



Advanced Machine Learning. Flash Simulation and bleeding edge applications

FlashSim: November status report

Lucio Anderlini

Istituto Nazionale di Fisica Nucleare, Sezione di Firenze

External Partner



Who we are

Staff members:

- Alessandro Bombini ^j, INFN
- Giuseppe Piparo ^l, INFN
- Maurizio Martinelli ^a, Università Milano Bicocca
- Simone Capelli ^a, Università Milano Bicocca
- Federica Maria Simone ⁱ, Politecnico di Bari
- Nicola De Filippis ⁱ, Politecnico di Bari
- Vieri Candelise ^h, Università di Trieste
- Giuseppe Della Ricca ^h, Università di Trieste
- Valentina Zaccolo ^k, Università di Trieste
- Mattia Faggin ^k, Università di Trieste
- Lorenzo Rinaldi ^e, Università di Bologna
- Piergiulio Lenzi ^g, Università di Firenze
- Vitaliano Ciulli ^g, Università di Firenze
- Sharam Rahatlou ^h, Università Roma 1
- Daniele del Re ^h, Università Roma 1
- Lorenzo Capriotti ^f, Università di Ferrara
- Francesco Conventi ^e, Università di Napoli
- Francesco Ciotto ^e, Università di Napoli

PhD students:

- Francesco Vaselli ^c, Scuola Normale Superiore di Pisa
- Matteo Barbetti ^b, Università di Firenze
- Muhammad Numan Anwar ^j, Politecnico di Bari
- Benedetta Camaiani ^g, Università di Firenze
- Alkis Papanastassiou ^g, Università di Firenze
- Antonio D'Avanzo ^e, Università di Napoli

External collaborators:

- Andrea Rizzi ^c, Università di Pisa

KPIs

KPI ID	Description	Acceptance threshold	2024-09-24
KPI2.2.1.1	N_{MC} billion events obtained from ML-based simulation, as demonstrated by official links in experiments' simulation databases	$N_{MC} \geq 1$	2.3 M events (completed: 0.2%)
KPI2.2.1.2	N_{EXP} experiments have tested a machine-learning based simulation	$N_{EXP} \geq 2$	3 experiment (completed: 150%)
KPI2.2.1.3	Machine-learning use-cases tested in the context of the CN were presented at N_{CONF} international and national events	$N_{CONF} \geq 3$	17 use-cases (since Sept. '23) (completed: 567%)
KPI2.2.1.4	N_{UC} different machine-learning use-cases were tested in the context of the CN and made available in git repositories	$N_{UC} \geq 5$	5 use-cases (completed: 100%)



Risk Analysis

Identifier	Description	Update
R1	The CN is unable to provide the needed resources	We have access to Leonardo resources, and the provisioning model enabling offloading via InterLink has been validated in Integration PoC. Offloading from the AI_INFN Platform is still being commissioned. We have recently gained access to Tier1 resources. HPC Bubbles should be installed in 2024 Q4 and the provisioning model is under discussion.
R2 TODAY	The provisioning model is not ready for production	The provisioning model has been validated on CPU and using simplified backends such as Docker or Kubernetes itself, and recently extended to HTCondor backend. Polishing and refinements are needed to enable offloading to CINECA Leonardo.
R3	The recruitment process has limited or delayed success due to the large number of ML positions opening	All postdoc selected to work on the subject of this flagship planned to start on October 2024 withdrew. A new selection is ongoing, with planned start January 2025.

Flash Simulation

From M. Barbetti's talk @ ICHEP 2024 [\[LINK\]](#)

Fast simulation vs. Flash simulation

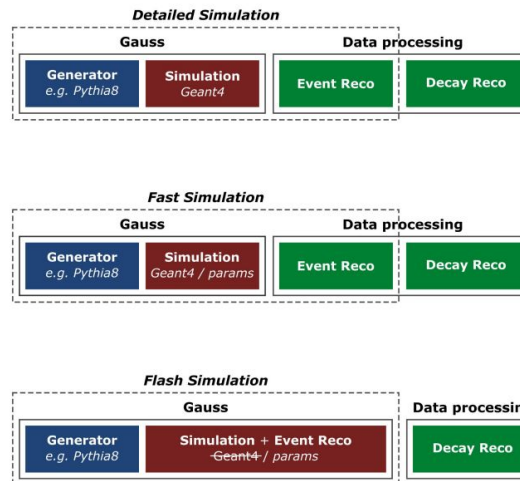
Methods to **speed up** the Geant4-based simulation productions:

- upgrade of the simulation framework (including multi-threading)
- leveraging GPU-acceleration (e.g., use AdEPT, Celeritas)
- reuse of the not-signal part of the event, **ReDecay** [2]

Fast Simulation techniques to parameterize the detector low-level response without relying on Geant4:

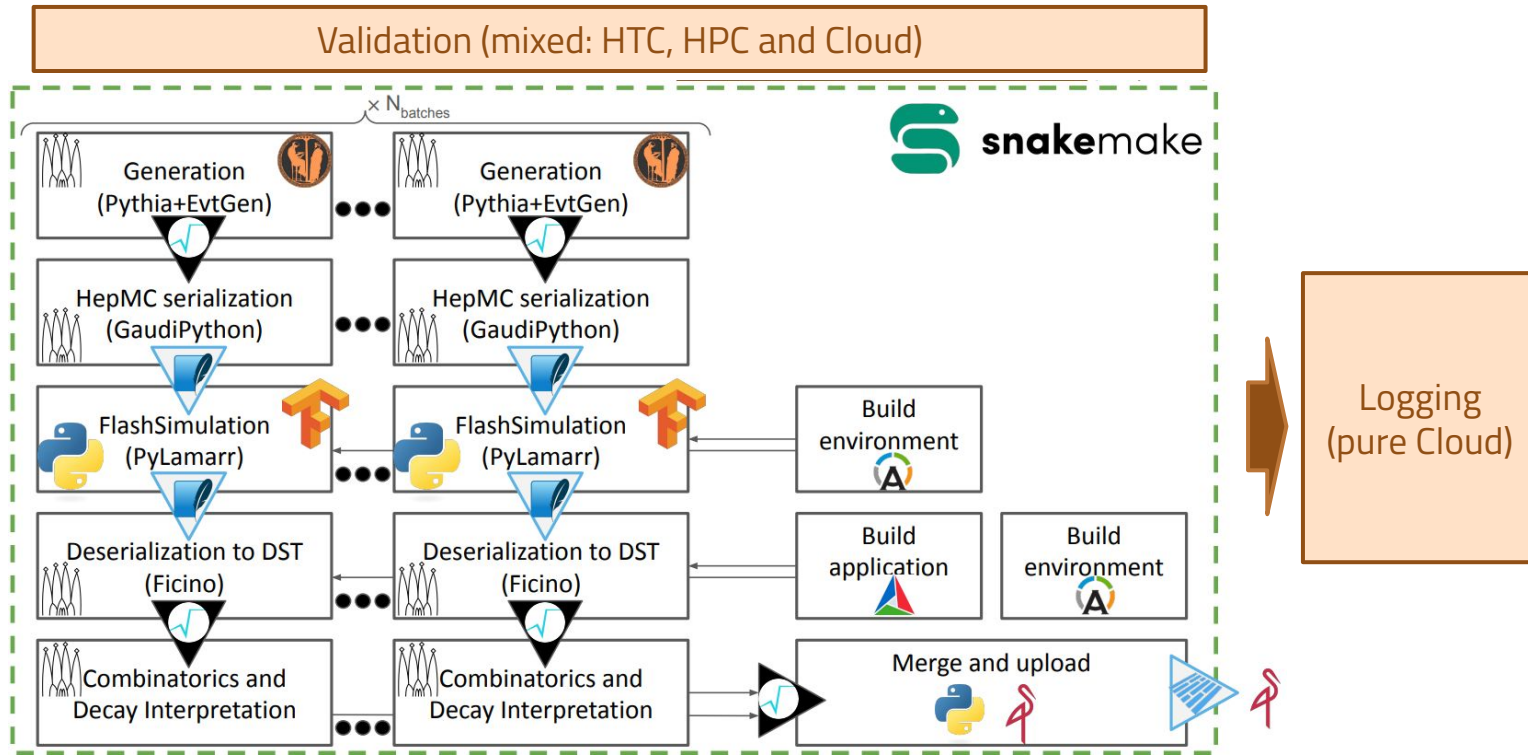
- **Point library** for Calorimeters energy deposits [3]
- **Generative Models** (e.g., GAN, VAE) for Calorimeters energy deposits [4]

Flash Simulation (also called *Ultra-Fast* or *parametric*) defines a more radical approach by replacing Geant4 and reconstruction with parameterizations able to **directly transform** generator-level particles into analysis-level reconstructed objects



3

Flash Simulation Workflow



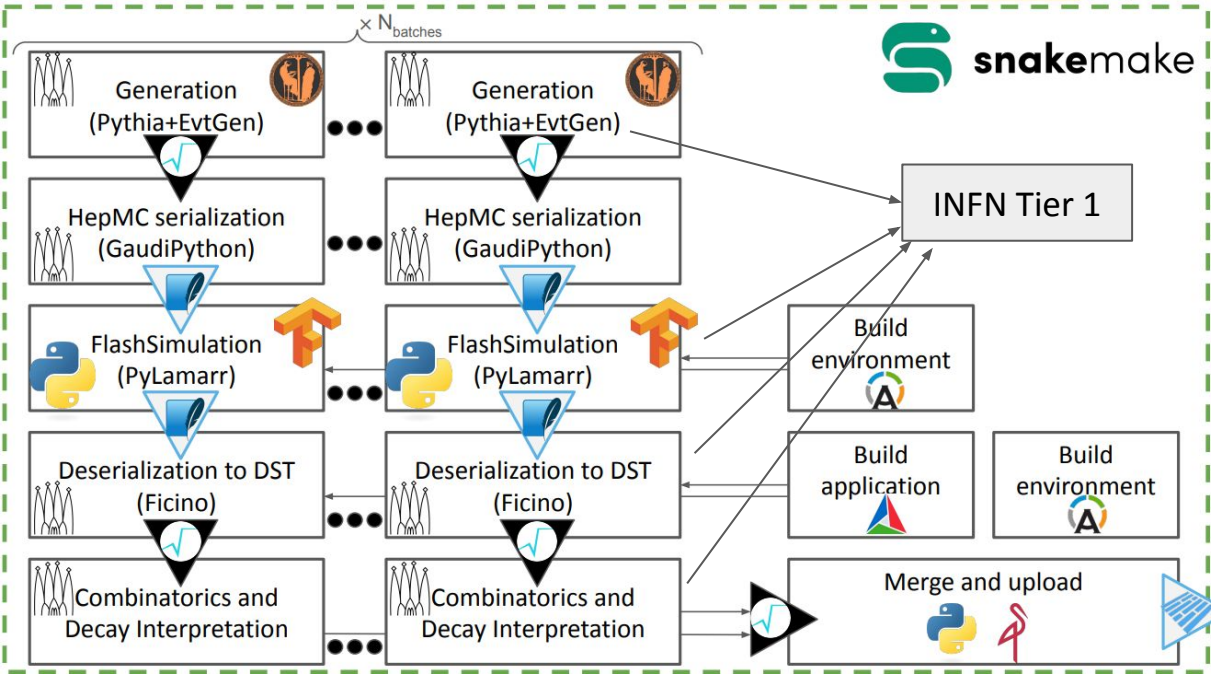
Flash Simulation Workflow – offloaded

Training
(pure HPC)



Leonardo

Validation (mixed: HTC, HPC and Cloud)



Logging
(pure Cloud)



Status of the integration of INFN-T1 resources

Developing the **HERD Computing Model**, CNAF defined a CondorCE submitting jobs from remote locations through authentication.

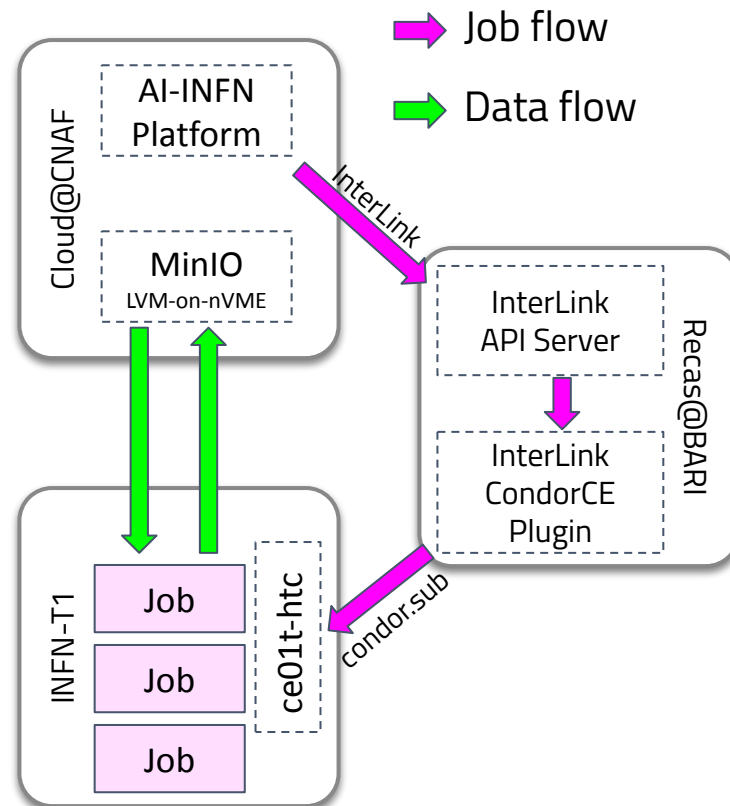
Unfortunately, the CondorCE is not reachable from Cloud@CNAF for network policies, but it is, for example, from ReCaS@BARI.

We developed an **InterLink plugin** sitting in a VM in Bari, accepting InterLink submissions from Cloud@CNAF and forwarding them to CNAF Tier-1 test CE.

The plugin converts the **Kubernetes Pod** specifications into a (possibly rather long) shell script running **Apptainer** containers in multiple subprocesses.

Input and output data is managed through a self-managed **MinIO instance on LVM-on-nVME** hosted in **Cloud@CNAF**.

 **landerlini/interlink-condorce-plugin**



(re)Defining cvmfs and fuse volumes

Converting Pod's requests to access cvmfs or fuse data should be responsibility of the plugin, as different compute backend may be subject to different rules.

In CondorCE plugin I use generic annotations to define volumes.

For Leonardo, this require hacking the singularity submission command (very verbose).

```

apiVersion: v1
kind: Pod
metadata:
  name: cern-vm-fs
  annotations:
    cvmfs.vk.io/my-volume: sft.cern.ch
spec:
  containers:
  - name: main
    image: ubuntu:latest
    command:
      - /bin/bash
      - -c
      - ls /
    volumeMounts:
      - name: my-volume
        mountPath: /cvmfs
        readOnly: True
  volumes:
  - name: my-volume
    persistentVolumeClaim:
      claimName: intentionally-not-existing
  
```

```

apiVersion: v1
kind: Pod
metadata:
  name: fuse-vol
  annotations:
    fuse.vk.io/my-fuse-vol: |
      cat << EOS > /tmp/rc1one.conf
      [example]
      type = local
      EOS

      # Mimic a remote
      mkdir -p /tmp
      echo "hello world" > /tmp/file.txt

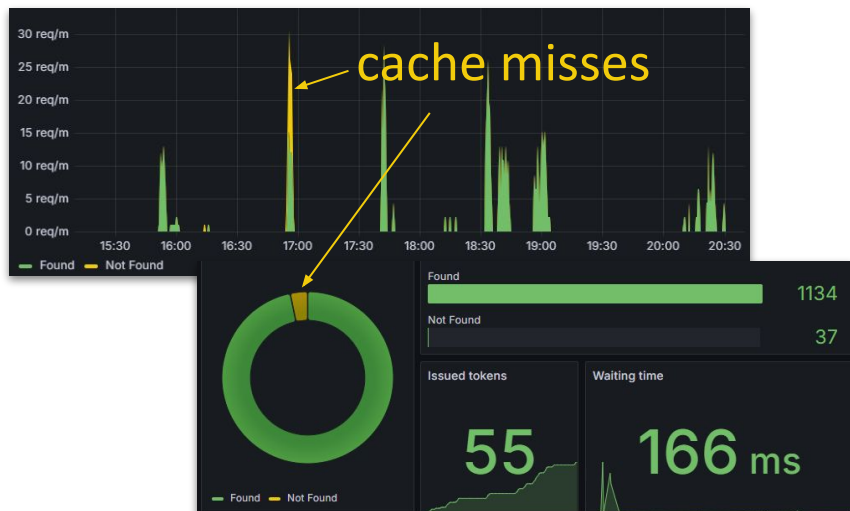
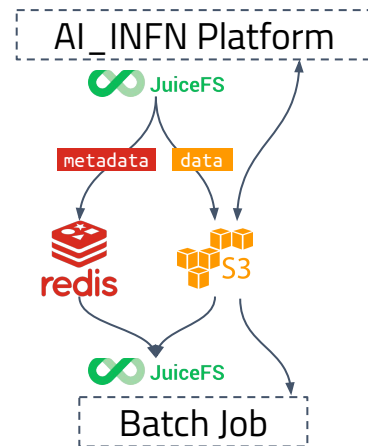
      # Mount the remote
      rclone mount2 \
        --config /tmp/rc1one.conf \
        --allow-non-empty example:/tmp \
        $MOUNT_POINT
spec:
  containers:
  - name: main
    image: rclone/rclone:latest
    command:
      - cat
    args:
      - /mnt/fuse-vol/file.txt
    volumeMounts:
      - name: my-fuse-vol
        mountPath: /mnt/fuse-vol
  volumes:
  - name: my-fuse-vol
    persistentVolumeClaim: # deliberately fake pvc
      claimName: csi.example.com
  
```

Solving the distributed cache problem

Focus on data flow

A **shared virtual file system** is mounted by the condor nodes with fuse using JuiceFS.

JuiceFS falls back on **MinIO** for the data and **Redis** (part of the AI-INFN platform) for the metadata



[ShubProxy]

Downloading and building docker images into SIF for each jobs

- would cause a periodic bans of CNAF by DockerHub;
- cause large inefficiency in short jobs.

We deployed a simple web application defining a shared cache.

If the image is not available in S3, the web app schedule its build, otherwise it return the built artifact from cache.



Status of the integration with Leonardo

The slurm plugin in production in Leonardo, does not accept Pod requests from the Flash Simulation workflow.

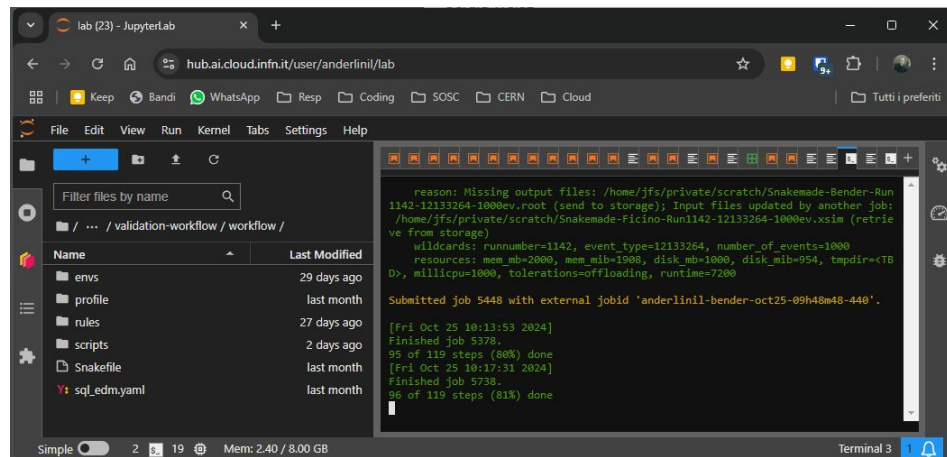
All the building blocks were tested separately and we expect no fundamental reason for the plugin not to work.

Still, some polishing would be needed, probably in a joint debugging session.

Alternatively, we may try to use the CondorCE plugin submitting to slurm.

Combining CNAF Tier-1 and CINECA Leonardo resources

- Kubernetes cluster
 - AI_INFN Platform (RKE2 with Kubernetes 1.27)
 - Virtual Nodes installed manually (no helm chart)
 - Snakemake as workload manager
- Tier1 setup
 - CondorCE (originally developed for HERD) mapped to ce01t
 - InterLink server and dedicated plugin running in a VM in ReCaS
- Leonardo setup
 - Slurm submission from edge node icsc01
 - Official interlink slurm plugin



```

reason: Missing output files: /home/jfs/private/scratch/Snakemake-Bender-Run
1142-12133264-1000ev.root (send to storage); Input files updated by another job:
/home/jfs/private/scratch/Snakemake-Ficino-Run1142-12133264-1000ev.xs.in (retrie
ve from storage)
wildcards: runnumber=1142, event_type=12133264, number_of_events=1000
resources: mem_mb=2000, mem_mib=1908, disk_mb=1000, disk_mib=954, tmpdir=<TB
D>, millicpu=1000, tolerations=offloading, runtime=7200

Submitted job 5448 with external jobid 'anderlini1-bender-oct25-09H40m48-440'.

[Fri Oct 25 10:13:53 2024]
Finished job 5378.
95 of 119 steps (80%) done
[Fri Oct 25 10:17:31 2024]
Finished job 5738.
96 of 119 steps (81%) done
  
```

```

(miniconda3)[Lucio@pchlcb06 v3]$ k get nodes
NAME                                STATUS    ROLES
hub-a100-2                          Ready    <none>
hub-a100-3                          Ready    <none>
hub-a102-b                          Ready    <none>
hub-cpu-2                          Ready    <none>
hub-master                          Ready    control-plane,etcd,master
hub-rtx-2                          Ready    <none>
hub-rtx-3                          Ready    <none>
hub-storage                         Ready    <none>
vk infn-t1                          Ready    agent
vk leonardo-virtual-node            Ready    agent
  
```

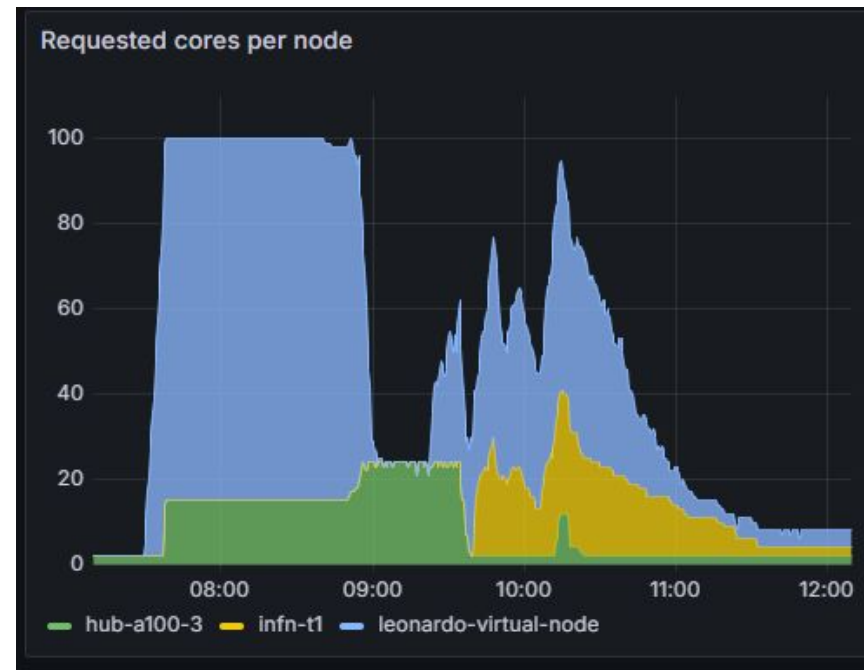
First workflow combining Leonardo DCGP and Tier-1

On October 25th, we run a first workflow combining CPU resources from Tier-1, Leonardo and a local node.

Will perform scalability tests soon.

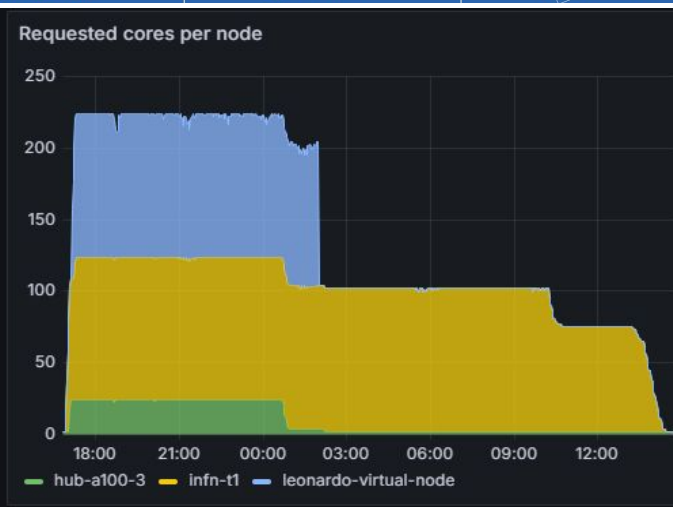
Known scalability boundary is the size of the allocated buffer (1 TB), $n\text{CPU} < 1\text{k}$.

Offloading to Leonardo booster (with GPU payloads) coming soon.



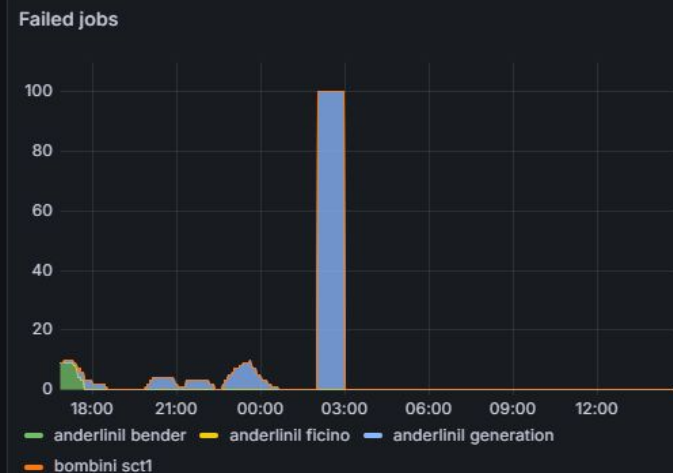
Longer run

Observed a bottleneck in the submission system, we tried submitting a bulk of “long” jobs.



Without entering details, very interesting dynamics, due to the interplay of:

- Node limits
- Kueue Resource Flavor limits
- Backend priority policies



AI_INF N Machine Learning Hackathon with ICSC support

AI_INF N is organizing the AI_INF N hackathon next week covering code-examples for the topics:

- Flash simulation and unfolding with GANs
- Reconstructing experimental data (LHCf)
- Processing of time-dependent NMR images
- Quantum Machine Learning

When: 26 – 28 November 2024

Where: Padova

Link to the agenda: agenda.infn.it/event/43129/

1st AI-INF N Advanced Hackathon

26-28 Nov 2024
University of Padua, Complesso Paolotti
Europe/Rome timezone

Public Europe/Rome L. Anderlini

Enter your search term

Overview

- Timetable [Preliminary]
- Registration
- Experts and tutors
- Access to cloud resources

Contact

✉ ml-infn-hackathons@ist...

Welcome to the First edition of the Advanced Artificial Intelligence @ INF N (ALINF N) hackathon, dedicated to INF N affiliates. This edition is hosted at INF N Sezione di Padova.

Notably, it is the third Hackathon to happen in Person, so please apply only if you are planning to come to Padua. The logistics allow for ~ 20 participants.

ALINF N hackathons are developed in continuity with ML_INF N hackathons. You may want to check the indicio pages of the [first](#) (entry level), [second](#) (entry level), [third](#) (advanced level), [fourth](#) (entry level) and [fifth](#) (advanced level) editions of ML_INF N hackathons, with most of the talks attached as video files.

The [mandatory registration](#) process will be open soon.

In case of a number of registrations exceeding the available positions, the applications will be ranked and selected on the basis of the scientific CV of the applicants and of the order of registration.

The successful applicant will be informed by November 10th. **Please do not book hotel/flight before a positive confirmation.**

The course is to be considered as "advanced level" for Machine Learning topics. The hackathon will be organized over 3 days, distributed as

School of Open Science Cloud

Deep involvement of WP2 people also in the organization of SOSC, including advanced machine learning use-cases of a cloud-based infrastructure.

- Containerization
- Data management
- Computer Vision and Machine Learning
- Distributing workflows

When: 2 – 6 December 2024

Where: Bologna

Link to the agenda: agenda.infn.it/event/40829

SOSC 2024 Sixth International School on Open Science Cloud

2-6 Dec 2024

University of Bologna. Department of Physics and Astronomy
Europe/Rome timezone

Overview

Timetable

Contribution List

Registration

Travel & Accomodation

School contact - mail

sosc24-pc@lists.infn.it

The theme of the sixth International School on Open Science Cloud is "Computing Models for Scientific Experiments"

The SOSC24 is organized by



ALMA MATER STUDIORUM
UNIVERSITÀ DI BOLOGNA



Centro Nazionale di Ricerca in HPC,
Big Data and Quantum Computing



UNIVERSITÀ DEGLI STUDI
DI PERUGIA



The 6th edition of the **International School on Open Science Cloud (SOSC 2024)** will be held in Bologna, from **02 to 06 December 2024**. The school is organized by INFN, Department of Physics and Astronomy "Augusto Righi" of the University of Bologna, the Departments of Physics and Geology of the University of Perugia and the ICSC Foundation.

The School is multi-disciplinary and targeted at postgraduate researchers including bachelor degree or equivalent in fields such as physics, statistics, computer science, computer vision, biology, medicine, bioinformatics, engineering, working at any research institute, with some experience and interest in data analysis, in computing or in related fields. Applications by university students (undergraduate) will be considered depending on availability and must be accompanied by a letter of reference from a university professor. **We embrace diversity and strongly encourage qualified and curious individuals from all nationalities and backgrounds to apply.**

Important dates

- **Monday 3rd of June** - applications open
- **Thursday 5th of September** – acceptance notification sent to the participants
- **Saturday 5th of October** - application closes
- Application confirmation will be sent to participants by **October 20**
- **Friday 1st of November** - registration fee payment deadline
- **Monday 2nd December** - student arrivals at Bologna
- **Friday 6th of December** - departure

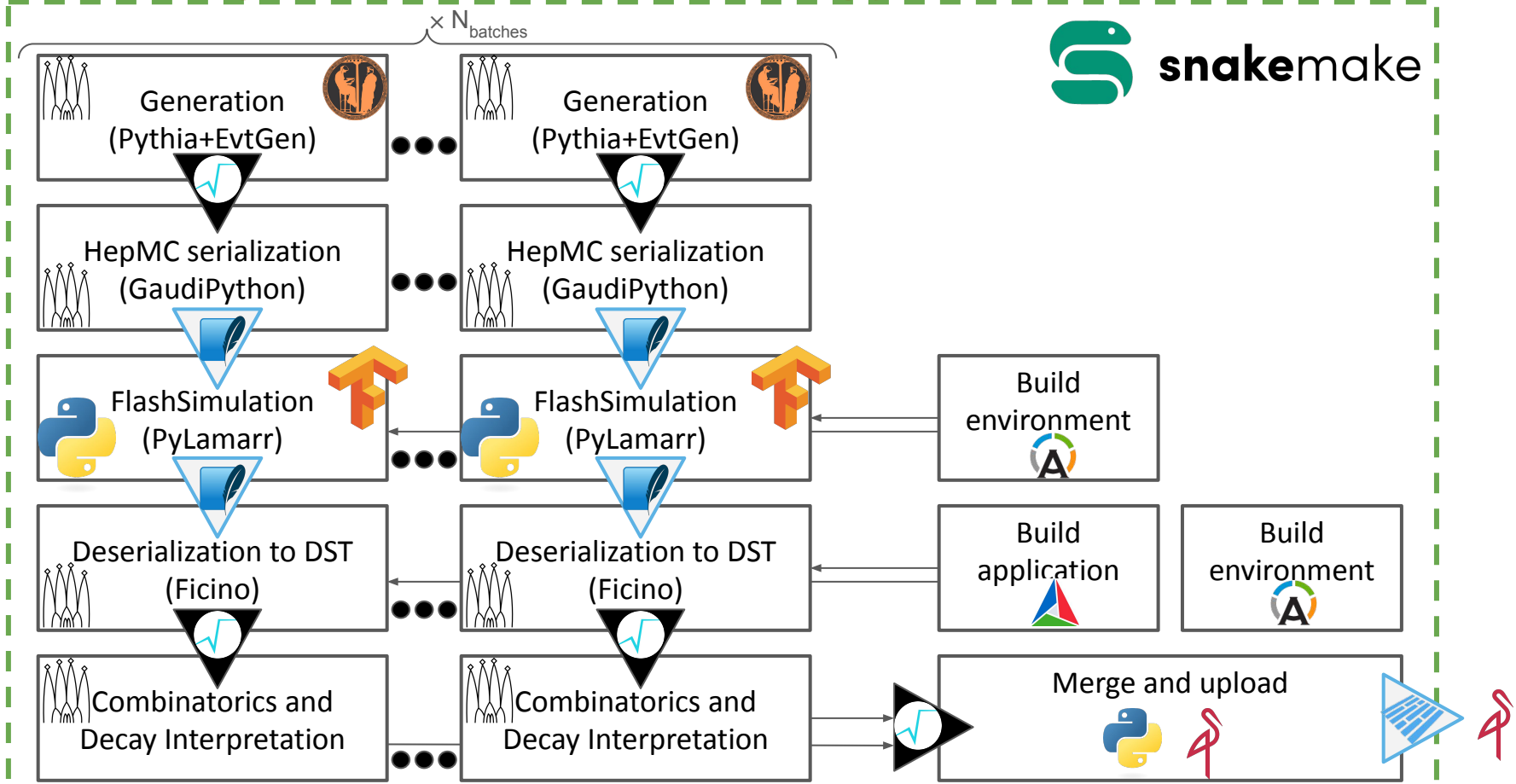
The SOSC 2024 is also supported by the INFN Commissione Scientifica Nazionale 5 (CSN5) through the initiative "ALINFN"





Backup

Recalling the production strategy



Resources

Pythia8 (full event)

Generates the whole proton-proton collision event, with pileup and spill-over.
Then processes all particles with Lamarr and Bender to produce nTuples.

1M events (on 50 parallel jobs) require:

- $O(48h) \times 50$ CPUs
- 0.8 TB of buffer in S3.



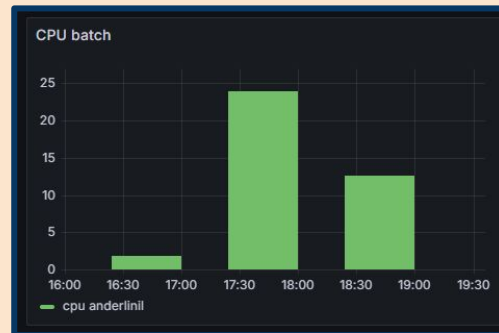
Particle Gun (signal-only)

Generates only the heavy hadron decay.
Then processes particles with Lamarr and Bender to produce nTuples.

Less tested than Pythia8 productions

1M events (on **up to** 50 parallel jobs) require:

- $O(1h)$, *limited by submission latency*
- 4 GB of buffer in S3



Requests for the validation part

Resource	Full Request	Strictly required for KPI 1 (Full-Pythia option)
CPU on INFN Cloud	2 M CPU hours	2.4 M CPU hours*
GPU on INFN Cloud	4 H200 for 18 months	0
GPU on Leonardo Booster via InterLink	10000 hours	0
Storage	25 TB	10 TB

Preliminary

- 0.5 M hours from opportunistic borrowing from AI_INFN Platform