconocido como Unidad de Investigación Consolidada de la Junta de Castilla y León UIC018, <https://diarium.usal.es/sinumcc>

Financiación del proyecto

Subvención del programa de apoyo a proyectos de investigación de la Conserjería de Educación de la Junta de Castilla y León, cofinanciadas por el Fondo Europeo de Desarrollo Regional (FEDER), al amparo de la orden de 25 de febrero de 2016 (tramo 2016-2018), www.educa.jcyl.es

Justificación del proyecto

Todos los informes internacionales sobre incendios forestales (FAO, EFFIS) señalan que los incendios forestales suponen una amenaza actual y futura para los ecosistemas del planeta. Estos informes inciden en el aumento del riesgo de incendios forestales debido a las anomalías climáticas que se vienen produciendo recientemente, en cuanto a temperaturas más elevadas y sequías más severas. El cambio climático pone de nuevo en primer término el riesgo que suponen para nuestros bosques los incendios forestales, puesto que además, los sistemas forestales son a su vez fundamentales para la reducción de los gases del efecto invernadero, y por tanto son un punto clave en el reto de frenar el cambio climático.

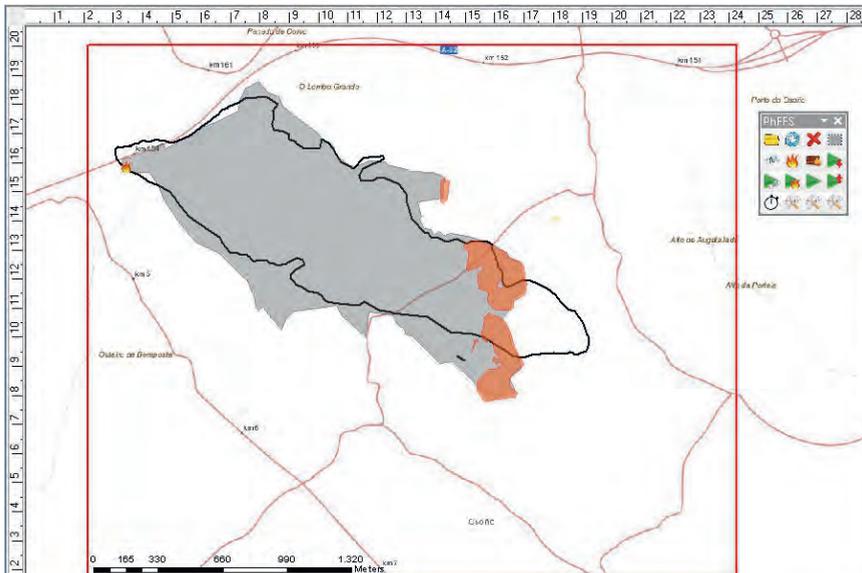
España es uno de los países de la cuenca mediterránea más afectados por los incendios forestales. Según fuentes del Ministerio de Agricultura, Alimentación y Medio Ambiente, en la primera década del siglo XXI, el fuego arrasó más de un millón de hectáreas de superficie forestal en nuestro país; y en el año 2012, especialmente dramático, más de 200.000 hectáreas de superficie forestal, de las cuales más de 40.000 hectáreas pertenecían a Castilla y León.

Los esfuerzos para luchar contra los incendios forestales deben reforzarse a todos los niveles, desde todas las administraciones pero también desde todos los sectores que puedan aportar una

mejora en el conocimiento de los incendios forestales y por lo tanto en la lucha para atajar este problema. En este sentido, el sector científico, con la colaboración de investigadores de diversas áreas de conocimiento, y la participación del sector tecnológico y empresarial, junto con las administraciones públicas, puede proporcionar importantes avances.

Desde el punto de vista de la investigación científica, la modelización matemática se ha convertido hoy en día en una herramienta esencial para el análisis y la predicción de un gran número de fenómenos físicos, en particular el tema que compete a este proyecto sobre modelización de la propagación de incendios forestales.

El aumento de la potencia de cálculo de los sistemas informáticos actuales, así como la mejora de las capacidades de la actual tecnología de información espacial (sistemas de información geográfica, sistemas de detección remota), ofrecen un enorme potencial para la simulación efectiva del comportamiento de los incendios forestales, permitiendo mejorar tanto los modelos de propagación del fuego como los modelos de combustibles.



Comparación de la simulación proporcionada por el modelo PhFFS (gris-área quemada, rojo-frente activo) y el perímetro real (línea negra) del incendio ocurrido en Osoño (Orense) en agosto de 2009.

Líder del proyecto

GRUPO SINUMCC (UIC018), <https://diarium.usal.es/sinumcc>, el grupo SINUMCC es un grupo de Investigación reconocido de la Universidad de Salamanca desde 2005, y Unidad de Investigación Consolidada de La Junta de Castilla y León desde 2015, y está formado por 5 investigadores senior de la Universidad de Salamanca y 2 investigadores senior de la Universidad de las Palmas de Gran Canaria, que suman 20 sexenios de investigación. Además actualmente hay 2 contratados a cargo de diversos proyectos de investigación y contratos con empresas.

La investigación del grupo se centra en diversos aspectos de la simulación numérica y el cálculo científico, con especial interés en las aplicaciones a problemas medioambientales. En concreto se han desarrollado diversos modelos de simulación: campos de viento, incendios forestales, dispersión de contaminantes y radiación solar.

El grupo trabaja habitualmente en colaboración con otros grupos de investigación nacionales como el Instituto Universitario de Sistemas Inteligentes y Aplicaciones Numéricas en Ingeniería de la Universidad de Las Palmas de Gran Canaria y el Laboratorio de Cálculo Numérico de la Universidad Politécnica de Barcelona; e internacionales como Ricardo H. Nochetto del Institute for Physical Science and Technology de la University of Maryland, USA; y Erwin Hernández del Grupo de Análisis y Modelamiento Matemático Valparaíso de la Universidad Técnica Federico Santa María de Chile.

Funciones de SCAYLE

Algunos aspectos clave para la mejora del modelo de simulación de propagación de incendios forestales PhFFS son el análisis de sensibilidad del modelo, el ajuste de sus parámetros y la utilización de técnicas de asimilación de datos. Todos estos aspectos requieren de la ejecución de miles de simulaciones dentro de un tiempo razonable y es ahí donde el supercomputador de SCAYLE juega un papel fundamental. Caléndula permite realizar los cálculos requeridos por el modelo utilizando recursos en paralelo y con un gran número de procesadores con altas prestaciones de memoria. Gracias a SCAYLE se está consiguiendo obtener resultados de forma rápida y eficiente que permiten mejorar el modelo PhFFS en los aspectos señalados, y en consecuencia mejorar sustancialmente la ejecución del proyecto.



Código SA020U16

Cirugía Oncológica: modelo de gestión y formalización del conocimiento (casos de estudio cáncer colorrectal - CCR)

Objetivo del proyecto

En el proyecto se plantean dos metas:

- La creación de un sistema de gestión de pacientes específico para las unidades médicas de coloproctología.
- El análisis posterior usando inteligencia artificial (IA) y herramientas estadísticas.

Ambos objetivos están orientados a aumentar la supervivencia y disminuir las complicaciones de los pacientes con cáncer de colon.

Hasta el momento, la gestión de los pacientes se lleva a cabo con varias aplicaciones, pero ninguna se ajusta a las necesidades concretas de las unidades de coloproctología, lo cual hace que se pierda información muy significativa para el equipo médico al no encontrarse correctamente vinculada. Debido a ello surge la necesidad de la creación de un sistema de gestión de pacientes específico para las unidades de coloproctología de toda Castilla y León que ayude a la atención de los pacientes con cáncer de colon.

Este sistema permitirá la creación de una base de datos consistente, usable y útil para facilitar la práctica diaria de la unidad médica de coloproctología. Los datos recabados se usarán en la segunda fase del proyecto para estudiar la supervivencia a largo plazo de los tratamientos quirúrgicos aplicados al cáncer de colon, junto con las complicaciones derivadas de las mismas. De esta forma se mejora la ayuda al modificar las tasas de supervivencia y de complicaciones usando herramientas de última generación que hagan uso de la inteligencia artificial sobre los datos almacenados de los pacientes.

El sistema de gestión de pacientes será una aplicación basada en tecnologías Web 2.0, al que se podrá acceder por medio de cualquier navegador web, permitiendo que pueda ser usado desde distintos dispositivos con acceso a internet (ordenadores, móviles y tablets), con lo cual se obtendrá un sistema que soportará muchos usuarios concurrentes y en multiplataforma.

El objeto del convenio es establecer un marco de colaboración en el que SCAYLE presta servicios y recursos para el desarrollo de diferentes proyectos de informática médica que requieren el acceso a datos de carácter personal contenidos en ficheros automatizados pertenecientes al Complejo Asistencial Universitario de León, (CAULE).



Código FI4VDI-SOE4/P3/E804

Periodo de ejecución

Fase I. Desarrollo de sistema de gestión de pacientes, iniciado en 2011 y en fase de ejecución.

Fase II. Adaptación de la aplicación al entorno *cloud*, julio 2013 y en fase de ejecución.

Financiación del proyecto

Fase I. Desarrollo de sistema de gestión de pacientes, financiación propia de SCAYLE.

Fase II. Adaptación de la aplicación al entorno *cloud* e integración en el plan de sistemas (Red de Infraestructuras Federadas), Financiación a través de Fondos FEDER en el marco del Proyecto FI4VDI-SOE4/P1/E804 y cofinanciado por SCAYLE - Consejería de Fomento y Medio Ambiente de la Junta de Castilla y León.

Participantes del proyecto

Complejo Asistencial Universitario de León, (CAULE)

Departamento de Ciencias Biomédicas (Área de Medicina Preventiva y Salud Pública) de la Universidad de León, <http://www3.unileon.es/dp/e/>

Supercomputación de Castilla y León (SCAYLE), www.scayle.es

Red de Centros - Red de Infraestructuras Federadas, Proyecto FI4VDI-SOE4/P1/E804

Justificación del proyecto

Según datos del Centro Nacional de Epidemiología del Instituto de Salud Carlos III, en España hay 25.000 casos nuevos y 13.000 defunciones por CCR al año (López Abente, 2005).

En cuanto a la distribución municipal, según los datos del mismo Instituto, hay un núcleo de municipios de mayor riesgo en Cataluña (Barcelona y Gerona), en Baleares, en Castilla y León (León y Palencia) y en Andalucía (municipios de Sevilla y Cádiz), y aunque la tendencia a la supervivencia ha mejorado debido al diagnóstico más temprano, la mayor eficacia de los tratamientos adyuvantes y neoadyuvantes y un abordaje quirúrgico más agresivo con las recidivas locales y metástasis, el CCR sigue siendo un problema de salud pública y de consumo de recursos de primera magnitud.

En el Complejo Asistencial Universitario de León, según el registro de tumores del Complejo, se diagnostican algo más de 350 CCR al año y en la actividad del Servicio de Cirugía, más de 400 intervenciones están motivadas por el CCR.

Todo lo anterior ha motivado la puesta en marcha de un proyecto cuyo objetivo es el diseño y desarrollo de un Sistema de gestión, predicción y análisis inteligente de cáncer colorrectal en Castilla y León.



Hospital de León.

Líneas futuras de desarrollo

Dentro del marco del proyecto-convenio se establecen como desarrollo de líneas futuras el estudio de la obesidad, actividad física y adherencia a la dieta mediterránea en el cáncer colorrectal en el área de salud de León, y su interacción con variantes y vías genéticas (*pathways*) relacionadas con el metabolismo. Será imprescindible el apoyo de Caléndula en cuanto a capacidad de almacenamiento, facilitando la creación de una base de datos y sus copias de seguridad, y en cuanto a potencia de cálculo mediante la ejecución de las herramientas de inteligencia artificial y de análisis estadístico.

Funciones de SCAYLE

La potencia de cálculo y el almacenamiento de Caléndula permitirán ejecutar el servidor web y el motor de la base de datos. La gran capacidad de almacenamiento del superordenador posibilitará alojar la base de datos, y facilitará la creación de copias de seguridad del sistema de gestión de pacientes, garantizando la seguridad de los datos, su disponibilidad y consistencia.

En la etapa del estudio de la supervivencia y de las complicaciones, el poder de cálculo de Caléndula será imprescindible para ayudar a ejecutar con éxito las herramientas de inteligencia artificial y de análisis estadístico.

Integración en el plan de sistemas desarrollado en el marco del Proyecto FI4VDI-SOE4/P1/E804.

Líder del proyecto

COMPLEJO ASISTENCIAL UNIVERSITARIO DE LEÓN (CAULE). El Complejo Asistencial Universitario de León, www.saludcastillayleon.es/CHLeon/es, está integrado por diversos centros asistenciales, como son el Hospital de León -Complejo Sanitario integrado por los Hospitales Virgen Blanca de la Seguridad Social, el Princesa Sofía de la Diputación Provincial de León y el edificio San Antonio Abad-, la Residencia de Santa Isabel y el Hospital Monte San Isidro.

Lidera el proyecto la dirección del Servicio de Cirugía del CAULE.

Colaboraciones en proyectos I+D+i a través de la RES

First-principles modelling of irradiation effects on light species retention and diffusion in materials relevant for nuclear technologies (2nd period) - QCM-2017-3-0018

Grupo líder del proyecto: University of Oviedo

CO2 adsorption and activation by early transition metal oxycarbides - QCM-2017-3-0002

Grupo líder del proyecto: University of Barcelona

Effects of time-periodic spanwise fluidic actuation on the flow around a cylinder - FI-2017-3-0009

Grupo líder del proyecto: Universitat Politècnica de Catalunya

Understanding the molecular interactions that allow drug molecules to passively diffuse across the blood brain barrier - BCV-2017-3-0008

Grupo líder del proyecto: King's College London

Insights into the inactivation dynamics of γ -arrestin 1 and 2 - BCV-2018-1-0012

Grupo líder del proyecto: Universidad Pompeu Fabra

Metadynamics calculations of free energy landscapes of tryptophan bound to phospholipid cell membranes - FI-2018-1-0007

Grupo líder del proyecto: Universidad Politécnica de Cataluña

DNS of the impact of a rising drop against a solid circular cylinder - FI-2018-1-0036

Grupo líder del proyecto: Universidad Politécnica de Cataluña

Reversible Breathing Single-Crystal to Single-Crystal Metal-Organic Frameworks: A Theoretical Study - QCM-2018-1-0010

Grupo líder del proyecto: Universidad de Barcelona

Atomic ordering in bimetallic nanoparticles of fcc-type: Pt-Cu - QCM-2018-1-0029

Grupo líder del proyecto: Institució Catalana de Recerca i Estudis Avançats

Modeling of RuO₂ surface-liquid water and RuO₂ nanoparticle-liquid water interfaces: Influence of surface morphology and nanoparticle size - QCM-2018-1-0044

Grupo líder del proyecto: Universidad Autónoma de Barcelona

Rational structure-based design of anti-EGF biological therapeutics - BCV-2018-3-0002

Grupo líder del proyecto: Institute for Research in Biomedicine

Transport of particles and heat in the core of magnetically confined thermo-nuclear fusion plasmas - FI-2018-3-0024

Grupo líder del proyecto: Laboratorio Nacional de Fusión (CIEMAT)

Carbon-based materials as support in heterogeneous catalysis (4th period) - QCM-2018-3-0016

Grupo líder del proyecto: Institute of Chemical Research of Catalonia

Menciones a SCAYLE en publicaciones científicas

Controlling the polarization and vortex charge of attosecond high-harmonic beams via simultaneous spin-orbit momentum conservation

Authors: Kevin M. Dorney, Laura Rego, Nathan J. Brooks, Julio San Román, Chen-Ting Liao, Jennifer L. Ellis, Dmitry Zusin, Christian Gentry, Quynh L. Nguyen, Justin M. Shaw, Antonio Picón, Luis Plaja, Henry C. Kapteyn, Margaret M. Murnane and Carlos Hernández-García.

Journal: Nature Photonics,
www.nature.com/naturephotonics, © 2018
Macmillan Publishers Limited, part of Springer
Nature. All rights reserved.

DOI: <https://doi.org/10.1038/s41566-018-0304-3>

Polarization control of isolated high-harmonic pulses

Authors: Pei-Chi Huang, Carlos Hernández-García, Jen-Ting Huang, Po-Yao Huang, Chih-Hsuan Lu, Laura Rego, Daniel D. Hickstein, Jennifer L. Ellis, Agnieszka Jaron-Becker, Andreas Becker, Shang-Da Yang, Charles G. Durfee, Luis Plaja, Henry C. Kapteyn, Margaret M. Murnane, A. H. Kung and Ming-Chang Chen.

Journal: Nature Photonics,
www.nature.com/naturephotonics, © 2018
Macmillan Publishers Limited, part of Springer
Nature. All rights reserved.

DOI: <https://doi.org/10.1038/s41566-018-0145-0>

On parallel computation of centrality measures of graphs

Authors: J. F. García y M.V. Carriegos.

Journal: The Journal of Supercomputing (2018).

DOI: <https://doi.org/10.1007/s11227-018-2654-5>

ISSN (Online) 1573-0484

Palladium nanodendrites uniformly deposited on the surface of polymers as an efficient and recyclable catalyst for direct drug modification via Z-selective semihydrogenation of alkynes

Authors: José García-Calvo, Patricia Calvo-Gredilla, Saúl Vallejos, José Miguel García, José Vicente Cuevas-Vicario, Gabriel García-Herbosa, Manuel Avella and Tomás Torroba.

Journal: Green Chem., 2018, 20, 3875-3883. © The Royal Society of Chemistry 2018

DOI: 10.1039/c8gc01522h

A nanoscopic approach on benzene-toluene-xylenes extraction by sulfolane

Authors: Gutierrez, A.; Atilhan, M.; Aparicio, S.

Journal: Molecular Liquids 2018, 249, 1039-1046.

DOI: 10.1016/j.molliq.2017.11.112

A theoretical study on mixtures of amino acid-based ionic liquids

Authors: Herrera, C.; Atilhan, M.; Aparicio, S.

Journal: Physical Chemistry 2018, 20, 10213-10223.

DOI: 10.1039/c7cp08533h

Insights on the mixtures of imidazolium based ionic liquids with molecular solvent

Authors: Gutierrez, A.; Atilhan, M.; Alcalde, R., Trenzado, J. L., Aparicio, S.

Journal: Journal of Molecular Liquids 2018, 255, 199-207.

DOI: 10.1016/j.molliq.2018.01.167

On the properties and structure of 2-hydroxyethylammonium formate ionic liquid

Authors: Hosseini, S. M.; Alavianmehr, M.M.; Gutierrez, A.; Khalifeh, R.; Moghadasi, J.; Aparicio, S.
Journal: *Molecular Liquids* 2018, 249, 233-244.

DOI: 10.1016/j.molliq.2017.10.122

Microscopic characterization of mixtures of amino acid ionic liquids and organic solvents

Authors: Herrera, C.; Alcalde, R.; Atilhan, M.; Aparicio, S.
Journal: *Molecular Liquids* 2018, 250, 111-120.

DOI: 10.1016/j.molliq.2017.11.131

Molecular dynamics and experimental characterization of [BMIM][BF4] and [BMIM][PF6] with ether cosolvent binary mixtures

Authors: Trenzado, J.L.; Alcalde, R.; Atilhan, M.; Aparicio, S.
Journal: *Molecular Liquids* 2018, 271, 65-73.

DOI: 10.1016/j.molliq.2018.08.031

Theoretical Study on Molten Alkali Carbonate Interfaces

Authors: Gutierrez, A.; Garroni, S.; Souentie, S.; Cuesta-López, S.; Yakoumis, I.; Aparicio, S.
Journal: *Langmuir* 2018, 34, 13065-13076.

DOI: 10.1021/acs.langmuir.8b02907

Structural Elucidation of Covalent Organic Polymers (COP) and Their Linker Effect on Gas Adsorption Performance via Density Functional Theory Approach

Authors: Aparicio, S.; Atilhan, M.
Journal: *Chemistry Select* 2018, 3, 8294-8305.

DOI: 10.1002/slct.201801849

Molecular Dynamics Simulations of Metal Nanoparticles in Deep Eutectic Solvents

Authors: Atilhan, M.; Aparicio, S.
Journal: *Journal of Physical Chemistry C* 2018, 122, 18029-18039.

DOI: 10.1021/acs.jpcc.8b02582

A Theoretical Study of Oil Desulfuration by Ammonium-Based Deep Eutectic Solvents

Authors: Gutierrez, A.; Atilhan, M.; Aparicio, S.
Journal: *Energy Fuels* 2018, 32, 7497-7507.

DOI: 10.1021/acs.energyfuels.8b01403

Molecular Insights into Benzimidazole-Linked Polymer Interactions with Carbon Dioxide and Nitrogen

Authors: Aparicio, S.; Yavuz, C.T.; Atilhan, M.
Journal: *Chemistry Select* 2018, 3, 3691-3701.

DOI: 10.1002/slct.201800253

Theoretical Study of Low Viscous Ionic Liquids at the Graphene Interface

Authors: Atilhan, M.; Aparicio, S.
Journal: *Physical Chemistry C* 2018, 122, 1645-1656.

DOI: 10.1021/acs.jpcc.7b10434

A theoretical study on lidocaine solubility in deep eutectic solvents

Authors: Gutierrez, A.; Atilhan, M.; Aparicio, S.
Journal: *Chemical Physics* 2018, 20, 27464-27473.

DOI: 10.1039/c8cp05641b

Molecular Modeling Analysis of CO2 Absorption by Glymes

Authors: Gutierrez, A.; Atilhan, M.; Aparicio, S.
Journal: *Physical Chemistry B* 2018, 122, 1948-1957.

DOI: 10.1021/acs.jpcc.7b10276

Accountability in Mobile Service Robots.

Autor/es: Ángel Manuel Guerrero-Higuera¹, Francisco J. Rodríguez-Lera, Francisco Martín-Rico, Jesús Balsa-Comerón, Vicente Matellán.
XIX Workshop en Agentes Físicos (WAF), Extended abstract. Madrid, Universidad Carlos III, noviembre 2018.

DOI: 10.1007/978-3-319-99885-5_17.

Aplicaciones contenedorizadas en entornos HPC. Caso práctico: aprendizaje automático aplicado a ciberseguridad en robots móviles.

Autor/es: Ángel Manuel Guerrero-Higueras, Jesús Lorenzana, María del Carmen Calvo Olivera, Claudia Álvarez-Aparicio, y Vicente Matellán.

Proceedings SARTECO'18-Jornadas Sarteco, Extended abstract. Teruel, Universidad de Zaragoza, septiembre 2018.

Predictive models of academic success: a case study with version control systems.

Autor/es: Ángel Manuel Guerrero-Higueras, Noemí Decastro-García, Vicente Matellán and Miguel Ángel Conde.

Proceedings TEEM'18 - Sixth Edition Technological Ecosystems for Enhancing Multiculturality (Teem 2018), Track 5. Learning Analytics: The good, the bad and the ugly, Extended abstract. Salamanca, Universidad de Salamanca, octubre 24-26.

Model for evaluating student performance through their interaction with version control systems.

Autor/es: Ángel Manuel Guerrero-Higueras, Vicente Matellán, Gonzalo Esteban Costales, Camino Fernández-Llamas, Francisco Jesús Rodríguez-Sedano, and Miguel Ángel Conde.

Learning Analytics Summer Institute Spain Workshop (LASI 2018), Extended abstract. León, Universidad de León with the collaboration of SNOLA (Spanish Network of Learning Analytics), June 18-19

Deploying containerized applications on HPC. A machine learning for cybersecurity example.

Autor/es: Ángel Manuel Guerrero-Higueras, Jesús Lorenzana, María del Carmen Calvo Olivera, Claudia Álvarez-Aparicio, y Vicente Matellán.

HPC Knowledge Meeting'18, Ponencia invitada. Barcelona, HPCNow, junio 2018.

Benchmark Dataset for evaluation of Range-Based people tracker Classifiers in Mobile Robots

Authors: Claudia Álvarez-Aparicio¹, Ángel Manuel Guerrero-Higueras¹, Maria Carmen Calvo Olivera¹, Francisco J. Rodríguez-Lera², Francisco Martín³ and Vicente Matellán¹

¹ Grupo de Robótica, Universidad de León, León, Spain

² University of Luxembourg, Luxembourg, Luxembourg

³ Robotics Lab, Universidad Rey Juan Carlos, Madrid, Spain

Journal: Frontiers in Neurorobotics

DOI: 10.3389/fnbot.2017.00072

PMID: 29379431

Abstract

Detecting and tracking people is a very useful skill for different systems, in particular, for improving navigation social for mobile robots, or to facilitate more socially acceptable robots. Many solutions in the literature try to solve this problem using a multimodal approach, typically vision and range sensors, as seen in Arras et al. (2012). Vision sensors are more expensive and are more likely to gather contradictory information. For this reason, systems based only on range sensors are desirable. Regarding the classifiers for processing sensor data, convolutional neural networks are becoming a very popular solution, according to Long et al. (2015).

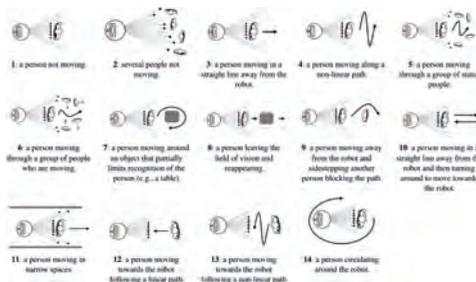
Laser imaging detection and ranging (LIDAR) sensors are reliable and currently affordable range sensors that provide information about the environment at good rates (~20–30 Hz.) for dynamic environments. They are easy to process on-board in real-time because each scan consists an array of only a few 100 integers.

Usually mobile robots mount laser scanners in a low position (~30–50 cm from the ground) to detect dynamic obstacles. They are also used to construct occupancy maps and navigate. The information provided allows estimating the distance at precise angles (resolution of 0.5°). This means that objects such as table or chair legs, trunks of plants, etc., may be easily confused with legs of persons. It is also difficult to keep track of a particular person (i.e., a pair of legs) in a crowded environment because many obstructions can result.

Fitting neural networks requires a good training dataset. Collecting and organizing a training set requires time as well as domain-specific knowledge. There is a large collection of robotic datasets available from various mobile robots, vehicles, and handheld sensors, such as Repository of robotics and computer vision datasets¹ for Mobile Robot Programming Toolkit (MRPT). However, most datasets may not be suitable for training neural networks. This data report summarizes a benchmark dataset, which can be used to evaluate the performance of different approaches for detecting and tracking people by using LIDAR sensors. Information contained in the dataset is specially suitable for use as training data for neural network-based classifiers.

Data actually contained in the dataset allow evaluating two people trackers, both neural network-based: leg detector (LD), a widely used solution by the Robot Operating System (ROS) community, see Quigley et al. (2009); and a people-tracker tool developed by the Robotics Group at the University of Leon, known as PeTra.

The rest of this paper is organized as follows: Section 2 describes the systems and the environment where data were gathered. It also specifies the procedure and tools employed. Section 3 explains how the dataset can be accessed and enumerates some applications and limitations for the data included in the dataset.



Recognition situations recorded.

Supercomputers to improve the performance in higher education: A review of the literature

Authors: Álvaro Fernández^a, Camino Fernández^b, José Ángel Miguel-Dávila^b, Miguel Ángel Conde^b, Vicente Matellán^a

^a University of Leon-Supercomputing Center of Castile and Leon (SCAYLE), Spain

^b University of Leon, Spain

Journal of Computers & Education

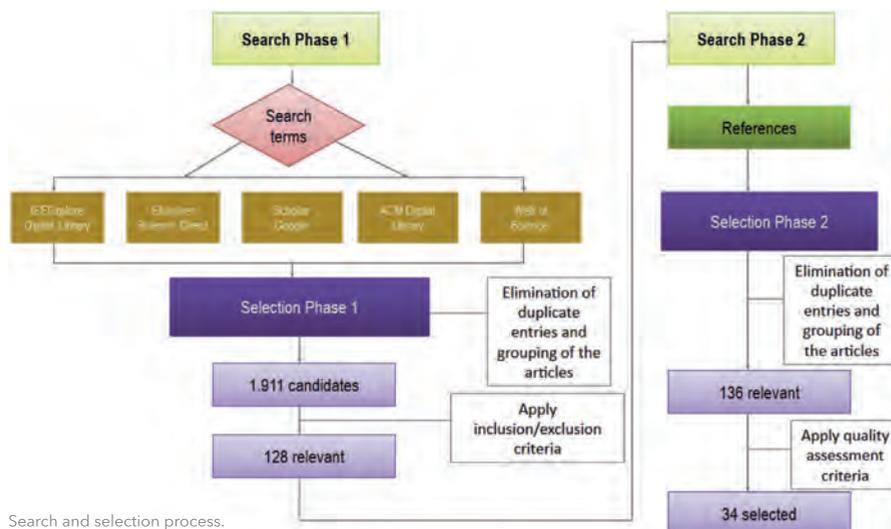
DOI: 10.1016/j.compedu.2018.10.004

ISSN: 03601315

Abstract

The use of Supercomputers is currently very widespread, constituting an essential component in many fields of science. The interest in the use of high performance computation is increasing in a wider and more diverse population of higher education students, mainly senior undergraduates and postgraduates, because the use of these infrastructures allows learners to improve their skills and the results of their training. For this reason, the demand of courses related to supercomputing increases continuously. In this paper we propose, through a wide review of primary studies, several questions that have been considered as a way of knowing the most widely-used contents in Supercomputing training. We have focused on the factors considered for improving training in Supercomputing, in order to improve the results of researchers in higher education organizations, to identify the limitations of Supercomputing training, and to provide solutions for these limitations. During the search procedure for answering research questions, 1911 studies were considered in the first selection.

Through the definition of inclusion and exclusion codes in the results of searching databases, 136 published articles were studied. Finally, using quality criteria, 34 studies were identified as relevant in answering the research questions. Several factors were described, such as the way in which courses related to Supercomputing are organized, the adaptations that are currently being applied in curricula related to the students of these techniques, the use of problem-solving training and the qualification of teachers, among the most relevant ones, as well as several limitations of this type of training and the identification of solutions for these limitations. Data was collected by searching keywords related to Supercomputing training and education in the most important databases used in Computational Science, finding empirical evidence to support the positive effect of High Performance Computers (HPC) on educators and researchers. The implications of this study are: first, it provides a summary of the most relevant factors in improving training, as well as the factors that improve the results through the use of a Supercomputer; and second, it provides the analysis of the limitations found for a better performance of learners and the solutions for these limitations.



Characterization of Spread in a Mesoscale Ensemble Prediction System: Multiphysics versus Initial Conditions

Authors: Sergio Fernández-González¹, Mariano Sastre², Francisco Valero², Andrés Merino³, Eduardo García-Ortega³, José Luis Sánchez³, Jesús Lorenzana⁴ and María Luisa Martín⁵

¹State Meteorological Agency (AEMET), Madrid, Spain

²Department of Earth Physics and Astrophysics, Faculty of Physics, Complutense University of Madrid, Madrid, Spain

³Atmospheric Physics Group, IMA, University of León, León, Spain

⁴SCAYLE Supercomputación Castilla y León, Spain

⁵Department of Applied Mathematics, Faculty of Computer Engineering, University of Valladolid, Segovia, Spain

Meteorologische Zeitschrift

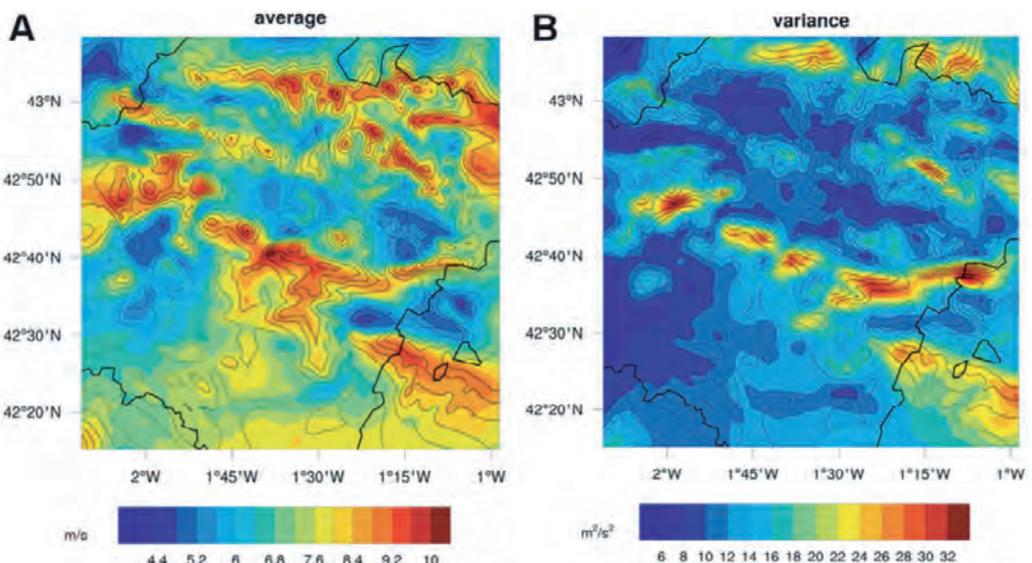
DOI: 10.1127/metz/2018/0918

Gebrüder Borntraeger Science Publishers, Stuttgart, www.borntraeger-cramer.com

© 2018 The authors

Abstract

In this research, uncertainty associated with initial and boundary conditions is evaluated for short-term wind speed prediction in complex terrain. The study area is the Alaiz mountain range, a windy region in the northern Iberian Peninsula. A multiphysics and multiple initial and boundary condition ensemble prediction system (EPS) was generated using the Weather Research and Forecasting model. Uncertainty of the EPS is analyzed using an index based on the spread between ensemble members, considering its behavior under different wind speed and direction events, and also during distinct atmospheric stability conditions. The results corroborate that physical parameterization uncertainty is greater for short-term forecasts (63.5%). However, it is also necessary to consider the uncertainty associated with initial conditions, not only for its quantitative importance (36.5%) but also for its behavior during thermal inversion conditions in the narrow valleys surrounded by mountains.



Temporal average (A) and variance (B) of wind speed at 90m a.g.l. in domain 4, simulated by WRF ensemble mean during study period.

Sensitivity analysis of WRF model: wind-resource 1 assessment for complex terrain

Authors: Sergio Fernández-González^a, María Luisa Martín^b, Eduardo García-Ortega^c, Andrés Merino^c, Jesús Lorenzana^d, José Luis Sánchez^c, Francisco Valero^a and Javier Sanz Rodrigo^e

^a Department of Earth Physics, Astronomy and Astrophysics II, Faculty of Physics, Complutense University of Madrid, Madrid, Spain.

^b Department of Applied Mathematics, Faculty of Computer Engineering, University of Valladolid, Segovia, Spain.

^c Atmospheric Physics Group, IMA, University of León, León, Spain.

^d Supercomputing Center of Castile and León, University of León, 24071 León, Spain.

^e National Renewable Energy Centre (CENER), Sarriguren, Spain.

Journal of Applied Meteorology and Climatology

AMERICAN METEOROLOGICAL SOCIETY

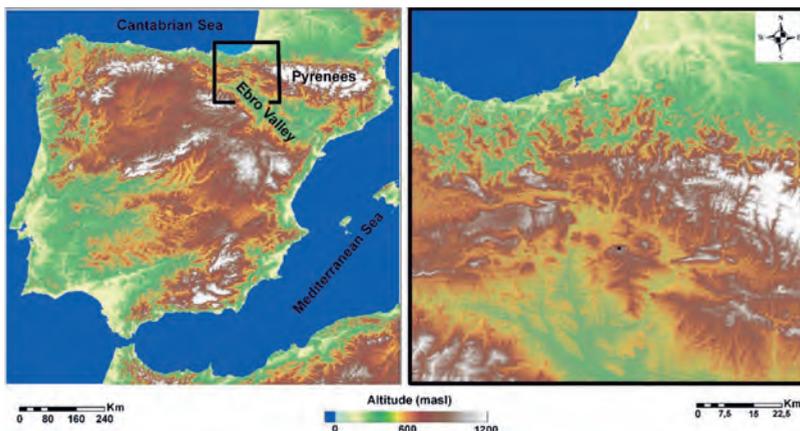
DOI:10.1175/JAMC-D-17-0121.1.

© 2017 American Meteorological Society

Abstract

Wind energy requires accurate forecasts for adequate integration into the electric grid system. In addition, global atmospheric models are not able to simulate local winds in complex terrain, where wind farms are sometimes placed. For this reason, the use of mesoscale models is vital for estimating wind speed at wind turbine hub height. In this regard, the Weather Research and Forecasting (WRF) model allows a user to apply different initial and boundary conditions, as well as physical parameterizations. In this research, we performed a sensitivity analysis of several physical schemes and initial and boundary conditions for the Alaiz mountain range in the northern Iberian Peninsula, where several wind farms are located.

Model performance was evaluated under various atmospheric stabilities and wind speeds. For validation purposes, a mast with anemometers installed at 40, 78, 90 and 118 m above ground level (m AGL) was used. The results indicate that performance of the Global Forecast System (GFS) analysis and ERA-Interim reanalysis as initial and boundary conditions was similar, although each performed better under certain meteorological conditions. Regarding physical schemes, there is no single combination of parameterizations that performs best during all weather conditions. Nevertheless, some combinations have been identified as inefficient so their use is discouraged. As a result, the validation of an ensemble prediction system (EPS) composed of the best 12 deterministic simulations shows the most accurate results, obtaining relative errors in wind speed forecasts that are < 15%.



Orography of Iberian Peninsula 856 (left panel). A zoomed view around Alaiz mountain range (black dot) is shown in right panel

FORMACIÓN

Curso de Introducción a la Dinámica de Fluidos Computacional con OpenFOAM

Basic Computational Skills for Metagenomics Analysis

Use of genomic sequence data in animal breeding

Curso Práctico de Iniciación al uso de la Supercomputación aplicado al Análisis de datos RNA-Seq - 4ª edición

Curso Práctico de iniciación del uso de la Supercomputación aplicado a la Metagenómica y Genómica comparada - 6ª edición

Curso Práctico avanzado del uso de la Supercomputación aplicado a la Metagenómica y Genómica comparada

Otras colaboraciones en materia de formación de SCAYLE

Curso de Introducción a la Dinámica de Fluidos Computacional con OpenFOAM

Objetivos

Proporcionar formación básica en el uso de técnicas computacionales para la resolución numérica de movimientos fluidos, tanto teórica como práctica. Introducir el uso de OpenFOAM, un paquete de código abierto que implementa funcionalidades interesantes tanto desde el punto de vista académico como profesional.

Destinatarios

Alumnos universitarios (titulaciones técnicas con contenidos específicos sobre fluidos), investigadores y profesionales noveles del sector aeronáutico e industrial y, en general, cualquier persona afín a la temática que se quiera iniciar en las técnicas CFD.

Profesorado

Jesús Gonzalo de Grado.

Departamento de Ingeniería Mecánica, Informática y Aeroespacial. Área de Ingeniería Aeroespacial. Universidad de León.

Joaquín Fernández Francos.

Departamento de Ingeniería Mecánica, Energética y de los materiales. Universidad de Extremadura.

Diego Domínguez Fernández.

Departamento de Ingeniería Mecánica, Informática y Aeroespacial. Área de Ingeniería Aeroespacial. Universidad de León.

Deibi López Rodríguez.

Departamento de Ingeniería Mecánica, Informática y Aeroespacial. Área de Ingeniería Aeroespacial. Universidad de León.

Jesús Lorenzana Campillo.

Fundación Centro de Supercomputación de Castilla y León.

Duración 30 horas

Fecha y lugar

22 al 26 de enero de 2018. Universidad de León.

Contenidos

Supercomputación y CFD.

- *Presentación del Centro de Supercomputación.*
- *Introducción a la supercomputación aplicada al CFD.*
- *Uso de supercomputadores para resolver problemas de CFD.*

Introducción al CFD y presentación de Open FOAM.

- *Definición, conceptos clave y capacidad de las técnicas CFD.*
- *Ecuaciones de la mecánica de fluidos.*
- *Modelos de turbulencia.*
- *Métodos numéricos: Elementos Finitos, Volúmenes Finitos.*
- *Algoritmos de resolución.*

Preparación de casos para simular.

- *Instalación, características, prestaciones y ventajas de OpenFOAM.*
- *Mallado.*
- *Condiciones iniciales y de contorno.*
- *Utilities de OpenFOAM.*

Procesado de datos.

- *Solvers en OpenFOAM.*
- *Ejecución y monitorización de simulaciones.*

Post-Procesado.

- *Visualización e interpretación de resultados.*
- *Localización de errores.*

Técnicas avanzadas en Open FOAM.

- *Ejecución de tareas en paralelo.*
- *Modelos de turbulencia.*
- *Modificación de Códigos.*

Trabajo individual tutorizado y prueba final.

- *Realización de prácticas y prueba final de evaluación.*

Basic Computational Skills for Metagenomics Analysis

Objetivos

The course will teach the basic computational skills required to perform analyses in metagenomics. The module will be eminently practical and hands-on, and will teach participants skills ranging from the installation and basic use of Linux systems to the building of analysis pipelines. The course will teach the basics in Next Generation Sequencing data quality assessment, metagenomic assembly and statistics, annotation process in standard and supercomputing environments.

Profesorado

Giuseppe D'Auria.

Fundación para el Fomento de la Investigación Sanitaria y Biomédica de la Comunitat Valenciana. (FISABIO), Valencia.

Duración 36 horas

Fecha y lugar

5 al 9 de febrero de 2018. Università degli Studi di Pavia, Italia.

Contenidos

Introduction to Linux and NGS.

- *Next generation sequencing technologies: Introduction to metagenomics.*
- *Linux system, installing a new distro, virtual environments.*
- *From office automation to computation for NGS.*
- *Unix commands, moving around the system, building pipelines.*

Basic steps in Metagenomics.

- *NGS Bioinformatics formats, managing data, practice.*
- *Metagenome assembly and quality control.*
- *Reads mapping and ORFs finding.*

Metagenomes validation.

- *Metagenomes validation.*
- *Scaffolds analysis.*
- *Metagenomics analysis (wrap up section).*

Metagenomics using a computer cluster.

- *Bash scripting and qsub environment for computer cluster queuing system in metagenomics.*
- *running metagenomics analysis using a computer cluster computer.*

Metataxonomy analysis (16S rDNA amplicons).

- *Quimera check and taxonomic annotation.*
- *Descriptive and statistics analysis of 16S rDNA data.*

Use of genomic sequence data in animal breeding

Objetivos

Los recientes avances en las tecnologías de secuenciación del ADN de alto rendimiento están abriendo nuevas oportunidades para el uso de la información de la secuencia genómica en la mejora genética animal. Estas tecnologías producen grandes cantidades de datos que requieren herramientas bioinformáticas apropiadas para su análisis e interpretación. Entender la aplicación de estas herramientas es esencial para que los especialistas en mejora genética puedan evaluar el potencial de estos nuevos enfoques y su posible aplicación en cada caso particular. Este curso proporcionará formación práctica básica en bioinformática utilizando ejemplos de datos reales de secuencias de animales domésticos. El objetivo es facilitar a los participantes el uso de los datos de secuenciación de sus propios proyectos o de las bases de datos públicas para su utilización en sus propios programas de investigación.

Al final del curso los participantes estarán en posición de hacer:

1. Manejar grandes conjuntos de datos resultantes de tecnologías genómicas, en un entorno Linux.
2. Identificar y explorar los recursos públicos de datos y bioinformática relevantes para los animales domésticos.
3. Explorar datos de secuenciación de próxima generación, incluyendo anotación e identificación de variantes.
4. Evaluar diferentes alternativas para el análisis de datos de variación y cartografía.
5. Utilizar herramientas que les permita utilizar la variabilidad genómica en análisis posteriores.

Duración 16 horas

Fecha y lugar

11 al 13 de junio de 2018. Università degli Studi di Pavia, Italia.

Profesorado

Juan José Arranz Santos.

Dpto. Producción Animal, Universidad de León.

Jesús Lorenzana Campillo.

SCAYLE Supercomputación de Castilla y León.

Beatriz Gutiérrez Gil.

Dpto. Producción Animal, Universidad de León.

Cristina Esteban Blanco.

Dpto. Producción Animal, Universidad de León.

Héctor Marina García.

Dpto. Producción Animal, Universidad de León.

Aroa Suárez Vega.

The Department of Animal Biosciences, University of Guelph, Canadá).

Contenidos

Trabajar en entorno Linux/Unix. Tutorial interactivo

- *Fundamentos de Linux.*
- *Gestión de archivos grandes.*
- *Ejecutar grandes trabajos.*
- *Ejecución de Linux desde su PC.*

Bioinformática básica. Introducción: Formato de ficheros NGS. Uso de ficheros de bases de datos públicas (formato SRA).

Workflow de análisis WGSeq para identificación de variantes en secuencias genómicas.

- *Trabajo con ficheros SRA.*
- *Control de calidad (QC) de los datos brutos: Trimmomatic.*
- *Alineación de secuencia frente al genoma de referencia.*
- *Manipulación de ficheros con Samtools.*
- *Identificación de SNVs con GATK.*
- *Anotación funcional con VEP.*
- *Manejo de archivos vcf y utilización de la herramienta PLINK.*

Curso Práctico de Iniciación al uso de la Supercomputación aplicado al Análisis de datos RNA-Seq-4^a edición

Objetivos

En este curso se proporcionará una formación básica para el manejo e interpretación de datos de expresión génica global procedentes de *Next Generation Sequencing* (RNA-Seq). Para ello, además de explicar las bases teóricas de la generación de los datos y del proceso de análisis, se pretende trabajar con datos reales de expresión génica en los que se realizará: el control de calidad, el alineamiento frente al genoma de referencia, ensamblado, cuantificación y normalización de la expresión génica, análisis de expresión diferencial y análisis de enriquecimiento funcional.

Destinatarios

El curso está dirigido a investigadores interesados en estudios genómicos, a profesionales del sector de las Ciencias Computacionales, Biología y/o Biotecnología relacionados con el diagnóstico genético y a Alumnos Universitarios (titulaciones técnicas del ámbito experimental y/o económico) de posgrado y, en general, cualquier persona afín a la temática tanto en la dimensión de la investigación, como de la innovación y el desarrollo.

Profesorado

Aroa Suárez Vega.

Dpto de Producción Animal, Universidad de León.

Beatriz Gutiérrez Gil.

Dpto de Producción Animal, Universidad de León.

Beatriz Rosón Burgo.

Instituto Karolinska, Estocolmo, Suecia.

Cristina Esteban Blanco

Supercomputación de Castilla y León, León.

Juan José Arranz Santos.

Dpto de Producción Animal, Universidad de León.

Fecha y lugar

18 al 22 de junio de 2018. Edificio CRAI-TIC, Campus de Vegazana, Universidad de León

Duración 36 horas

Contenidos

Introducción acceso a Caléndula.

Introducción al entorno Linux.

NGS y RNA-Seq Supercomputación.

Control de Calidad y Trimming (FAstQC, otras herramientas Trimmomatic, etc.).

Alineamiento de lecturas (TopHat) y visualización (IGV).

Manipulación de secuencias (SamTools).

Transcript assembly (Cufflinks).

Cuantificación de lecturas (Cufflinks y HTSeq).

Introducción a R y Bioconductor. Toma de contacto.

Qué es el análisis de expresión diferencial.

Análisis de expresión diferencial de RNAseq.

Introducción a las anotaciones funcionales.

Bases de datos y ontologías para anotación funcional (KEGG, GO, INTERPRO).

Análisis de enriquecimiento funcional (DAVID and GeneTermLinker).

Redes funcionales (FGNet package).

Visita al Superordenador Caléndula (voluntario).

SUPERCOMPUTACIÓN
CASTILLA Y LEÓN

Curso Práctico de iniciación del uso de la Supercomputación aplicado a la Metagenómica y Genómica comparada-6ª edición

Objetivos

Se proporciona la formación necesaria para el análisis de datos procedentes de técnicas de *Next Generation Sequencing*, centrada particularmente en su aplicación al estudio metagenómico de muestras de diversos ambientes y emplear la supercomputación en la recopilación y ensamblado de los fragmentos de ADN secuenciados, así como su posterior anotación y análisis.

Destinatarios

El curso está dirigido a investigadores interesados en estudios genómicos, a profesionales del sector de las Ciencias Computacionales, Biología y/o Biotecnología relacionados con el diagnóstico genético y a Alumnos Universitarios (titulaciones técnicas del ámbito experimental y/o económico) de posgrado y, en general, cualquier persona afín a la temática tanto en la dimensión de la investigación, como de la innovación y el desarrollo.

Profesorado

Cristina Esteban Blanco.

Dpto de Producción Animal, Universidad de León.

Giuseppe D'Auria.

Fundación para el Fomento de la Investigación Sanitaria y Biomédica de la Comunitat Valenciana. (FISABIO), Valencia.

Javier Tamames de la Huerta.

Centro Nacional de Biotecnología (CNB), Madrid.

Duración 36 horas

Fecha y lugar

15 al 19 de octubre de 2018. Edificio CRAI-TIC, Campus de Vegazana, Universidad de León

Contenidos

Introducción al entorno Linux.

Control de calidad de las secuencias.

Asignación taxonómica.

Estimación de abundancia.

Análisis de datos de taxonomía.

Ensamblaje y coensamblaje de metagenomas.

Predicción de genes y búsqueda de homología.

Asignación funcional y taxonómica.

Mapping de lecturas sobre contigs para estimar abundancias.

Binning: Separación de MAGs (genomas de especies individuales).

Estimación de expresión de genes presentes en el metagenoma.

Comparación de niveles de expresión en diferentes condiciones.

Visita al Superordenador Caléndula (voluntario).

Curso Práctico avanzado del uso de la Supercomputación aplicado a la Metagenómica y Genómica comparada

Objetivos

Este curso está destinado a profundizar en los métodos de análisis metagenómico.

La primera parte está enfocada en métodos de ensamblaje híbridos, combinando diferentes tipos de secuencias, especialmente secuencias largas y cortas. En la segunda parte se trabajará en profundidad con los resultados de binning, para poder recuperar con gran precisión genomas muy completos a partir del metagenoma. Se enseña como depurar los bins, como completarlos, y como combinar resultados de diferentes métodos.

La última parte se centra en el análisis estadístico de los resultados, usando R para obtener asociaciones entre abundancias de taxones/genes/rutas metabólicas y tipos de muestra.

Destinatarios

El curso está dirigido a investigadores interesados en estudios genómicos, a profesionales del sector de las Ciencias Computacionales, Biología y/o Biotecnología relacionados con el diagnóstico genético y a Alumnos Universitarios (titulaciones técnicas del ámbito experimental) de posgrado y, en general, cualquier persona afín a la temática.

Fecha y lugar

5 al 9 de noviembre de 2018. Edificio CRAI-TIC, Campus de Vegazana, Universidad de León.

Profesorado

Fernando Puente Sánchez

Centro Nacional de Biotecnología (CNB), Madrid.

Giuseppe D'Auria.

Fundación para el Fomento de la Investigación Sanitaria y Biomédica de la Comunitat Valenciana. (FISABIO), Valencia.

Javier Tamames de la Huerta.

Centro Nacional de Biotecnología (CNB), Madrid.

Duración 36 horas

Contenidos

Introducción a la secuenciación de tercera generación.

Control de calidad de secuencias de tercera generación tipo Oxford Nanopore (Minlon).

Ensamblajes mixtos: PacBio/Illumina/Minlon.

Ensamblado de secuencias largas (tercera generación).

Corrección de errores con secuencias cortas (segunda generación).

Acabado del ensamblado.

Anotación funcional del ensamblado.

Depuración de bins: Detección de contigs no pertenecientes al bin; Reclutado de nuevos contigs.

Combinación de métodos de binning.

Anotación metabólica de bins.

Manejo de R.

Detección de funciones metabólicas con abundancia diferencial en conjuntos de metagenomas/metatranscriptomas con R/DeSeq2.

Análisis multivariante y redes de co-abundancia.

Caso práctico 1: Búsqueda de auxotrofías en biosíntesis de aminoácidos y vitaminas en bins metagenómicos.

Caso práctico 2: Caracterización de genes de transportadores de membrana.

Caso práctico 3: Modelado conceptual de comunidades microbianas a partir de datos metagenómicos.

Visita al Superordenador Caléndula (voluntario).

Otras colaboraciones en materia de formación de SCAYLE

Asignatura de Computación de Altas Prestaciones, Máster Universitario en Ingeniería Informática.

Departamento de Ingeniería Mecánica, Informática y Aeroespacial, Área de Ingeniería Aeroespacial.
Universidad de León.

11 de noviembre de 2017 al 17 de enero de 2018.

Asignatura de Computación grid y supercomputación. 4º de Graduado en Ingeniería Informática

Departamento de Ingeniería Mecánica, Informática y Aeroespacial, Área de Ingeniería Aeroespacial.
Universidad de León.

14 de marzo al 6 de julio 2018.

Asignatura de Computación de Altas Prestaciones, Máster Universitario en Ingeniería Informática.

Departamento de Ingeniería Mecánica, Informática y Aeroespacial, Área de Ingeniería Aeroespacial.
Universidad de León.

2 de octubre de 2017 al 10 de febrero 2018.

DATOS DE SCAYLE

10° aniversario SCAYLE

Datos técnicos

Datos económicos

Protocolos y convenios de colaboración

Visitas a SCAYLE

Actividades de promoción, difusión y comunicación

10º aniversario SCAYLE



El 11 de enero de 2018 la Fundación Centro de Supercomputación de Castilla y León celebró su décimo aniversario.

Durante estos primeros diez años se han ido realizando continuas actualizaciones técnicas que han permitido aportar recursos vitales a la investigación, siempre respaldados por un equipo humano altamente comprometido y profesional para apoyar los avances de la investigación.

Por todo ello, y para celebrarlo, SCAYLE organizó un evento en el que participaron el presidente del patronato y Consejero de Fomento y Medio Ambiente de la Junta de Castilla y León D. Juan Carlos Suarez-Quiñones, el Vicepresidente y Rector de la Universidad de León D. Juan Francisco Garcia Marín, el Alcalde de León y anterior presidente de SCAYLE D. Antonio Silván Rodríguez, la Directora General de Telecomunicaciones de la Junta de Castilla y León Dña. Mª Victoria Seco Fernández, así como autoridades, investigadores y público general.

Todos ellos se congregaron para celebrar la evolución de SCAYLE a lo largo de estos diez años, y durante las exposiciones que realizaron el personal de SCAYLE, los investigadores y las empresas que participaron, se puso de manifiesto lo que a lo largo de estos años ha apotado disponer de un centro regional de supercomputación. Así lo indicaron la empresa Tecnosylva y el grupo de investigación de Producción animal de la Facultad de Veterinaria de la Universidad de León, ambos ponentes el día de la celebración y usuarios de la infraestructura desde los inicios de SCAYLE. También se quiso contar para la ocasión con representantes de otros centros de Supercomputación, como en el caso del CESGA, el Centro de Supercomputación de Galicia, que puso de manifiesto la importancia que tiene el uso de este tipo de infraestructuras y la facilidad de acceso que se tiene a ellas, tanto a través de la RES, mediante convocatorias de financiación, o a través de grupos de investigación de empresas u otras instituciones independientes.



De izquierda a derecha D. Vicente Matellán Olivera, D. Juan Francisco Garcia Marín, D. Juan Carlos Suarez-Quiñones, D. Antonio Silván Rodríguez y Dña. Mª Victoria Seco Fernández. Fuente: Diario de León

También durante el evento se presentó la nueva imagen de la Fundación Centro de Supercomputación de Castilla y León, que cambió su acrónimo de FCSCCL a SCAYLE

El nuevo logo hace referencia a los laterales del superordenador en perspectiva, y los colores representan los cambios de temperatura que se sufre en el interior de Caléndula provocados por la velocidad del proceso de los datos, utilizando para ello colores identificados como colores digitales.

2018: de FCSCCL a SCAYLE



Con motivo de este aniversario, desde SCAYLE se han realizado una serie de actividades a lo largo del año como ha sido la participación en varias jornadas y congresos a través de las cuales se ha tratado de transmitir como se puede trabajar con SCAYLE y los beneficios que ello puede reportar en los avances de la investigación tanto en Castilla y León, ámbito principal de actuación de SCAYLE, como a nivel nacional e internacional con la participación del centro en diferentes convocatorias de financiación. Así destaca la participación en congresos como el "Congreso de la Sociedad Española de Medicina Interna (SEMI) en Madrid, UNVEX 2018- Security & Defense en León, o en 12ENISE también en León.

Las jornadas que se han realizado en las universidades públicas de Castilla y León, como la jornada "La Supercomputación en la I+D+i de la Universidad" que se realizó en la Universidad de Burgos y en la Universidad de Salamanca, o jornadas en la administración como en el caso de la jornada "Supercomputación para la competitividad" realizada en la consejería de Fomento y Medio Ambiente de Castilla y León, siempre han tenido como finalidad el poder acercar un poco al ciudadano una infraestructura como Caléndula y brindar la oportunidad de utilizarla a todos aquellos que pudieran necesitar una infraestructura como la que se tiene en León.

También se aprovechó la participación en eventos en las que otras entidades nos invitan para celebrar el décimo aniversario. Así se hizo en "LEÓN4UP! Innovación & Talento" en León, o "StartUp Olé" en Salamanca, en los que mediante coloquios se explicó que aporta el uso de centros de datos como SCAYLE.

A estas actividades le sumaron todas aquellas que de forma habitual se realizan desde el centro, como son la organización de cursos a los que asisten investigadores que necesitan procesar sus datos en un centro especializado, porque sus equipos ya no son capaces trabajar con el volumen de datos que han ido adquiriendo en el avance de sus investigaciones, este perfil de alumnado necesita poder aprender a trabajar en un superordenador para seguir avanzando en el desarrollo de sus investigaciones.

Utilización del Sistema

Cálculo científico

A lo largo del año 2018 se han ejecutado en el sistema de colas gestionado por SGE 9.997.340 horas de CPU y en el nuevo, gestionado por Slurm 826.502 horas.

Durante el primer semestre se instaló un servidor dotado con procesadores gráficos (GPUs), que desde el mes de julio se puso a disposición de los investigadores.

La incorporación de esta tecnología también ha supuesto el despliegue de aplicaciones contenedorizadas por primera vez en Caléndula. Esta forma de desplegar aplicaciones va a permitir ampliar aún más el espectro de investigadores que pueden hacer uso de la infraestructura.

Otro de los hitos ha sido la llegada de los nodos del superordenador "MareNostrum 3" cedidos por el Centro Nacional de Supercomputación. Se han recibido 3 racks de IBM, cada uno de ellos con 84 nodos de 2 procesadores de 8 cores y 32 Gigabytes de memoria.

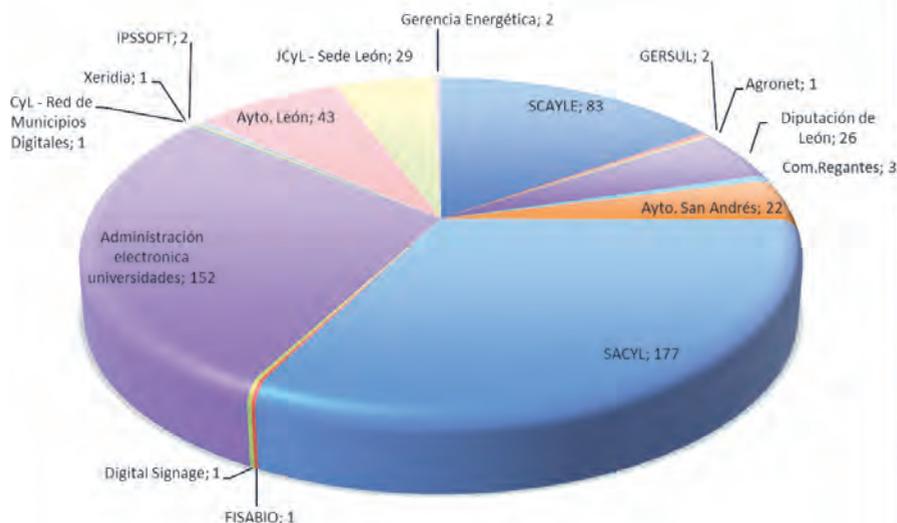
Durante el segundo semestre se ha realizado la instalación física de los servidores, lo que ha implicado la retirada del servicio de los nodos de la primera versión de Caléndula que ya habían superado su vida útil. En total se retiraron 244 servidores con una potencia de cálculo de 23,4 TeraFLOPS y se han instalado 168 servidores con una potencia pico de cálculo de 56 TeraFLOPS.

También se ha migrado el sistema de gestión de colas de Caléndula. El gestor SGE se ha cambiado por Slurm. Este cambio permite implementar las políticas de prioridades y límites más granulares en los tiempos de ejecución.

Servicios Cloud

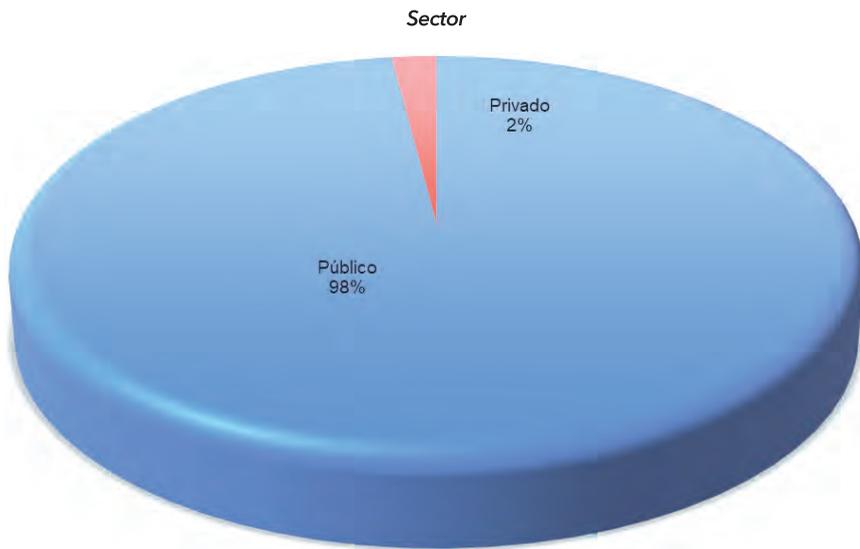
Se encuentran desplegados 546 servidores virtuales en la nube de SCAYLE:

Distribución de Máquinas Virtuales



Fuente: Área Técnica SCAYLE, 31/12/2018

La distribución por sectores es:



Fuente: Área Técnica SCAYLE, 31/12/2018.

Actualización tecnológica

La plataforma de virtualización sigue funcionando de manera estable ofreciendo buen rendimiento y disponibilidad, que en el periodo ha seguido siendo mayor del 99,9%.

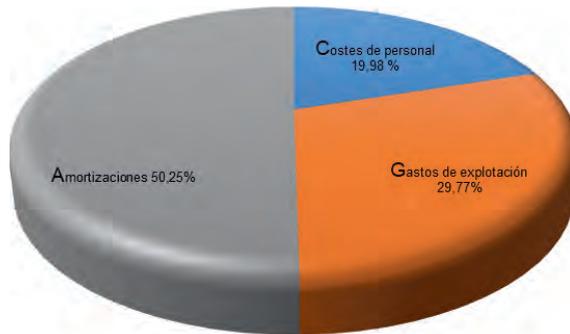
Durante el primer semestre de 2018 se han realizado las siguientes actividades:

- Instalación de un servicio de ficheros compartidos de uso interno basado en la nube.
- Configuración y puesta en servicio de los nuevos firewalls así como migración de todos los clientes.
- Ampliación de los sistemas de backup de SCAYLE.
- Instalación del sistema de cálculo basado en GPUs.
- Migración de servidores de correo y web del dominio fcsc.es al dominio scayle.es
- Instalación de la actualización del equipamiento de red de SCAYLE.

datos económicos

Gastos de SCAYLE, ejercicio 2018: **1.562.417,80 €**

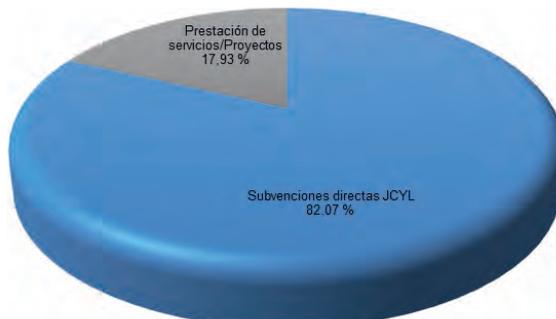
La principal partida de gastos de la SCAYLE en 2018 continúa siendo la de amortizaciones, debido a la contabilización de la depreciación de los equipos de la Red Regional de Ciencia y Tecnología de Castilla y León, así como del resto de inversiones financiadas con cargo a FEDER. Los gastos de explotación incluyen los costes de subcontratas y de los principales suministros como electricidad y comunicaciones, son la segunda partida en importancia. Los costes de personal suponen el resto de gastos, siendo la de menor importe de las tres.



Fuente: Área Administrativo-Financiera SCAYLE, 2018

Ingresos de SCAYLE, ejercicio 2018: **1.562.417,80 €**

Los principales ingresos proceden, al igual que en años anteriores, de las aportaciones de la Junta de Castilla y León, Consejería de Fomento y Medio Ambiente y Consejería de Educación, para financiar tanto los gastos de explotación como las inversiones, cofinanciadas estas últimas a través de fondos FEDER, representando más de un 80 % de los ingresos. La prestación de servicios a empresas y los ingresos procedentes de la participación en proyectos suponen el resto de ingresos de SCAYLE.



Fuente: Área Administrativo-Financiera SCAYLE, 2018

protocolos y convenios de colaboración

Protocolo entre el Ayuntamiento de San Andrés del Rabanedo y la Fundación centro de supercomputación de Castilla y León para establecer un marco adecuado que agilice los proyectos de colaboración para el desarrollo tecnológico, científico y socioeconómico y para la adaptación de tecnologías emergentes de la información y de las comunicaciones en los campos que son propios de los fines de ambas instituciones.

Comodato equipos informaticos entre BSC-CNS y SCAYLE.

Convenio entre la Universidad de León y la Fundación Centro de Supercomputación de Castilla y León para la Colaboración en Materia de Docencia Reglada, Formación, Investigación y Desarrollo Científico-Tecnológico.

- 19-01-18 - IES Juan del Enzina, Dpto. Tecnología.
- 26-01-2018 - HP-CDS.
- 8-02-2018 - Centro de F.P. María Auxiliadora de León.
- 12-02-2018 - Master Internacional Económicas, European Master in Business Studies (EMBS), Universidad de León.
- 16-02-2018 - Curso OpenFOAM-Visita de alumnos.
- 02-03-2018 - Centro de Formación para el Empleo de León de las Carrizas, San Andrés del Rabanedo (León), curso "Sistemas Microinformáticos (IFCT0209)".
- 22-03-2018 - Escuela de adultos.
- 13-04-2018 - Centro de F.P. María Auxiliadora de León.
- 19-04-2018 - Colegio Beata Imelda - Santo Tomás de La Felguera - FESD, Principado de Asturias.
- 24-04-2018 - Universidad de León.
- 03-05-2018 - Colegio San José. Agustinas Misioneras de León.
- 10-05-2018 - Semana Internacional ULE.
- 11-05-2018 - Círculo Empresarial Leonés.
- 15-05-2018 - Alumnos Máster Universitario en Ingeniería Informática, Universidad de León.
- 25-05-2018 - Colegio Virgen Blanca de León.
- 01-06-2018 - Colegio Divina Pastora de León.
- 04-06-2018 - Ateneo Leonés- Centro de Formación.
- 14-06-2018 - IESO Ástura de Mansilla de las Mulas, León.
- 28-06-2018 - Campus Tecnológico de la Universidad de León.
- 05-07-2018 - Campus Tecnológico de la Universidad de León.
- 12-07-2018 - Campus Tecnológico de la Universidad de León.
- 19-07-2018 - Cybersecurity Summer Camp.
- 20-07-2018 - Cybersecurity Summer Camp.
- 23-07-2018 - Cybersecurity Summer Camp.
- 24-07-2018 - Cybersecurity Summer Camp.
- 24-09-2018 - Subdelegado Gobierno de León.
- 28-09-2018 - Grado de Informática de la Universidad de Valladolid.
- 19-10-2018 - Alumnos Curso Práctico de Iniciación al uso de la Supercomputación aplicado a la Metagenómica y Genómica comparada.
- 08-11-2018 - Instituto Nacional de Ciberseguridad.
- 09-11-2018 - IES Obispo Argüelles de Villablino, León.
- 09-11-2018 - Alumnos Curso Práctico Avanzado del uso de la Supercomputación aplicado a la Metagenómica y Genómica comparada.
- 12-11-2018 - Visita Semana de la Ciencia.
- 13-11-2018 - Visita Semana de la Ciencia.
- 15-11-2018 - Visita Semana de la Ciencia.
- 16-11-2018 - Visita Semana de la Ciencia.
- 29-11-2018 - I.E.S. La Providencia de Medina de Pomar, Burgos.
- 21-12-2018 - Colegio Nuestra Señora del Carmen, León.

actividades de promoción, difusión y comunicación

Cyber-Security Scientific Research Conference 2018.
León (22 de enero de 2018).

La Supercomputación en la I+D+i de la Universidad.
Burgos (13 de febrero de 2018).

Reunión Consejo Municipal de Nuevas Tecnologías.
León (14 de febrero de 2018).

Consejo Red Española Supercomputación.
Madrid (16 de marzo).

Cybersecurity Innovation HUB de Castilla y León.
León (20 de marzo de 2018).

StartUp Olé.
Salamanca (17 al 19 de abril de 2018).

UNVEX, Security & Defense - 2018.
León (29 al 31 de mayo 2018).

Jornada Supercomputación para la competitividad.
Valladolid (8 de junio de 2018).

HPDA Forum 2018.
Toulouse (18 al 20 de junio de 2018).

Cybersecurity Summer Camp.
León (17 de julio de 2018).

Atos Quantum Learning Machine.
Madrid (18 de septiembre 2018).

Reunión anual de usuarios de la Red Española de Supercomputación (RES).
Valencia (19 al 21 de septiembre de 2018).

León4UP.
León (3 de octubre de 2018).

12 ENISE-Encuentro Internacional de Seguridad de la Información.
León (23 y 24 de octubre de 2018).

Big data y Supercomputación en Medicina.
Burgos (21 de noviembre de 2018).

Supercomputación aplicada a la I+D+i de la Universidad de Salamanca.
Salamanca (31 de noviembre de 2018).

