270649 - SA - Supercomputers Architecture

Coordinating unit: 270 - FIB - Barcelona School of Informatics
Teaching unit: 701 - DAC - Department of Computer Architecture
Academic year: 2019
Degree: MASTER’S DEGREE IN INNOVATION AND RESEARCH IN INFORMATICS (Syllabus 2012). (Teaching unit Optional)
ECTS credits: 6 Teaching languages: English

Prior skills
Programming in C and Linux basics will be expected in the course. Prior exposure to parallel programming constructions, experience with linear algebra/matrices or machine learning knowledge, will be very helpful.

Teaching methodology
The theoretical part of the course will follow the slides designed by the teacher during theory class. The practical component is the most important part of this subject. In this course the "learn by doing" method is used, with a set of Hands-on, based on problems that the students must carry out throughout the course. The course will be marked by continuous assessment which ensures constant and steady work. The method is also based on teamwork and a 'learn to learn' approach reading and presenting papers. Thus the student is able to adapt and anticipate new technologies that will arise in the coming years.

Course Activities:
Class attendance and participation: Regular and consistent attendance is expected and to be able to discuss concepts covered during class.

Lab activities: Hands-on sessions will be conducted during lab sessions using supercomputing facilities. Each hands-on will involve writing a lab report with all the results to be delivered one week later.

Homework Assignments: Homework will be assigned weekly that includes reading the documentation that expands the concepts introduced during lectures, and periodically will include reading research papers related with the lecture of the week, and prepare presentations (with slides).

Assessment: There will be 2 short midterm exams (and could be some pop quiz) along the course.

Student presentation: Students/groups randomly chosen will present the homework (presentations/projects).

Learning objectives of the subject
1. To train students to follow by themselves the continuous development of supercomputing systems that enable the convergence of advanced analytic algorithms as artificial intelligence.
270649 - SA - Supercomputers Architecture

Study load

| Total learning time: 150h | Theory classes: 24h 16.00% | Practical classes: 12h 8.00% | Laboratory classes: 12h 8.00% | Guided activities: 6h 4.00% | Self study: 96h 64.00% |
# Content

## Course content and motivation

Degree competences to which the content contributes:

## Supercomputing Basics

Degree competences to which the content contributes:

## Supercomputer Building Blocks

Degree competences to which the content contributes:

## Supercomputer Software Stack

Degree competences to which the content contributes:

## Parallel Programming Models: OpenMP

Degree competences to which the content contributes:

## Parallel Programming Models: MPI

Degree competences to which the content contributes:

## Parallel Performance Metrics and Measurements

Degree competences to which the content contributes:

## Supercomputer Building Blocks for AI servers

Degree competences to which the content contributes:

## Coprocessors and Programming Models

Degree competences to which the content contributes:
# Powering Artificial Intelligence, Machine Learning and Deep Learning with Supercomputing

**Degree competences to which the content contributes:**

## Parallel platforms for AI and its software stack

**Degree competences to which the content contributes:**

## Distributed AI platforms and its software stack

**Degree competences to which the content contributes:**

## Towards Exascale Computing

**Degree competences to which the content contributes:**
## Planning of activities

| Course content and motivation | Hours: 4h  
| Theory classes: 2h  
| Practical classes: 0h  
| Laboratory classes: 0h  
| Guided activities: 0h  
| Self study: 2h |

### Specific objectives:

1

| Supercomputing Basics | Hours: 4h  
| Theory classes: 2h  
| Practical classes: 0h  
| Laboratory classes: 0h  
| Guided activities: 0h  
| Self study: 2h |

| HPC Building Blocks (general purpose blocks) | Hours: 4h  
| Theory classes: 2h  
| Practical classes: 0h  
| Laboratory classes: 0h  
| Guided activities: 0h  
| Self study: 2h |

| HPC Software Stack (general purpose blocks) | Hours: 4h  
| Theory classes: 2h  
| Practical classes: 0h  
| Laboratory classes: 0h  
| Guided activities: 0h  
| Self study: 2h |

| Parallel Programming Models: OpenMP | Hours: 4h  
| Theory classes: 2h  
| Practical classes: 0h  
| Laboratory classes: 0h  
| Guided activities: 0h  
| Self study: 2h |
# Parallel Programming Models: MPI

**Hours:** 4h  
Theory classes: 2h  
Practical classes: 0h  
Laboratory classes: 0h  
Guided activities: 0h  
Self study: 2h

---

# Parallel Performance Metrics and Measurements

**Hours:** 4h  
Theory classes: 2h  
Practical classes: 0h  
Laboratory classes: 0h  
Guided activities: 0h  
Self study: 2h

---

# HPC Building Blocks for AI servers

**Hours:** 4h  
Theory classes: 2h  
Practical classes: 0h  
Laboratory classes: 0h  
Guided activities: 0h  
Self study: 2h

---

# Coprocessors and Programming Models

**Hours:** 4h  
Theory classes: 2h  
Practical classes: 0h  
Laboratory classes: 0h  
Guided activities: 0h  
Self study: 2h

---

# Powering Artificial Intelligence, Machine Learning and Deep Learning with Supercomputing

**Hours:** 4h  
Theory classes: 2h  
Practical classes: 0h  
Laboratory classes: 0h  
Guided activities: 0h  
Self study: 2h
| Parallel AI platforms and its software stack | Hours: 4h  
Theory classes: 2h  
Practical classes: 0h  
Laboratory classes: 0h  
Guided activities: 0h  
Self study: 2h |
| Distributed AI platforms and its software stack | Hours: 4h  
Theory classes: 2h  
Practical classes: 0h  
Laboratory classes: 0h  
Guided activities: 0h  
Self study: 2h |
| Conclusions and remarks: Towards Exascale Computing | Hours: 4h  
Theory classes: 2h  
Practical classes: 0h  
Laboratory classes: 0h  
Guided activities: 0h  
Self study: 2h |
| 1- Supercomputing Building Blocks: Marenostrom visit | Hours: 4h 12m  
Theory classes: 0h  
Practical classes: 0h  
Laboratory classes: 2h  
Guided activities: 0h 12m  
Self study: 2h |
| 2- Getting Started with Supercomputing | Hours: 4h 12m  
Theory classes: 0h  
Practical classes: 0h  
Laboratory classes: 2h  
Guided activities: 0h 12m  
Self study: 2h |
### 3- Getting Started with Parallel Programming Models

**Hours:** 4h 06m  
- Theory classes: 0h  
- Practical classes: 0h  
- Laboratory classes: 2h  
- Guided activities: 0h 06m  
- Self study: 2h

### 4- Getting Started with Parallel Performance Metrics

**Hours:** 4h 12m  
- Theory classes: 0h  
- Practical classes: 0h  
- Laboratory classes: 2h  
- Guided activities: 0h 12m  
- Self study: 2h

### 5- Getting Started with Parallel Performance Model - I

**Hours:** 4h 12m  
- Theory classes: 0h  
- Practical classes: 0h  
- Laboratory classes: 2h  
- Guided activities: 0h 12m  
- Self study: 2h

### 6- Getting Started with Parallel Performance Model - II

**Hours:** 4h 06m  
- Theory classes: 0h  
- Practical classes: 0h  
- Laboratory classes: 2h  
- Guided activities: 0h 06m  
- Self study: 2h

### 7- Getting Started with GPU based Supercomputing

**Hours:** 4h 06m  
- Theory classes: 0h  
- Practical classes: 0h  
- Laboratory classes: 2h  
- Guided activities: 0h 06m  
- Self study: 2h
<table>
<thead>
<tr>
<th>Topic</th>
<th>Hours</th>
</tr>
</thead>
<tbody>
<tr>
<td>8- Getting Started with CUDA programming model</td>
<td>4h 12m</td>
</tr>
<tr>
<td></td>
<td>Theory classes: 0h</td>
</tr>
<tr>
<td></td>
<td>Practical classes: 0h</td>
</tr>
<tr>
<td></td>
<td>Laboratory classes: 2h</td>
</tr>
<tr>
<td></td>
<td>Guided activities: 0h 12m</td>
</tr>
<tr>
<td></td>
<td>Self study: 2h</td>
</tr>
<tr>
<td>9- Getting Started with Deep Learning Frameworks in a Supercomputer</td>
<td>4h 06m</td>
</tr>
<tr>
<td></td>
<td>Theory classes: 0h</td>
</tr>
<tr>
<td></td>
<td>Practical classes: 0h</td>
</tr>
<tr>
<td></td>
<td>Laboratory classes: 2h</td>
</tr>
<tr>
<td></td>
<td>Guided activities: 0h 06m</td>
</tr>
<tr>
<td></td>
<td>Self study: 2h</td>
</tr>
<tr>
<td>10- Getting Started with Deep Learning basic model</td>
<td>4h 12m</td>
</tr>
<tr>
<td></td>
<td>Theory classes: 0h</td>
</tr>
<tr>
<td></td>
<td>Practical classes: 0h</td>
</tr>
<tr>
<td></td>
<td>Laboratory classes: 2h</td>
</tr>
<tr>
<td></td>
<td>Guided activities: 0h 12m</td>
</tr>
<tr>
<td></td>
<td>Self study: 2h</td>
</tr>
<tr>
<td>11- Getting Started with a Deep Learning real problems and its solutions</td>
<td>4h 06m</td>
</tr>
<tr>
<td></td>
<td>Theory classes: 0h</td>
</tr>
<tr>
<td></td>
<td>Practical classes: 0h</td>
</tr>
<tr>
<td></td>
<td>Laboratory classes: 2h</td>
</tr>
<tr>
<td></td>
<td>Guided activities: 0h 06m</td>
</tr>
<tr>
<td></td>
<td>Self study: 2h</td>
</tr>
<tr>
<td>12- Getting Started with parallelization of a Deep Learning problems</td>
<td>4h 06m</td>
</tr>
<tr>
<td></td>
<td>Theory classes: 0h</td>
</tr>
<tr>
<td></td>
<td>Practical classes: 0h</td>
</tr>
<tr>
<td></td>
<td>Laboratory classes: 2h</td>
</tr>
<tr>
<td></td>
<td>Guided activities: 0h 06m</td>
</tr>
<tr>
<td></td>
<td>Self study: 2h</td>
</tr>
</tbody>
</table>
13- Getting Started with a distributed Deep Learning problems

| Hours: 4h 06m |
| Theory classes: 0h |
| Practical classes: 0h |
| Laboratory classes: 2h |
| Guided activities: 0h 06m |
| Self study: 2h |

Qualification system

The evaluation of this course will take into account different items (tentative):

- Attendance (minimum 80% required) & participation in class will account for 15% of the grade.
- Homework, papers reading, paper presentations, will account for 25% of the grade.
- Exams will account for 15% of the grade.
- Lab sessions (+ Lab reports) will account for 45% of the grade.

Bibliography

Basic:

Jordi Torres. Class handouts and materials associated with this class. (can be found on the Racó-FIB web server), 2019.

Others resources:

Hyperlink

https://torres.ai/SA-MIRI/