

Document version v1.2 (6 April 2023)

HEPScore23 (HS23) run procedures

WLCG HEP-SCORE Deployment Task Force

[1. Introduction](#)

[2. How to run HEPscore23](#)

[ulimit configuration \(reason and procedure\)](#)

[2.1 Run the HEP Benchmark Suite](#)

[2.2 Script procedure](#)

[3 Accounting procedures](#)

[3.1 Accounting strategy in action](#)

[Example1](#)

[Example2](#)

[References](#)

1. Introduction

This document provides instructions on how to execute the HEPscore23 benchmark and how to publish the results in the accounting system.

HEPScore23 will progressively replace HS06 starting April 2023. The accounting migration procedure has been officially endorsed by the WLCG MB during the December 20th, 2022 meeting [1].

The HEPscore TF strongly recommends following the outlined procedures to ensure a smooth transition. If assistance is needed, the support unit of HEPscore can be reached via GGUS tickets. In that case the Responsible unit to be selected is Benchmarking.

2. How to run HEPscore23

It is crucial that the server is fully dedicated to the benchmarking activity during the run, to ensure accurate measurements and prevent potential errors.

The server must have a minimum hardware configuration and include the

following packages:

- Container engine Apptainer (version 1.1.6 or higher),
- Python version 3.9 or higher
- python3-pip
- git

Hardware requirements:

- A disk space proportional to the number of available cores on the server (1 GB per logical core) is necessary to temporarily store the results;
- The server must have at least 2GB of RAM per logical core
- ulimit configuration (see details below)

The user will need pip and git to install HEPscore23 as a Python package.

ulimit configuration (reason and procedure)

*A workload of the HEPscore23 benchmark uses a multi-service approach for the reconstruction and starts multiple processes per core that stay idle waiting for their turn of processing. For machines with more than 100 CPU cores, this translates into more than 4096 processes, which is the default for normal (non-root) users on CentOS7. Therefore, HEPscore23 should run as root, or the user should be able to start more processes. This can be set with **ulimit** on CentOS7 by adding the line*

benchmark soft nproc unlimited

in /etc/security/limits.d/20-nproc.conf

It is necessary to start a new shell session after that change before running HEPscore23.

CVMFS (as image repository) configuration

Although it's not part of the standard configuration, it is possible to get the container images for the benchmark from CVMFS instead of from the gitlab registry. Some workloads of the HEPscore23 benchmark access several files in parallel on /cvmfs, about 200 files per CPU core. For bigger machines (more than 60 cores), it is necessary to adjust the CVMFS config and set the maximal number of open files (CVMFS_NFILES in /etc/cvmfs/default.local) value to about 200 times the number of cores. The new value is active after a remount of the CVMFS repository on the machine.

2.1 Run the HEP Benchmark Suite

While it is possible to install HEPscore23 manually, it is recommended to use the HEP Benchmark Suite alongside HEPscore23 to include in the benchmark report metadata about the server's running conditions. The metadata includes details about the server's CPU, RAM, disks, IP addresses, and other relevant information. The HEP Benchmark Suite can be installed using pip and git.

2.2 Script procedure

A bash script [2] has been developed to streamline the installation and running process. This script provides a fully comprehensive running procedure and enables the system administrator to install the HEP Benchmark Suite and HEPscore23, run the HEP Benchmark Suite, which in turn extracts the necessary metadata from the server, executes HEPscore23 and produces a final output document.

The HEP Benchmark Suite also offers the added benefit of being able to submit the benchmark results to the WLCG Benchmark DB (based on OpenSearch/ElasticSearch). To accomplish this, a valid X509 certificate (service, robot, user) must be available, and the certificate's DN must be authorized for the publication of the results.

To declare the DN users should open a ticket via GGUS. The procedure to extract the DN from the certificate is detailed in [3].

To use the bash script, users will need to provide a few mandatory custom parameters to declare the specific WLCG site on which the benchmark is running and to publish the results. The SITE parameter is essential for ensuring that the results are accurately attributed to the correct site when integrated into the WLCG Benchmark DB.

To run the script, users can use the following command line.

```
./run_HEPscore.sh -s SITE -p -c ./cert.pem -k ./key.pem
```

By default, the script will use the current directory to create a working directory where all necessary files will be stored, including container images, benchmark outputs, and temporary workload results. The working directory can be modified using the parameter `-w target_folder`.

During the execution the script reports the stdout of the HEP Benchmark Suite. If the execution completes successfully, it will print at the end information such as

```
<date>, hepbenchmarksuite.hepbenchmarksuite:cleanup [INFO] Successfully
completed all requested benchmarks
=====
BENCHMARK RESULTS FOR <hostname>
=====
Suite start: <start_date>
Suite end: <end_date>
Machine CPU Model: <name>
HEPScore Benchmark = <value>
```

Using the bash script ensures that the entire process is performed correctly, and it is recommended that users utilize it when installing and running HEPscore23.

3 Accounting procedures

The migration strategy for the accounting side is detailed by the Accounting TF in the document [4]. This strategy involves implementing software changes on the site side as well as APEL, EGI portal, WAU sides. To streamline the process and minimize the number of changes, several strategic approaches have been discussed within the WLCG collaboration, in particular at the Lancaster Workshop. These approaches have been endorsed by the WLCG Management Board [1].

To summarize, the transition from HS06 to HEPscore23 should be gradual and seamless. This will be achieved through the following measures:

- The HEPscore23 benchmark will use the same scale factor as HS06, which is fixed on a reference server.
- Sites are only expected to benchmark new resources with HEPscore23. Old servers do not need to be re-benchmarked for accounting purposes. This ensures that the installed capacity pledged by the sites will remain unchanged.
- Sites are free to re-benchmark their servers if they wish, but they are not required to submit this information to the accounting portal. However, they can still use the HEP Benchmark Suite to publish their results in the benchmark database, which is separate from the accounting infrastructure.

3.1 Accounting

How do these procedures reflect what is done in a given WLCG site?

Below we describe how to calculate the benchmarking factor depending on site configuration and how the report would look like in accordance with the new [specification](#).

Example1: Site with a different cluster per CPU model.

New resources won't be mixed with old resources.

| Cluster | Model | Num | Num of logical threads | Score per node | | Total score | | Score for accounting |
|---------|-----------------------------|-----|------------------------|----------------|------|-------------|--------|----------------------|
| | | | | HS06 | HS23 | HS06 | HS23 | |
| Old | 2x AMD EPYC 7702 64-core | 29 | 256 | 2643 | 2546 | 76647 | 73834 | 76647 |
| New | 2x AMD EPYC 7742 64-Core | 188 | 256 | 2917 | 2972 | 548396 | 558736 | 558736 |
| Total | | | | | | | 632570 | 635383 |

Suppose the site has 2 separate clusters, each cluster consisting of servers with the same CPU model. Labels “Old” and “New” identify the clusters included in production before 1st of April 2023 (Old) or after 1st of April 2023 (New). The table below summarizes the HS06 and HS23 scores per node and for the total installation.

Therefore, if the same benchmark would be used for both clusters, the site would provide 625043 HS06 or 632570 HEPscore23, but following the agreement to translate the HS06 score of the old machines with 1:1 ratio, the final total score accounted for that site is **635383**.

For the accounting reporting the same input numbers and configuration translate into the following reported **benchmarking factor**

| Cluster | Model | Num | Num of logical threads | Score per node | | Benchmarking factor (score per 1 processor core) |
|---------|-----------------------------|-----|------------------------|----------------|------|---|
| | | | | HS06 | HS23 | |
| Old | 2x AMD EPYC 7702 64-core | 29 | 256 | 2643 | 2546 | $2643/256=10.32$ |
| New | 2x AMD EPYC 7742 64-Core | 188 | 256 | 2917 | 2972 | $2972/256=11.6$ |

Since resources are not mixed they are accessible through different CEs. In

this case in accordance with the new accounting record [specification](#) the job accounting records will look like:

For old resources

```
APEL-summary-job-message: v0.4
Site: SOME-SITE
SubmitHost: <old_cluster_ce>
Month: 4
Year: 2023
GlobalUserName: <...>
WallDuration: 47248
CpuDuration: 46871
Processors: 1
NumberOfJobs: 3
InfrastructureType: grid
EarliestStartTime: ...
LatestEndTime: ...
ServiceLevel: {hepspec: 10.32}
```

For new resources

```
APEL-summary-job-message: v0.4
Site: SOME-SITE
SubmitHost: <new_cluster_ce>
Month: 4
Year: 2023
GlobalUserName: <...>
WallDuration: 47248
CpuDuration: 46871
Processors: 1
NumberOfJobs: 3
InfrastructureType: grid
EarliestStartTime: ...
LatestEndTime: ...
ServiceLevel: {HEPscore23: 11.6}
```

If using the APEL client, HEPscore will be configurable locally in the new version of the client in its client.cfg file. Example of the spec_updater section shown below:

```
site_name = MY-SITENAME
manual_spec1 = <old_cluster_ce>, hepspec, 10.32
manual_spec2 = <new_cluster_ce>, HEPscore23, 11.6
```

Example2

A site with a single cluster mixing all CPU models. We take exactly the same set of HW as in the previous example, just resources are all mixed.

Then first we need to calculate the contribution of both sets of resources in the overall capacity.

Fraction of old resources:

76647/635383=0.12

Correspondingly, fraction of new resources is **0.88**

Benchmarking factor for the mixed cluster will be **10.32*0.12 + 11.6*0.88=11.45**

The accounting job record will look like:

```
APEL-summary-job-message: v0.4
Site: SOME-SITE
SubmitHost: <old_cluster_ce>
Month: 4
Year: 2023
GlobalUserName: <...>
WallDuration: 47248
CpuDuration: 46871
Processors: 1
NumberOfJobs: 3
InfrastructureType: grid
EarliestStartTime: ...
LatestEndTime: ...
ServiceLevel: {HEPscore23: 11.45}
```

Please, pay attention, that in case we have a cluster with mixed resources having part of them benchmarked with HEPscore23, we do report as if the whole cluster has been benchmarked with HEPscore23.

If using the APEL client, HEPscore will be configurable locally in the new version of the client in its client.cfg file. Example of the spec_updater section shown below:

```
site_name = MY-SITENAME
manual_spec1 = <new_cluster_ce>, HEPscore23, 11.45
```

References

- [1] [WLCG Management Board \(20 Dec. 2022\)](#)
- [2] [Bash script on Gitlab](#)
- [3] [Procedure for the DN authorization](#)
- [4] [Operation Coordination changes for HEPscore](#)