

Guia docent

270647 - SCA - Supercomputació per Aplicacions Exigents

Última modificació: 16/07/2024

Unitat responsable: Facultat d'Informàtica de Barcelona
Unitat que imparteix: 701 - DAC - Departament d'Arquitectura de Computadors.
Titulació: MÀSTER UNIVERSITARI EN INNOVACIÓ I RECERCA EN INFORMÀTICA (Pla 2012). (Assignatura optativa).
Curs: 2024 **Crèdits ECTS:** 6.0 **Idiomes:** Anglès

PROFESSORAT

Professorat responsable: CARLOS ALVAREZ MARTINEZ
Altres: Primer quadrimestre:
CARLOS ALVAREZ MARTINEZ - 10
JOSE RAMON HERRERO ZARAGOZA - 10
JOSEP LARRIBA PEY - 10

CAPACITATS PRÈVIES

Comprensió bàsica de les arquitectures paral·leles, inclosos els sistemes multiprocessador de memòria compartida i distribuïda.

Habilitats de programació útils d'algun model de programació paral·lela.

COMPETÈNCIES DE LA TITULACIÓ A LES QUALS CONTRIBUEIX L'ASSIGNATURA

Específiques:

CEC3. Capacitat per aplicar solucions innovadores i realitzar avanços en el coneixement que explotin els nous paradigmes de la Informàtica, particularment en entorns distribuïts.

CEE3.3. Capacidad para entender las necesidades computacionales de problemas de disciplinas distintas de la informática y efectuar contribuciones significativas en equipos multidisciplinarios que usen la computación.

CEE4.2. Capacidad de analizar, evaluar, diseñar y optimizar software considerando la arquitectura y de proponer nuevas técnicas de optimización.

Genèriques:

CG1. Capacidad para aplicar el método científico en el estudio y análisis de fenómenos y sistemas en cualquier ámbito de la Informática, así como en la concepción, diseño e implantación de soluciones informáticas innovadoras y originales.

CG3. Capacidad para el modelado matemático, cálculo y diseño experimental en centros tecnológicos y de ingeniería de empresa, particularmente en tareas de investigación e innovación en todos los ámbitos de la Informática.

CG5. Capacidad para aplicar soluciones innovadoras y realizar avances en el conocimiento que exploten los nuevos paradigmas de la Informática, particularmente en entornos distribuidos.

METODOLOGIES DOCENTS

Durant el curs es realitzaran dos tipus d'activitats:

- Activitats centrades en l'adquisició de coneixements teòrics.
- Activitats centrades en l'adquisició de coneixements mitjançant experimentació amb la implementació i avaluació empírica al laboratori dels mecanismes explicats a nivell teòric.

Les activitats teòriques inclouen classes expositives participatives on s'expliquen els continguts bàsics del curs. Les activitats pràctiques inclouen seminaris de laboratori emprant l'ordinador portàtil de l'alumne/a, on els alumnes implementen els mecanismes descrits a les classes expositives. Els seminaris requereixen d'una preparació prèvia mitjançant la lectura de l'enunciat i la documentació de suport, i una elaboració posterior de les conclusions obtingudes en un informe.

OBJECTIUS D'APRENTATGE DE L'ASSIGNATURA

- 1.L'estudiant ha de ser capaç de comprendre la complexitat dels diferents algorismes, identificar les parts computacionalment intensives d'una simulació o processament de dades, i decidir quines parts necessiten optimitzar i paralelitzar-se.
- 2.L'estudiant ha de ser capaç de dissenyar i implementar algorismes eficients de simulació paral·lela i processament de dades usant un model de programació paral·lela.
- 3.L'estudiant ha de ser capaç d'avaluar els diferents compromisos (robustesa, cost computacional, escalabilitat) per a seleccionar un algoritme específic per a un problema de simulació o processament de dades

HORES TOTS DE DEDICACIÓ DE L'ESTUDIANTAT

Tipus	Hores	Percentatge
Hores aprenentatge autònom	96,0	64.00
Hores grup gran	54,0	36.00

Dedicació total: 150 h

CONTINGUTS

Introducció

Descripció:

- Introducció: descripció general
- El mètode científic modern
- Simulació i optimització
- HPC vs. HTC
- Simulacions numèriques
- Límits de la paral·lelització
- Evolució i límits dels sistemes HPC

Introducció a les Simulacions numèriques

Descripció:

- De models a algorismes
- Discretització i PDE
- Tipus de PDEs: el·líptiques, parabòliques, hiperbòliques
- Del problema a les matemàtiques i la solució
- Esquemes numèrics: Explicit vs Implicit
- Diferències finites i elements finits

Resolució directe de sistemes lineals d'equacions a gran escala

Descripció:

- Sistemes triangulars i paral·lelització
- Eliminació de Gauss
- Factorització LU
- Mètodes de particions
- HPL

Resolució iterativa de sistemes lineals d'equacions a gran escala

Descripció:

- Mètodes directes vs iteratius
- Jacobi
- Paral·lelització de mètodes iteratius
- Gauss-Seidel i SOR
- Mètodes de Krylov i preconditionament: HPCG
- Programari per mètodes numèrics: BLAS, LAPACK, etc.

Casos pràctics de sistemes numèrics

Descripció:

- Models de programació en paral·lel
- Paral·lelisme i granularitat
- Paral·lelisme de blocs
- Resolució sistemes dispersos

Introducció als problemes de ML

Descripció:

- Introducció a les operacions comunes de les DNNs
- Aplicacions de DNNs més comuns
- Arquitectures per DNNs
- Optimitzacions per DNNs
- DNNs i precisió
- Paral·lelisme en DNNs

Alineació de seqüència

Descripció:

Es tracta d'una forma d'ordenar les seqüències d'ADN, ARN o proteïna per identificar regions de semblança que poden ser conseqüència de relacions funcionals, estructurals o evolutius entre les seqüències.

Les estratègies de cerca, puntuació i de paral·lelització són importants per superar aquest repte.

Casos Pràctics

Descripció:

- DNNs i precisió
- Paral·lelisme en DNNs
- Algorismes de Sequence Alignment

Introducció al nucli de DBMSs i l'execució de consultes en aquests sistemes

Descripció:

Aquest tema té com a objectiu comprendre les diferents capes de programari d'un DBMS i com interactuen, les diferents complexitats que encarnen i com la seva interacció determina el rendiment d'aquests SGBD.

Aplicacions de Big Data i bases de dades de grafs

Descripció:

Aquesta sessió té l'objectiu de presentar als estudiants el disseny d'aquests sistemes de base de dades Big Data i grafs, com es construeixen i com la seva estructura es veu influenciada per l'estructura de les seves diferents capes de programari.

Benchmarking per a bases de dades

Descripció:

El benchmarking és un dels problemes més importants en el disseny i l'evolució de la base de dades. Aquesta part del curs serà dissenyada per comprendre els diferents esforços realitzats en benchmarking dels EUA i Europa per a bases de dades relacionals i grafs.

Casos pràctics de Big Data

Descripció:

Casos pràctics de Big Data

ACTIVITATS

Introducció

Descripció:

Seguiu les conferències, estudeu els materials i les pràctiques.

Objectius específics:

1

Competències relacionades:

CEE3.3. Capacidad para entender las necesidades computacionales de problemas de disciplinas distintas de la informática y efectuar contribuciones significativas en equipos multidisciplinares que usen la computación.

CG1. Capacidad para aplicar el método científico en el estudio y análisis de fenómenos y sistemas en cualquier ámbito de la Informática, así como en la concepción, diseño e implantación de soluciones informáticas innovadoras y originales.

CG3. Capacidad para el modelado matemático, cálculo y diseño experimental en centros tecnológicos y de ingeniería de empresa, particularmente en tareas de investigación e innovación en todos los ámbitos de la Informática.

Dedicació: 2h

Grup gran/Teoria: 2h

Part I: Aplicacions Numèriques

Descripció:

Seguiu les conferències, estudiï els materials i les pràctiques.

Objectius específics:

1, 2, 3

Competències relacionades:

CEC3. Capacitat per aplicar solucions innovadores i realitzar avanços en el coneixement que explotin els nous paradigmes de la Informàtica, particularment en entorns distribuïts.

CEE3.3. Capacidad para entender las necesidades computacionales de problemas de disciplinas distintas de la informática y efectuar contribuciones significativas en equipos multidisciplinares que usen la computación.

CEE4.2. Capacidad de analizar, evaluar, diseñar y optimizar software considerando la arquitectura y de proponer nuevas técnicas de optimización.

CG5. Capacidad para aplicar soluciones innovadoras y realizar avances en el conocimiento que exploten los nuevos paradigmas de la Informática, particularmente en entornos distribuidos.

CG1. Capacidad para aplicar el método científico en el estudio y análisis de fenómenos y sistemas en cualquier ámbito de la Informática, así como en la concepción, diseño e implantación de soluciones informáticas innovadoras y originales.

CG3. Capacidad para el modelado matemático, cálculo y diseño experimental en centros tecnológicos y de ingeniería de empresa, particularmente en tareas de investigación e innovación en todos los ámbitos de la Informática.

Dedicació: 18h

Grup gran/Teoria: 8h

Aprenentatge autònom: 10h

Laboratori d'aplicacions numèriques

Dedicació: 22h

Grup petit/Laboratori: 8h

Activitats dirigides: 2h

Aprenentatge autònom: 12h

Lliurable: assignació en aplicacions numèriques

Descripció:

Assignació per a la part d'Aplicacions Numèriques. Per ser lliurat en Racó.

Objectius específics:

1, 2, 3

Competències relacionades:

CEC3. Capacitat per aplicar solucions innovadores i realitzar avanços en el coneixement que explotin els nous paradigmes de la Informàtica, particularment en entorns distribuïts.

CEE3.3. Capacidad para entender las necesidades computacionales de problemas de disciplinas distintas de la informática y efectuar contribuciones significativas en equipos multidisciplinares que usen la computación.

CEE4.2. Capacidad de analizar, evaluar, diseñar y optimizar software considerando la arquitectura y de proponer nuevas técnicas de optimización.

CG5. Capacidad para aplicar soluciones innovadoras y realizar avances en el conocimiento que exploten los nuevos paradigmas de la Informática, particularmente en entornos distribuidos.

CG1. Capacidad para aplicar el método científico en el estudio y análisis de fenómenos y sistemas en cualquier ámbito de la Informática, así como en la concepción, diseño e implantación de soluciones informáticas innovadoras y originales.

CG3. Capacidad para el modelado matemático, cálculo y diseño experimental en centros tecnológicos y de ingeniería de empresa, particularmente en tareas de investigación e innovación en todos los ámbitos de la Informática.

Dedicació: 12h

Aprenentatge autònom: 12h

Part II: Aplicacions no numèriques: ML i Bioinformàtica

Descripció:

Segueix les classes magistrals, estudia els materials i pràctiques

Objectius específics:

1, 2, 3

Competències relacionades:

CEC3. Capacitat per aplicar solucions innovadores i realitzar avanços en el coneixement que explotin els nous paradigmes de la Informàtica, particularment en entorns distribuïts.

CEE3.3. Capacidad para entender las necesidades computacionales de problemas de disciplinas distintas de la informática y efectuar contribuciones significativas en equipos multidisciplinares que usen la computación.

CEE4.2. Capacidad de analizar, evaluar, diseñar y optimizar software considerando la arquitectura y de proponer nuevas técnicas de optimización.

CG5. Capacidad para aplicar soluciones innovadoras y realizar avances en el conocimiento que exploten los nuevos paradigmas de la Informática, particularmente en entornos distribuidos.

CG1. Capacidad para aplicar el método científico en el estudio y análisis de fenómenos y sistemas en cualquier ámbito de la Informática, así como en la concepción, diseño e implantación de soluciones informáticas innovadoras y originales.

CG3. Capacidad para el modelado matemático, cálculo y diseño experimental en centros tecnológicos y de ingeniería de empresa, particularmente en tareas de investigación e innovación en todos los ámbitos de la Informática.

Dedicació: 20h

Grup gran/Teoria: 10h

Aprenentatge autònom: 10h

Laboratori aplicacions no numèriques

Dedicació: 16h

Grup petit/Laboratori: 6h

Aprenentatge autònom: 10h

Lliurable: Problema no numèric

Objectius específics:

1, 2, 3

Competències relacionades:

CEC3. Capacitat per aplicar solucions innovadores i realitzar avanços en el coneixement que explotin els nous paradigmes de la Informàtica, particularment en entorns distribuïts.

CEE3.3. Capacidad para entender las necesidades computacionales de problemas de disciplinas distintas de la informática y efectuar contribuciones significativas en equipos multidisciplinares que usen la computación.

CEE4.2. Capacidad de analizar, evaluar, diseñar y optimizar software considerando la arquitectura y de proponer nuevas técnicas de optimización.

CG5. Capacidad para aplicar soluciones innovadoras y realizar avances en el conocimiento que exploten los nuevos paradigmas de la Informática, particularmente en entornos distribuidos.

CG1. Capacidad para aplicar el método científico en el estudio y análisis de fenómenos y sistemas en cualquier ámbito de la Informática, así como en la concepción, diseño e implantación de soluciones informáticas innovadoras y originales.

CG3. Capacidad para el modelado matemático, cálculo y diseño experimental en centros tecnológicos y de ingeniería de empresa, particularmente en tareas de investigación e innovación en todos los ámbitos de la Informática.

Dedicació: 12h

Aprenentatge autònom: 12h

Part III: Problemes de Big Data

Descripció:

Segueix les classes magistrals, estudia els materials i pràctiques

Dedicació: 20h

Grup gran/Teoria: 10h

Aprenentatge autònom: 10h

Laboratori Big Data

Dedicació: 16h

Grup petit/Laboratori: 6h

Aprenentatge autònom: 10h

Lliurable Big Data

Objectius específics:

1, 2, 3

Competències relacionades:

CEC3. Capacitat per aplicar solucions innovadores i realitzar avanços en el coneixement que explotin els nous paradigmes de la Informàtica, particularment en entorns distribuïts.

CEE3.3. Capacidad para entender las necesidades computacionales de problemas de disciplinas distintas de la informática y efectuar contribuciones significativas en equipos multidisciplinares que usen la computación.

CEE4.2. Capacidad de analizar, evaluar, diseñar y optimizar software considerando la arquitectura y de proponer nuevas técnicas de optimización.

CG5. Capacidad para aplicar soluciones innovadoras y realizar avances en el conocimiento que exploten los nuevos paradigmas de la Informática, particularmente en entornos distribuidos.

CG1. Capacidad para aplicar el método científico en el estudio y análisis de fenómenos y sistemas en cualquier ámbito de la Informática, así como en la concepción, diseño e implantación de soluciones informáticas innovadoras y originales.

CG3. Capacidad para el modelado matemático, cálculo y diseño experimental en centros tecnológicos y de ingeniería de empresa, particularmente en tareas de investigación e innovación en todos los ámbitos de la Informática.

Dedicació: 12h

Aprenentatge autònom: 12h

SISTEMA DE QUALIFICACIÓ

El curs s'avaluarà amb una nota parcial per bloc de continguts (Numèrics, No numèrics i Big Data). Cada bloc de continguts tindrà el mateix pes a la nota final:

$$\text{Nota} = A1 / 3 + A2 / 3 + A3 / 3$$

on

Ai: = Nota del bloc i (i d'1 a 3)

i

- Cada part del curs té alguns lliurables curts i parcials (PD) i un lliurable del projecte final (FPD).

BIBLIOGRAFIA

Bàsica:

- Burnett, D.S. Finite element analysis: from concepts to applications. Addison-Wesley, 1987. ISBN 0201108062.
- Chapra, S.C.; Canale, R.P. Numerical methods for engineers. 8th ed. New York: McGraw-Hill, 2021. ISBN 9781260571387.
- Pržulj, N. Analyzing network data in biology and medicine: an interdisciplinary textbook for biological, medical and computational scientists. Cambridge: Cambridge University Press, 2019. ISBN 9781108432238.
- Saad, Y. Iterative methods for sparse linear systems. 2nd ed. SIAM, 2003. ISBN 0898715342.
- Quinn, M.J. Parallel programming in C with MPI and OpenMP. International edition. McGraw-Hill, 2003. ISBN 0071232656.
- Briggs, W.L.; Henson, V.E.; McCormick, S.F. A multigrid tutorial. 2nd ed. Society for Industrial and Applied Mathematics, 2000. ISBN 0898714621.
- Tramontano, A. The ten most wanted solutions in protein bioinformatics. Chapman and Hall/CRC, 2005. ISBN 1584884916.
- Gray, J.; Reuter, A. Transaction processing: concepts and techniques. Morgan Kaufmann, 1993. ISBN 1558601902.

Complementària:

- Cohen, G.C. Higher-order numerical methods for transient wave equations. Przulj, Natasa. Springer, 2001. ISBN 354041598X.
- Fichtner, A. Full seismic waveform modelling and inversion. Springer, 2011. ISBN 9783642158063.
- Wesseling, P. An introduction to multigrid methods. Edwards, 2004. ISBN 1930217080.

RECURSOS

Enllaç web:

- <http://glaros.dtc.umn.edu/gkhome/views/metis/>. Mesh partitioner software
- <http://openmp.org/>. OpenMP
- <http://www.mgnet.org/>. MGNet
- <http://www.netlib.org/utk/people/JackDongarra/la-sw.html>. Freely Available Software for Linear Algebra
- <http://www.open-mpi.org/>. MPI library