

**Guía docente de la asignatura**

Asignatura	INFRAESTRUCTURAS PARA EL DESARROLLO DE APLICACIONES DE COMPUTACIÓN DE ALTAS PRESTACIONES		
Materia	SISTEMAS Y SERVICIOS EMPOTRADOS, UBÍCUOS Y DE ALTAS PRESTACIONES		
Módulo			
Titulación	MÁSTER EN INGENIERÍA INFORMÁTICA		
Plan	510	Código	53184
Periodo de impartición	2º CUATRIMESTRE	Tipo/Carácter	OPTATIVA
Nivel/Ciclo	MÁSTER	Curso	1º
Créditos ECTS	3 ECTS		
Lengua en que se imparte	CASTELLANO		
Profesor responsable	IVÁN SANTOS TEJIDO		
Datos de contacto	DESPACHO: 1D046, E.T.S.I. TELECOMUNICACIÓN TELÉFONO: 983 423000 ext. 5512 E-MAIL: ivasan@tel.uva.es		
Horario de tutorías	Véase www.uva.es → Centros → Campus de Valladolid → Escuela Técnica Superior de Ingeniería Informática → Tutorías		
Departamento	ELECTRICIDAD Y ELECTRÓNICA		

1. Situación / Sentido de la Asignatura**1.1 Contextualización**

La realización de cálculos computacionales intensivos es común en muchos ámbitos: en el procesado masivo de datos, en los modelos financieros, en el tratamiento y procesado de imágenes y videos, en la simulación de nuevos materiales y la evaluación de sus propiedades, en la ejecución de modelos climáticos para la predicción meteorológica, en la ejecución de modelos de dinámica de fluidos para cálculos aerodinámicos de nuevos diseños de coches o aviones...

Para poder obtener resultados en tiempos razonables, las aplicaciones utilizadas en estos campos suelen paralelizarse y ejecutarse en un gran número de procesadores a la vez (de decenas a miles de procesadores). Esto hace que los ordenadores personales comunes no sean los equipos más adecuados para este tipo de aplicaciones, y sea necesario recurrir a los denominados "high performance computing clusters" (HPC – clusters de alto rendimiento computacional) que son infraestructuras mucho más complejas con una mayor capacidad computacional.

En la asignatura de "Infraestructuras para el desarrollo de aplicaciones de computación de altas prestaciones" se abordarán los conceptos necesarios para diseñar y acondicionar las infraestructuras necesarias para poder ejecutar cálculos paralelos intensivos, tanto a nivel del software (requerimientos del sistema operativo, entornos de ejecución paralelos, librerías necesarias...) como del hardware (procesadores, memorias, coprocesadores



matemáticos, acondicionamiento eléctrico y térmico...) teniendo en cuenta las necesidades de cada tipo de aplicación que se quiera ejecutar.

Aunque muchos de los contenidos que se presentarán estarán centrados en infraestructuras dedicadas al cálculo intensivo, también son comunes a otros tipos de infraestructuras como son los centros de procesado de datos (CPDs), los servidores de virtualización o servidores web, los centros dedicados a la protección de datos... A lo largo de la asignatura se irán estableciendo paralelismos con estos otros entornos en los que también son necesarias infraestructuras computacionales complejas. De esta forma se pretende dar un enfoque más amplio sobre la utilidad de los contenidos desarrollados.

1.2 Relación con otras materias

La asignatura de "Infraestructuras para el desarrollo de aplicaciones de computación de altas prestaciones" pertenece a la materia de "Desarrollo de Infraestructuras para Sistemas y Servicios Empotrados, Ubicuos y de Altas Prestaciones" del Módulo de "Tecnologías Informáticas" del Máster de Ingeniería Informática. La asignatura se desarrollará durante la primera mitad del segundo semestre del Máster.

En esta asignatura se pretende dar una visión más aplicada de los contenidos desarrollados en la asignatura "Computación Paralela y Modelos Emergentes", que es la asignatura obligatoria de la materia "Desarrollo de Infraestructuras para Sistemas y Servicios Empotrados, Ubicuos y de Altas Prestaciones".

1.3 Prerrequisitos

Los conocimientos previos necesarios se desarrollan en las asignaturas obligatorias del Máster que los alumnos deben cursar.

2. Competencias

2.1 Generales

- CG1. Capacidad para proyectar, calcular y diseñar productos, procesos e instalaciones en todos los ámbitos de la ingeniería informática.
- CG4. Capacidad para el modelado matemático, cálculo y simulación en centros tecnológicos y de ingeniería de empresa, particularmente en tareas de investigación, desarrollo e innovación en todos los ámbitos relacionados con la Ingeniería en Informática.

2.2 Específicas

- CET1. Capacidad para modelar, diseñar, definir la arquitectura, implantar, gestionar, operar, administrar y mantener aplicaciones, redes, sistemas, servicios y contenidos informáticos.
- CET6. Capacidad para diseñar y evaluar sistemas operativos y servidores, y aplicaciones y sistemas basados en computación distribuida.
- CEG1. Capacidad para la integración de tecnologías, aplicaciones, servicios y sistemas propios de la Ingeniería Informática, con carácter generalista, y en contextos más amplios y multidisciplinares.

3. Objetivos

Al finalizar la asignatura el alumno deberá ser capaz de:

- Diseñar una infraestructura de computación de altas prestaciones teniendo en cuenta las necesidades de las aplicaciones que se van a ejecutar en ellas.
- Instalar y gestionar el software necesario para la infraestructura de computación de altas prestaciones.



4. Tabla de dedicación del estudiante a la asignatura

ACTIVIDADES PRESENCIALES	HORAS	ACTIVIDADES NO PRESENCIALES	HORAS
Clases teóricas	10	Investigación relacionada con las actividades realizadas durante las clases teóricas	5
Clases prácticas	0	Investigación relacionada con las prácticas de laboratorio y realización de las memorias de prácticas.	40
Laboratorios	20		
Prácticas externas, clínicas o de campo	0		
Seminarios	0		
Otras actividades	0		
Total presencial	30	Total no presencial	45





5. Bloques temáticos

Bloque 1: Contenidos Teóricos

Carga de trabajo en créditos ECTS:

a. Contextualización y justificación

En este bloque se presentan los contenidos de la asignatura. Se pretende dar una visión general al alumno sobre los diferentes tipos de sistemas computacionales de altas prestaciones que existen, de sus componentes, y de los aspectos que hay que tener en cuenta a la hora de diseñar una infraestructura para ejecutar aplicaciones paralelas intensivas.

b. Objetivos de aprendizaje

Al finalizar este bloque temático, el alumno deberá ser capaz de:

- Conocer los diferentes tipos de clusters de ordenadores que existen.
- Seleccionar los componentes hardware más adecuados para diseñar una infraestructura de computación de altas prestaciones en función del tipo de aplicación que se quiera ejecutar.
- Conocer los diferentes elementos software que se pueden utilizar para configurar una infraestructura de computación de altas prestaciones.
- Entender las diferentes posibilidades que existen para acondicionar una infraestructura computacional compleja.
- Conocer diferentes fuentes de información sobre los últimos avances en infraestructuras computacionales complejas.

c. Contenidos

Tema 1.- Estrategias de diseño de clusters

- 1.1.- Introducción: tipos de clusters y componentes.
- 1.2.- Centros de computación de altas prestaciones.
- 1.3.- Principales aplicaciones.
- 1.4.- Requerimientos hardware para cada tipo de aplicación.
- 1.5.- Ejemplos de centros de computación de altas prestaciones.

Tema 2.- Configuración y gestión de clusters

- 2.1.- Configuración de un cluster con Linux.
- 2.2.- Sistemas de ficheros para clusters.
- 2.3.- Sistemas de gestión de colas de trabajo.
- 2.4.- Monitorización del cluster.
- 2.5.- Soluciones integradas para clusters.

Tema 3.- Acondicionamiento de clusters

- 3.1.- Aspectos generales de diseño.
- 3.2.- Suministro eléctrico.
- 3.2.- Sistemas de refrigeración.
- 3.2.- Mejora de la eficiencia energética.



d. Métodos docentes

Se empleará:

- Clase magistral participativa y expositiva.
- Estudio de casos en aula.
- Aprendizaje colaborativo.

e. Plan de trabajo

Ver apartado 6 del presente documento.

f. Evaluación

Se evaluarán las entregas realizadas por los alumnos en las diferentes actividades que se propongan durante la exposición teórica de la asignatura. También se valorará el grado de participación e implicación del alumno en las actividades propuestas.

g. Bibliografía

Ver el apartado 8.c del presente documento.

h. Recursos necesarios

Ver el apartado 8.a del presente documento.

Bloque 2: Prácticas

Carga de trabajo en créditos ECTS:

a. Contextualización y justificación

Este bloque es totalmente práctico. Se pretende que el alumno aplique los conocimientos adquiridos en el Bloque 1 para configurar e instalar un cluster.

b. Objetivos de aprendizaje

Al finalizar este bloque temático, el alumno deberá ser capaz de:

- Configurar un cluster de ordenadores.
- Instalar el software necesario para ejecutar aplicaciones en paralelo.
- Administrar y gestionar un cluster.
- Ejecutar aplicaciones paralelas en un cluster.

c. Contenidos

Se plantearán diferentes hitos. Algunos de ellos son fundamentales y tienen que ser alcanzados para aprobar la asignatura. Otros son opcionales y servirán para poder establecer la nota final de la asignatura.



Hito Fundamental: configurar un cluster totalmente funcional. Esto incluye la configuración del sistema operativo, de un entorno de ejecución paralelo, de un sistema de colas, y de las librerías matemáticas fundamentales.

Hitos opcionales: a continuación se plantean algunas posibilidades para mejorar la funcionalidad del cluster.

- Evaluación de la capacidad computacional del cluster.
- Configuración de entornos paralelos adicionales.
- Implementación de un sistema de gestión energética del cluster.
- Integración de coprocesadores matemáticos en el cluster.
- Configuración de un sistema de ficheros paralelo.

Esta lista de hitos opcionales está abierta a sugerencias que pudieran darse durante el desarrollo de la asignatura.

d. Métodos docentes

Se empleará:

- Prácticas de laboratorio en grupos de alumnos.
- Aprendizaje por objetivos.
- Realización de memorias.
- Presentación de trabajos de grupo.

e. Plan de trabajo

Ver apartado 6 del presente documento.

f. Evaluación

La evaluación de la adquisición de competencias se basará en:

- Valoración de la actitud y participación del alumno en las prácticas.
- Valoración de los hitos alcanzados durante la realización de las prácticas.
- Informe realizado por grupos de alumnos sobre el caso práctico.
- Presentación oral de los aspectos más relevantes de la memoria realizada.

g. Bibliografía

Ver el apartado 8.c del presente documento.

h. Recursos necesarios

Ver el apartado 8.a del presente documento.



6. Temporalización

Primeramente se expondrán las clases teóricas durante las primeras 10 horas de la asignatura, y posteriormente se realizarán las sesiones de prácticas.

BLOQUE TEMÁTICO	CARGA ECTS	PERIODO PREVISTO DE DESARROLLO
Bloque 1: Contenidos teóricos	1.0 ECTS	Semanas 1 a 3
Bloque 2: Prácticas	2.0 ECTS	Semanas 3 a 7

7. Criterios de evaluación

INSTRUMENTO/PROCEDIMIENTO	PESO EN LA NOTA FINAL	OBSERVACIONES
Actitud y participación del alumno en las actividades desarrolladas en el aula	20%	Se valorará la realización de las actividades propuestas durante las clases de teoría.
Prácticas de laboratorio.	40%	Es condición necesaria para superar la asignatura que el alumno alcance el hito fundamental de las prácticas. También se valorará mediante observación sistemática en las propias sesiones de laboratorio el trabajo del alumno, y el cumplimiento de los diferentes objetivos marcados en las prácticas. Se tendrán en cuenta factores como la autonomía a la hora de resolver las dificultades que aparezcan y la iniciativa para incorporar objetivos novedosos a las prácticas.
Memoria final de las prácticas de laboratorio.	20%	Se valorará que la documentación de las actividades realizadas durante las prácticas sea completa y detallada, incluyendo todos aquellos factores que hayan podido afectar a la no consecución de los objetivos marcados, así como la discusión de las causas que los han ocasionado y posibles soluciones.
Exposición final del trabajo realizado en las prácticas.	20%	Se valorará que en la exposición sea concisa y ponga de relevancia los principales resultados y dificultades que se hayan tenido durante la realización de las prácticas. También se tendrán en cuenta las respuestas que se den a las preguntas que realice el profesor o cualquier otro alumno.

En el caso de que el alumno no consiga alcanzar el hito fundamental en las prácticas de laboratorio, obtendrá una calificación de 4 puntos sobre 10.

De forma totalmente excepcional, aquellos alumnos que no hayan alcanzado el hito fundamental podrán superar la asignatura siempre y cuando hayan mostrado suficiente interés durante las prácticas para intentar superar los problemas, y justifiquen detalladamente qué dificultades encontraron y qué estrategias siguieron para intentar superarlas. En este caso la calificación que obtendrán será de 5 puntos sobre 10.



8. Consideraciones finales

a. Recursos necesarios

El material que se vaya a utilizar en la asignatura estará disponible con suficiente antelación a través del Campus Virtual de la UVa (<http://campusvirtual.uva.es/>).

En la siguiente página web se muestra información de los recursos disponibles en el laboratorio: <http://www.ele.uva.es/~ivasan/HPC>

b. Métodos docentes

Clases de aula

Las clases de aula se desarrollarán a través de lección magistral para presentar a los alumnos de manera ordenada y sistemática los contenidos de la asignatura.

Consistirán en clases participativas que se desarrollarán mediante transparencias. En las transparencias se expondrán los conceptos utilizados para la exposición teórica, y estarán disponibles a los alumnos con suficiente antelación.

Durante el desarrollo de las clases de aula se propondrán diferentes actividades para que los alumnos recopilen, ordenen y presenten información sobre aspectos concretos de la asignatura.

Prácticas de laboratorio

Dependiendo del número de alumnos matriculados en la asignatura, las prácticas se realizarán de forma individual o en grupos.

Con las prácticas se pretende que los alumnos apliquen de forma práctica los conocimientos adquiridos durante las clases de aula.

Realización de memoria de prácticas

Una vez acabadas las prácticas, el alumno deberá realizar un informe en el que detalle y comente las actividades realizadas durante las prácticas. La memoria realizada tiene el objetivo de documentar de forma detallada el proceso de configuración del cluster y servir de guía para que otras personas puedan reproducir con éxito el trabajo desarrollado.

Sesión de exposiciones orales

Al finalizar la asignatura habrá una sesión dedicada a la exposición oral del trabajo realizado por los distintos grupos de prácticas. Cada grupo realizará una exposición oral al resto de grupos donde explicará de forma concisa el desarrollo de la configuración de su cluster, presentando los elementos más relevantes de la memoria de prácticas realizada. Estas presentaciones tienen como objetivo acostumbrar a los alumnos a sintetizar información y seleccionar la más relevante, y a realizar presentaciones orales. Durante las exposiciones orales el resto de alumnos podrán realizar las preguntas que consideren oportunas.



Visitas voluntarias a instalaciones

Se tiene pensado organizar las siguientes visitas voluntarias:

- Visita al Centro de Supercomputación de Castilla y León. Con esta visita se pretende que los alumnos conozcan las instalaciones del centro de computación de altas prestaciones de referencia de la comunidad.
- Visita al CPD de la Universidad de Valladolid. Con esta visita los alumnos podrán conocer las instalaciones de un centro de procesamiento de datos, que es el tipo de instalación más común que se podrán encontrar en una empresa.
- Visita al servidor de virtualización de la Escuela Técnica Superior de Ingeniería Informática de la Universidad de Valladolid. Con esta visita se pretende mostrar un ejemplo particular de instalación compleja dedicada a la virtualización.

Los detalles particulares de estas visitas se darán a conocer durante el desarrollo de la asignatura. Su realización dependerá del número de alumnos interesados, y de la financiación que se consiga para su realización.

c. Bibliografía

Durante la exposición teórica de la asignatura se indicarán todas aquellas referencias y documentos de utilidad para los alumnos, así como las principales fuentes de información que pueden consultar para mantenerse informados. A continuación se presenta una lista resumida que pretende abarcar los principales contenidos que se desarrollarán.

Bibliografía básica:

A continuación se indican una serie de libros que incluyen de una forma ordenada los aspectos fundamentales para configurar un cluster:

- Robert W. Lucke "Building clustered Linux systems", Prentice Hall, 2005.
- Charles Bookman, "Linux Clustering: Building and Maintaining Linux Clusters", New Riders Publishing, 2003.
- Thomas Sterling, "Beowulf Cluster Computing with Linux", MIT Press, 2001.

Bibliografía y recursos complementarios:

Hoy en día la información relativa a sistemas computacionales de altas prestaciones se encuentra dispersa en la red, y en la gran mayoría de las ocasiones es necesario acudir a los fabricantes de los componentes. A continuación se diferencian enlaces que contienen información de interés para la asignatura:

Centros de computación de altas prestaciones:

- Fundación Centro de Supercomputación de Castilla y León (<http://www.fcsc.es/index.php/es/>).
- Centro de Supercomputación de Barcelona (<http://www.bsc.es/>).
- Red Española de Supercomputación (<http://www.bsc.es/marenostrum-support-services/current-activities>).
- Lista Top 500 de los centros de supercomputación más potentes del mundo (<http://www.top500.org/>).
- Lista Top 500 de los centros de supercomputación más eficientes del mundo (<http://www.green500.org/>).
- Proyecto para crear un cluster de alta eficiencia a partir de una red de ordenadores (<http://boinc.berkeley.edu/>).
- Sección de IBM dedicada a sus productos para HPC (<http://www-03.ibm.com/systems/technicalcomputing/solutions/bluegene/index.html>).



- Cray Inc., empresa de equipos de HPC (<http://www.cray.com/products/computing.aspx>).
- Sección de Dell dedicada a sus soluciones de HPC (<http://www.dellhpc solutions.com/>).

Compiladores:

- GCC (GNU Compiler Collection - <http://gcc.gnu.org/>).
- Compiladores optimizados para procesadores AMD (<http://developer.amd.com/tools/cpu-development/x86-open64-compiler-suite/>).
- Compiladores optimizados para procesadores Intel (<http://software.intel.com/en-us/intel-compilers>).

Librerías Matemáticas:

- BLAS (Basic Linear Algebra Subprograms – <http://www.netlib.org/blas/>).
- LAPACK (Linear Algebra PACKage – <http://www.netlib.org/lapack/>).
- ScaLAPACK (Scalable Linear Algebra PACKage – <http://www.netlib.org/scalapack/>).
- Librerías matemáticas optimizadas para procesadores AMD (AMD Core Math Library – ACML – <http://developer.amd.com/tools/cpu-development/amd-core-math-library-acml/>).
- Librerías matemáticas optimizadas para procesadores Intel (Intel Math Kernel Library – IMKL – <http://software.intel.com/en-us/intel-mkl>).
- Otras librerías matemáticas de interés (<http://www.netlib.org/utk/people/JackDongarra/la-sw.html>).

Coprocadores matemáticos:

- Productos de NVidia (<http://www.nvidia.com/object/nvidia-kepler.html>, http://www.nvidia.com/object/tesla_software.html).
- Coprocesador Intel Xeon Phi (<http://software.intel.com/en-us/mic-developer>).

Sistemas de ficheros:

- NFS (<http://nfs.sourceforge.net/>).
- Lustre (http://wiki.lustre.org/index.php/Main_Page).
- GlusterFS (<http://www.gluster.org/>).

Sistemas de gestión de colas de trabajo:

- PBS (<http://www.pbsworks.com/>).
- TORQUE (<http://www.adaptivecomputing.com/products/open-source/torque/>).
- Grid Engine (<http://gridscheduler.sourceforge.net/>, <http://www.gridengine.info/>).
- SLURM (<https://computing.llnl.gov/linux/slurm/>).

Acondicionamiento de infraestructuras:

- APC (American Power Conversion – <http://www.apc.com>, <http://www.apcdistributors.com/white-papers/>).
- Rittal (<http://www.rittal-corp.com/>).

Varios:

- Recursos generales relacionados con clusters de altas prestaciones (<http://www.buyya.com/cluster/>).
- Red books de IBM: colección de libros relacionados con sus sistemas (<http://www.redbooks.ibm.com/>).