Certification Strategy and Contributions
HPCCF Virtual Workshop

Christian Meesters (+ HPC Certification Forum)

HPC Group – Johannes Gutenberg-University of Mainz

2020-05-18
I hope everyone is doing well.

I hope we all be teaching *in personam*, soon-ish, again.
Outline

1. Certification Strategy
2. Designing Questions
3. Contributing
Two Types of Users

Exaggeration Warning
Exaggeration Warning

Tier I
- Sims

Tier II
- Sims analysis
- bits of being

Tier III
- every code

HPC

HPC + HTC + JTC
Exaggeration Warning

- resources are always limited
- teaching resources even more
- integration into HPCCF might offer more (still needed) courses
How Joe User looks on HPC

Exaggeration Warning

Most users

- ... use 3\textsuperscript{rd} party applications ...
- ... will need (yet not always visit) an intro course ...
- ... perhaps a scripting course ...
- ... only \textit{really} interested in workflows tailored for their need.

And will never leave their site for other HPC courses!
Exaggeration Warning

Only power users

- ... will select their topics ...
- ... will care to travel for computing topics ...
- ... will rarely need intro courses ...
Teachers, basically, get two benefits:

1. simpler / better content and course objective transparency

2. feedback from HPCCF with regard to course quality ("Did my students pass? How good?")

3. Plus: We have learned, that users appreciate when a course content apparently is not conceived in isolation.
When conceiving a Certification Strategy the different views are in our mind.
Selecting Questions

Questions are randomly chosen from a pool:

- the pool may itself be a bundle of sub-branches of the skill tree
- each question will have a number of right and wrong answers in case of multiple choice questions

All examinations will be based on different sets of questions.
On Cheating

1. By confronting with random questions no perfect preparation can be accomplished.

2. There is a time-limit per question.

3. A registration prior to a test session is required.

No online system without ID checks and other measures is safe against cheating! Yet, our measures will raise awareness.
Boosting Acceptance

Want to hire a scientist?
We intend to provide a (sub)set of question for prospective employers. This way they will have an idea of the background, if a solicitant waves a HPCCF-certificate.
Disclaimer

Some examples are inspired by Greg Wilson's book

Teaching Tech Together (CRC Press, 2020)

Some ideas are based on own experience, some on other sources.
Before diving into Question Design, note:

- a question can be asked with a certain aim
- different courses ask for different knowledge / skills
- questions need to be designed and chosen with care
Multiple Choice Questions (MCQs) are popular when designing e-learning tests . . .

**Question**

When are they most suitable?

Suppose you are teaching children and you give them this MCQ:

**Exercise – Type: Testing Conceptions**

What is $37 + 15$?

- a) 52 correct
- b) 42 child did not understand “carrying”
- c) 412 child treated every column separately
- d) 43 knows she has to carry 1, but to wrong column
The Young-Child question rephrased for newbies to the SLURM batch system:

Think of a cluster with 20 core nodes. If a job is submitted with the following parameterisation, how many nodes are reserved?

```
#SBATCH -n 20
#SBATCH -c 2
```

- (a) 2 correct
- (b) 4 user did correctly multiply, but is not aware of the 20 cores
- (c) 1 user did not multiply by \(-c\) 2
- (d) unknown without N-flag user did not understand the concept
MCQs aren’t everything:

1. Freetext (if short and explicit)
2. Filling in blanks (for code; to be implemented)
3. Parson Problem (can by done as MCQ; shown in a minute)
4. Tracing (can by done as MCQ; in a minute)
Fill in Blanks

Filling Blanks is a (technical) variation on Freetext. It is more specific and the *blank screen of horror* issue is avoided, whilst the test might be testing “vocabulary”. An example:

**Exercise – Type: Bash Operators**

Which operator has to be filled in the place of ‘\_’ to print the statement in line 3?

```
1: number=4
2: if [ $((number _ 2 )) -eq 0 ]; then
3:     echo "$number is even"
4: fi
```

**Hint**

The answer is a single character.
Parsons Problems, too, avoid the *blank screen of horror* problem and also the vocabulary testing. Instead they allow the examinee to concentrate on the control flow.

**Exercise – Type: Bash Loop & Math**

Rearrange these lines to sum the values.

1: done
2: values=(1 2 3 4)
3: for v in $values[@]; do
   total=$((total + v))

Real tasks can be longer and intricate - allowing test of control flow understanding.

**Note**

The answer can be a free text, e.g.: “2 3 4 1”, which is easy to parse and check.
Contributions via the HPCCF-Wiki

K4-B Job Scheduling

Background

Parallel computers are operated differently than a normal PC, all users must share the system. Therefore, various operative procedures are in place. Users must understand these concepts and procedures to be able to use the available resources of a system to run a parallel application. A workload manager/job scheduler controls how available hardware resources are distributed among the user requests (jobs).

Aim

To enable practitioners to comprehend and describe the basic architecture and concepts of resource allocation for an HPC system

Outcomes

- explain the concepts and procedures for resource allocation and job execution in an HPC environment
- run interactive jobs and batch jobs
- comprehend and describe the expected behavior of job scripts
- charge provided job scripts and embed them into shell scripts to run a variety of parallel applications
- analyze the output generated from a job scheduler and describe the cause of typically generated errors
- comprehend accounting principles (billing for the jobs)
- comprehend scheduling strategies that increase productivity

Subskills

- K4.1-B Introduction to job scheduling
- K4.2-B SLURM Workload manager

Links

- Submit a proposal for an examination question
Contributions via the HPCCF-Wiki II

Each HPCCF wiki page contains a link. It leads to a little form asking for:

- contact mail
- to select a learning objective from a pre-formatted list
- to supply the question you thought of
- and (in case of a multiple choice question) the possible answers.

**Evaluation Process**

Now, HPCCF-member evaluate the submitted question. If approved, it will be formatted and merged into the pool of questions for the chosen topic / skillset.
Certification: Assessment Prototype

1. User takes multiple-choice test online (any time!)
   - A combination of JavaScript and a web service
   - System selects number of questions randomly from a pool
     - By submitting questions the related usage-allowance is granted to HPCCF
     - In case of sufficient numbers the system draws from a pool of different possible answers (MCQ-case).

2. Choices are submitted to the web server

3. Manual approval of the result

4. Automatic creation of certificate and returned by email
   - Permanent computer-verifiable proof is created about certification of skills
     - Return a text version with GPG signature
     - Return a link that can be verified on hpc-certification.org

5. Privacy: minimize information stored on servers, keep some for statistics
   - Includes some measure to prevent cheating and brute forcing (e.g., delay)
Contributing to the Question Pool

Certification: Certificate

Text representation

-----BEGIN PGP SIGNED MESSAGE-----
Hash: SHA512
HPC Certification Forum Certificate
This text confirms that "Jane Doe" has successfully obtained the certificate "HPC driving license" (id: 1) at 02/2019.
Verification URL: https://hpc-certification.org/[...]
-----BEGIN PGP SIGNATURE-----
[...]
-----END PGP SIGNATURE-----
Thank You for Your Attention!