

# CÉCI News by the Sysadmins

David Colignon, **Ariel Lozano** [ULiège]  
Juan Cabrera, Frédéric Wautelet [UNamur]  
Sébastien Skozlowskij [UMons]  
Raphaël Leplae, Michaël Waumans [ULB]  
Damien François, Bernard Van Renterghem, Olivier Mattelaer, Thomas Keutgen [UCL]

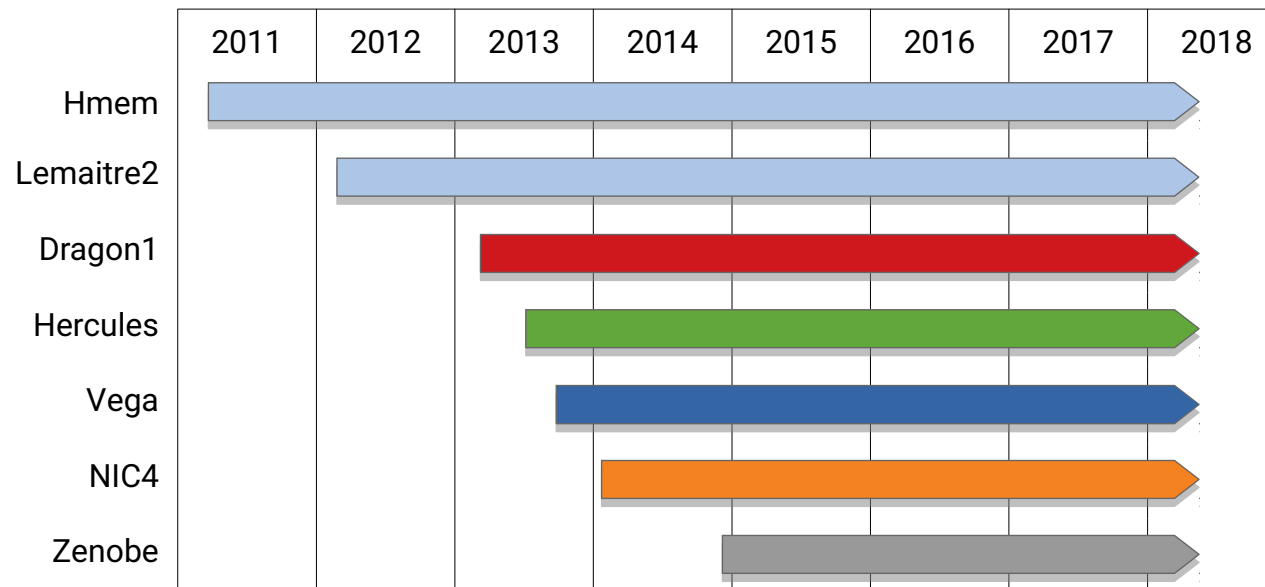
4<sup>th</sup> May 2018 - CÉCI Scientific Day @ UNamur

# Outline

- Clusters Upgrades
- CÉCI Slurm Federation
- CÉCI support page
- Slurm scripts generation tool
- SSH config wizard
- Common CÉCI storage

# Clusters upgrade

Cluster	Host	CPU type	CPU count*	RAM/node	Network	Filesystem**	Accelerator	Max time	Preferred jobs***
<b>NIC4</b>	ULg	SandyBridge 2.0 GHz IvyBridge 2.0 GHz	2048 (120 x 16 + 8 x 16)	64 GB	QDR Ib	FHGFS 144 TB	None	2 days	☒ MPI
<b>Vega</b>	ULB	Bulldozer 2.1 GHz	2752 (43 x 64)	256 GB	QDR Ib	GPFS 70 TB	2x Tesla M2090	14 days	☒ serial / ☒ SMP / ☒ MPI
<b>Hercules</b>	UNamur	SandyBridge 2.20 GHz Westmere 2.66 GHz	896 (32 x 16 + 32 x 12)	36..128 GB	GbE	NFS 20 TB	3x Tesla S2050	63 days	☒ serial / ☒ SMP
<b>Dragon1</b>	UMons	SandyBridge 2.60 GHz	416 (26 x 16)	128 GB	GbE	RAID0 1.1 TB	4x Tesla C2075	41 days	☒ serial / ☒ SMP
<b>Lemaitre2</b>	UCL	Westmere 2.53 GHz	1380 (115 x 12)	48 GB	QDR Ib	Lustre 120 TB	3x Quadro Q4000	3 days	☒ MPI
<b>Hmem</b>	UCL	MagnyCours 2.2 GHz	816 (17 x 48)	128..512 GB	QDR Ib	FHGFS 30 TB	None	15 days	☒ SMP



# Clusters upgrade: Lemaitre3 @ UCL

## ▫ Hard:

- **80 compute nodes** with 2 x Intel Xeon Gold 5118@2.3 Ghz (24 cores)
- Skylake Xeons with AVX, AVX2, AVX-512 extensions
- Total of **1920 cores**
- **96 GB of RAM** per node (4 GB per core)
- Intel Omni-Path Interconnect 56 Gb/s bandwidth between nodes
- **580 TB** Global Scratch with BeeGFS parallel file system
- Cluster with fast interconnect meant for **big MPI jobs**

## ▫ Soft:

- First cluster on the CÉCI Slurm Federation
- Support for Singularity containers
- Intel Parallel Studio XE 2018

# Clusters upgrade: Hercules and Dragon1

Planned for ~late 2018 early 2019

- Hercules2 @ UNamur:
  - New nodes will be added
  - High memory nodes will be available (at least 1TB of RAM)
  
- Dragon1 @ UMons:
  - New nodes will be added with high frequency Intel Xeons (> 3GHz)
  - Latest generation GPUs Nvidia Volta V100 (~4)

# Clusters upgrade: Vega2

- Vega2 @ ULB:
  - Will be oriented towards HTC (**High Throughput Computing**) and High Performance Data Analysis (HPDA) works & needs, covering **Big Data** and **Machine Learning**
  - If you are working or moving into those fields please fill the survey

The screenshot shows the CÉCI website interface. At the top is a blue navigation bar with links for Clusters, News, Training, FAQ, Documentation, Support, and Contact, along with a 'Create/Manage Account' button. Below this is a large banner for the 'Consortium des Équipements de Calcul Intensif' (C.E.C.I.), featuring the logo and the text '6 clusters, 10k cores, 1 login, 1 home directory'. The main content area is divided into three sections: 'About' (describing CÉCI as a consortium of high-performance computing centers from UCL, ULB, ULg, UMONS, and UNAMUR), '10th CÉCI Scientific Meeting' (announcing a meeting on May 4th in Namur), and 'Latest News' (highlighting a survey on Big Data and Machine Learning needs). A red arrow points to the survey news item, which is enclosed in a red box. The survey news item includes the date 'THURSDAY, 03 MAY 2018' and the text: 'Survey on Big Data and Machine Learning needs. We are conducting a survey about current and future High Performance Data Analysis (HPDA) works & needs, covering BigData, DeepLearning, MachineLearning, AI & co. Research groups already active in those fields are our primary center of interest. However, those moving or intending to move into those fields are welcome to fill the survey too. Our objective is to identify concrete hardware and software requirements for the future CECI Vega2 cluster which will be oriented towards HTC (High Throughput Computing) and HPDA. You are therefore cordially invited to follow this link and fill the survey.'

CÉCI Clusters News Training FAQ Documentation Support Contact [Create/Manage Account](#)

## C.E.C.I.

Consortium des Équipements de Calcul Intensif

6 clusters, 10k cores, 1 login, 1 home directory

### About

CÉCI is the 'Consortium des Équipements de Calcul Intensif'; a consortium of high-performance computing centers of [UCL](#), [ULB](#), [ULg](#), [UMons](#), and [UNamur](#). [Read more](#).

### 10th CÉCI Scientific Meeting

The next CÉCI scientific day will take place on May 4th in Namur. Details and registration [here](#)

### Latest News

**Survey on Big Data and Machine Learning needs**

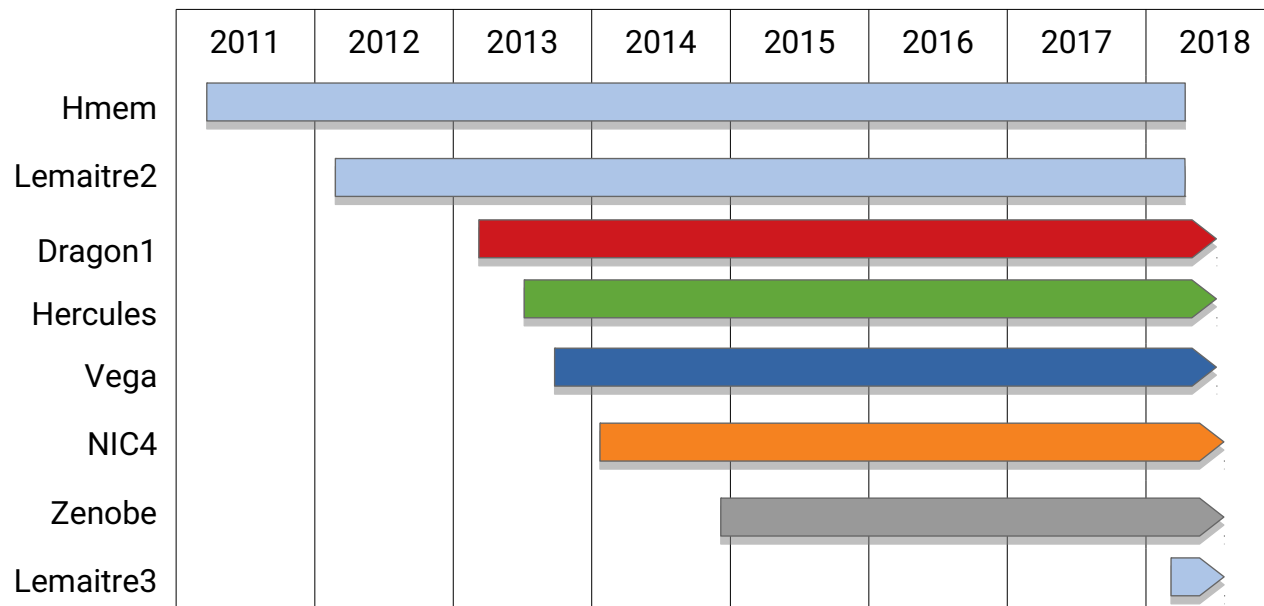
THURSDAY, 03 MAY 2018

We are conducting a survey about current and future High Performance Data Analysis (HPDA) works & needs, covering BigData, DeepLearning, MachineLearning, AI & co. Research groups already active in those fields are our primary center of interest. However, those moving or intending to move into those fields are welcome to fill the survey too. Our objective is to identify concrete hardware and software requirements for the future CECI Vega2 cluster which will be oriented towards HTC (High Throughput Computing) and HPDA.

You are therefore cordially invited to [follow this link](#) and fill the survey.

# Clusters upgrade: coming up timeline

- **Lemaitre3** already available since 30/04
- **Dragon1** and **Hercules** upgraded and expanded
- **Hmem** and **Lemaitre2** are planned to be **decommissioned!!**
- Only **Hmem 512 and 256 GB nodes** will temporarily stay until new high memory nodes are available on the CÉCI



# CÉCI Slurm Federation

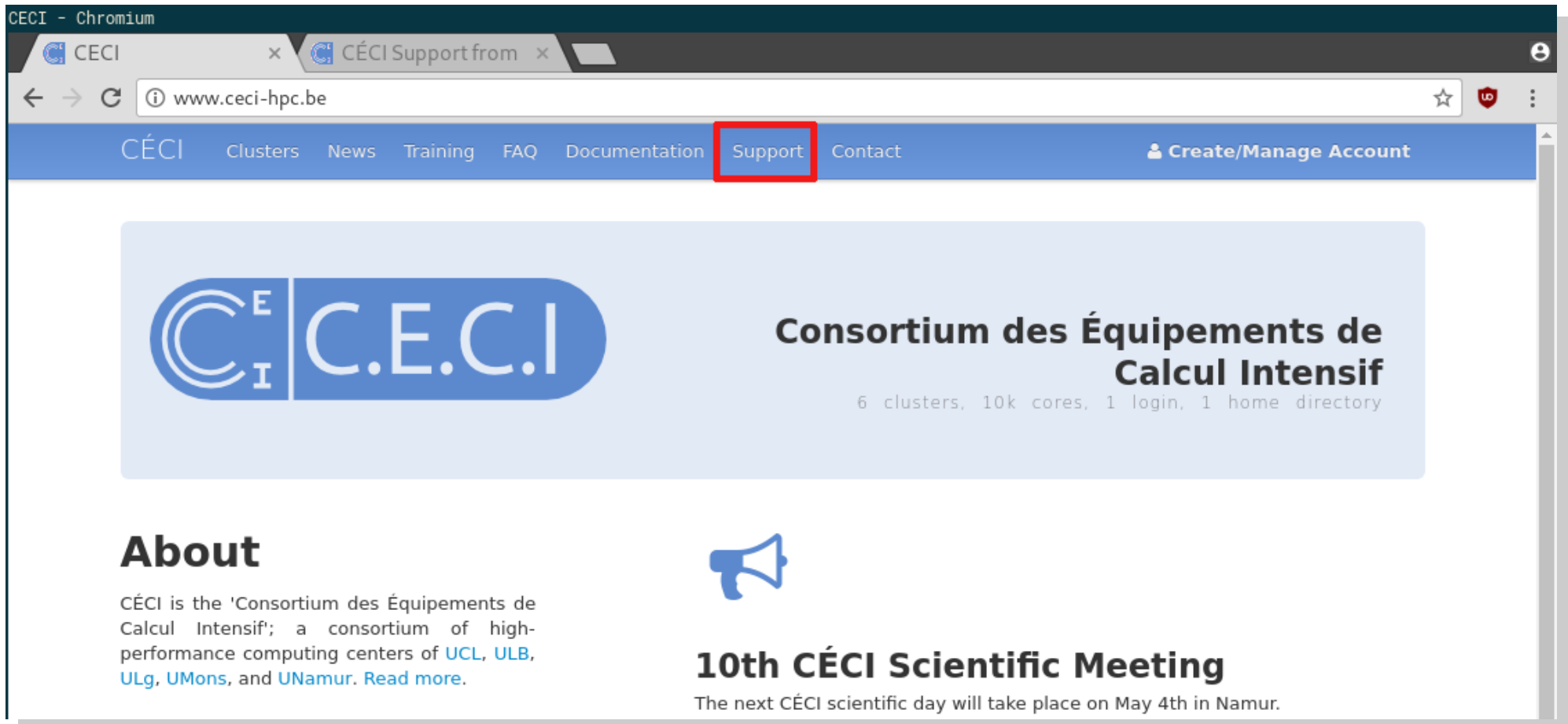
- New feature in Slurm as of November 2017
- Ability to see all jobs on all clusters with `squeue` on any of them
- Ability to submit a job to all clusters at once
  
- Lemaitre3 alone in the Federation at the moment
- New CÉCI clusters will join the Federation when installed

```
[dfr@lemaitre3 ~]$ sacctmgr show federation
Federation      Cluster ID      Features      FedS
-----
      ceci  lemaitre3  1  globalscratch,opa  A
```

```
[dfr@lemaitre3 ~]$ squeue --cluster lemaitre3
CLUSTER: lemaitre3
      JOBID PARTITION      NAME      USER ST
      67108991      batch mix_mpi_ dcoligno  R
```



# CÉCI support page



The screenshot shows a web browser window with the URL [www.cec-hpc.be](http://www.cec-hpc.be). The navigation menu includes 'CÉCI', 'Clusters', 'News', 'Training', 'FAQ', 'Documentation', 'Support', and 'Contact'. The 'Support' link is highlighted with a red rectangular box. Below the navigation bar, the main header features the C.E.C.I logo and the text 'Consortium des Équipements de Calcul Intensif' with the tagline '6 clusters, 10k cores, 1 login, 1 home directory'. The 'About' section describes CÉCI as a consortium of high-performance computing centers from UCL, ULB, ULg, UMon, and UNamur. A megaphone icon is positioned above the announcement for the '10th CÉCI Scientific Meeting', which is scheduled for May 4th in Namur.

- Since +1 year we have available the **CÉCI Support page** with a ticket system

# CÉCI support page

Welcome to the CÉCI support wizard. The wizard will help you help us help you ... by making sure you give us all the details we need in order to diagnose, and then solve, your problem. After you answer the wizard's questions, a button 'Generate email' will appear. Click on it and an email will open in your email program, ready to send. You can then modify it to your liking, and then click the 'Send' button of your email program. **Do not write your password in any field. This is not useful for us, and this compromises the security cluster.**

**Do you already have a CÉCI account ?**  
**If so, please enter your login below. Otherwise please enter your name.**

**What problem are you facing?**

- I cannot create or renew an account
- I cannot connect with SSH and/or copy files
- I have trouble compiling/installing software
- I can't submit a job to the cluster
- My job failed / A node crashed
- Other

- As you describe your problem you might be pointed to answers already provided in the FAQ/User docs
- Using the support wizard to report your issue **increases the amount of persons alerted**

# Interactive Slurm script generation

The screenshot shows a web browser window with the URL [www.cec-hpc.be/scriptgen.html](http://www.cec-hpc.be/scriptgen.html). The page has a blue header with the CÉCI logo and navigation links: Clusters, News, Training, FAQ, Documentation, Support, Contact, and a 'Create/Manage Account' button. A warning message states: 'Warning: this is still beta. Please send feedback to [damien.francois@uclouvain.be](mailto:damien.francois@uclouvain.be). Reload the page to reset.'

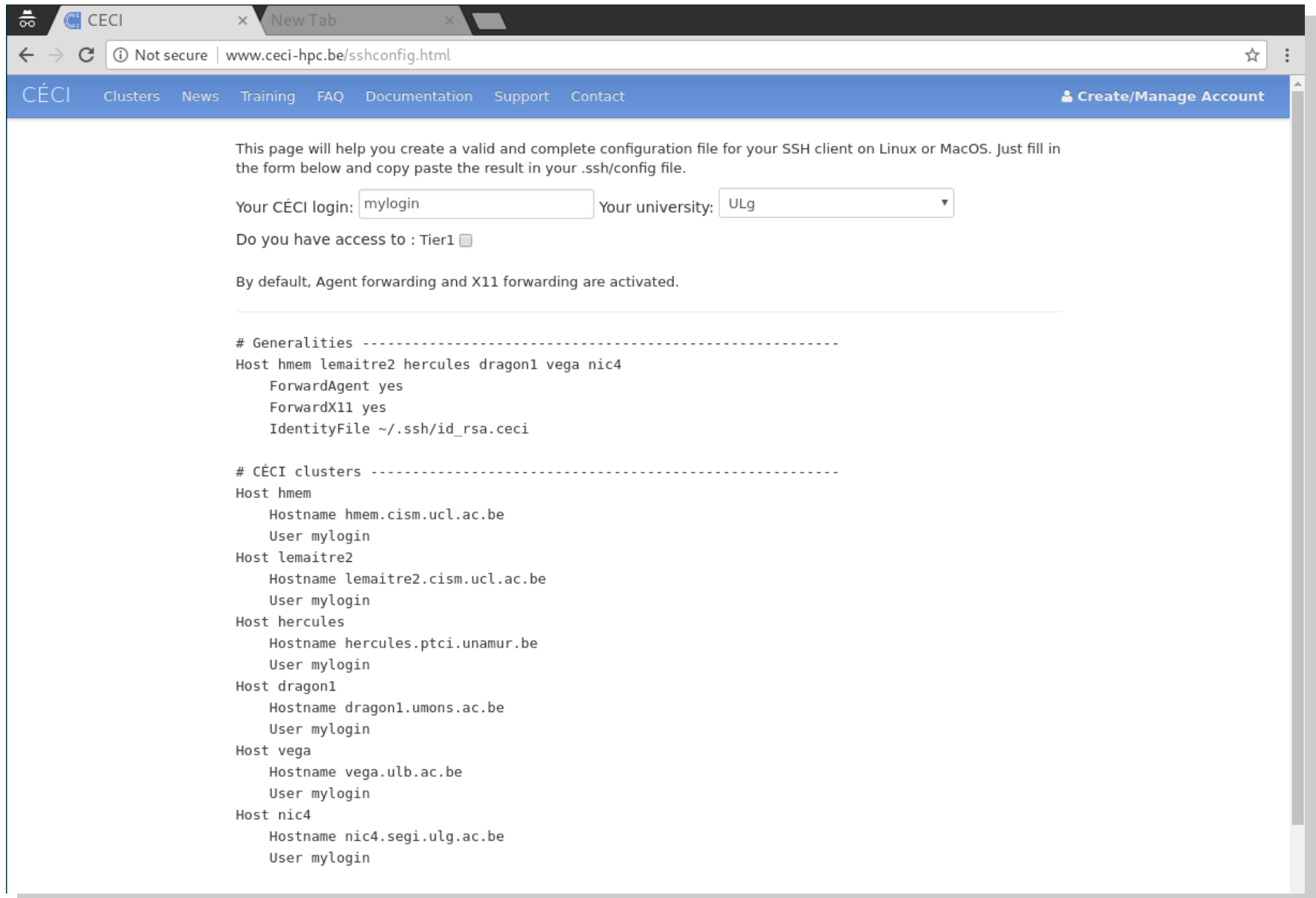
The main content is divided into three columns:

- 1. Describe your job**: Includes input fields for 'Email address' (user@example.com), 'Job name' (Some name), and 'Project' (Some project). It has a 'Parallelization paradigm(s)' section with three checkboxes: 'Embarassingly parallel / Job array', 'Shared memory / OpenMP', and 'Message passing / MPI'. The 'Job resources' section has input fields for 'Duration' (0 days, 1 hour, 0 minutes) and 'Memory' (2625 MB). The 'Filesystem' section has a dropdown menu set to '\$HOME'. A summary at the bottom shows: 'Total CPUs: 1 | Total Memory: 2625 MB | Total CPU.Hours: 1'.
- 2. Choose a cluster**: A list of radio buttons for cluster selection: NIC4 (selected), Vega, Lemaitre2, Hercules, Dragon1, HMEM, and Zenobe\*.
- 3. Copy-paste your script**: A text area containing a Slurm script:

```
#!/bin/bash
# Submission script for NIC4
#SBATCH --time=01:00:00 # hh:mm:ss
#
#SBATCH --ntasks=1
#SBATCH --mem-per-cpu=2625 # megabytes
#SBATCH --partition=defq
```

<http://www.cec-hpc.be/scriptgen.html>

# SSH config wizard



The screenshot shows a web browser window with the URL [www.cec-hpc.be/sshconfig.html](http://www.cec-hpc.be/sshconfig.html). The page title is "CÉCI" and the navigation menu includes "Clusters", "News", "Training", "FAQ", "Documentation", "Support", and "Contact". A "Create/Manage Account" link is visible in the top right corner. The main content area contains the following text:

This page will help you create a valid and complete configuration file for your SSH client on Linux or MacOS. Just fill in the form below and copy paste the result in your `.ssh/config` file.

Your CÉCI login:  Your university:

Do you have access to : Tier1

By default, Agent forwarding and X11 forwarding are activated.

---

```
# Generalities -----
Host hmem lemaitre2 hercules dragon1 vega nic4
  ForwardAgent yes
  ForwardX11 yes
  IdentityFile ~/.ssh/id_rsa.cec

# CÉCI clusters -----
Host hmem
  Hostname hmem.cism.ucl.ac.be
  User mylogin
Host lemaitre2
  Hostname lemaitre2.cism.ucl.ac.be
  User mylogin
Host hercules
  Hostname hercules.ptci.unamur.be
  User mylogin
Host dragon1
  Hostname dragon1.umons.ac.be
  User mylogin
Host vega
  Hostname vega.ulb.ac.be
  User mylogin
Host nic4
  Hostname nic4.segi.ulg.ac.be
  User mylogin
```

<http://www.cec-hpc.be/sshconfig.html>

# Common CÉCI storage

- Remember about the common storage for all CÉCI clusters:
  - `cd $CECIHOME`
  - `cd $CECITRSF`

The screenshot shows a web browser at [https://support.cec-hpc.be/doc/\\_contents/ManagingFiles/TheCommonFilesystem.html](https://support.cec-hpc.be/doc/_contents/ManagingFiles/TheCommonFilesystem.html). The navigation menu includes 'Documentation', which is highlighted with a red box and a red arrow pointing to the left sidebar. The sidebar contains a search bar and a list of topics under 'QUICK START - FIRST STEPS' and 'MANAGING FILES'. The 'Using the common filesystem' section is expanded in the sidebar. The main content area features the title 'Using the common filesystem' and a 'Note' box.

**Using the common filesystem**

All CÉCI clusters are connected to a central storage system that is visible to all compute nodes of all clusters. This system runs on a fast, dedicated, network. It will become the home of the users in the near future, but in this first phase, it is set up as an additional home besides the default, cluster-specific, home.

This storage system is installed at two CÉCI locations (at ULiège and UCL) and data are replicated synchronously on both locations to ensure data safety and a certain level of high availability. Moreover, on each site, a local cache is setup to mask the latencies of the network and make sure the user experience is as smooth as possible. Those caches are replicated asynchronously with the central storage, meaning that files that are written there will appear after some delay on the other clusters.

**Note**

There is no point in creating a file from a cluster and then quickly going to another cluster and running `ls` rapidly waiting for the file to appear. The file will be copied on the main storage only after 15 seconds, and the metadata will be copied on the other clusters after up to 60 seconds.

It also means that if you modify the same file from two different clusters, the result is undefined.

Thanks for listening!