Finanziato dall'Unione europea NextGenerationEU

Ministero dell'Università e della Ricerca

Italiadomani PIANO NAZIONALE DI RIPRESA E RESILIENZA

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

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

Benchmark interactive analyses ongoing at INFN Napoli

Adelina D'Onofrio, Elvira Rossi, Francesco Cirotto, Francesco Conventi, Orso Iorio, Antimo Cagnotta, Antonio D'Avanzo, Gianluca Sabella, Bernardino Spisso, Francesco G. Gravili

Spoke 2 - WP2 working meeting, 21st May 2024

Outline

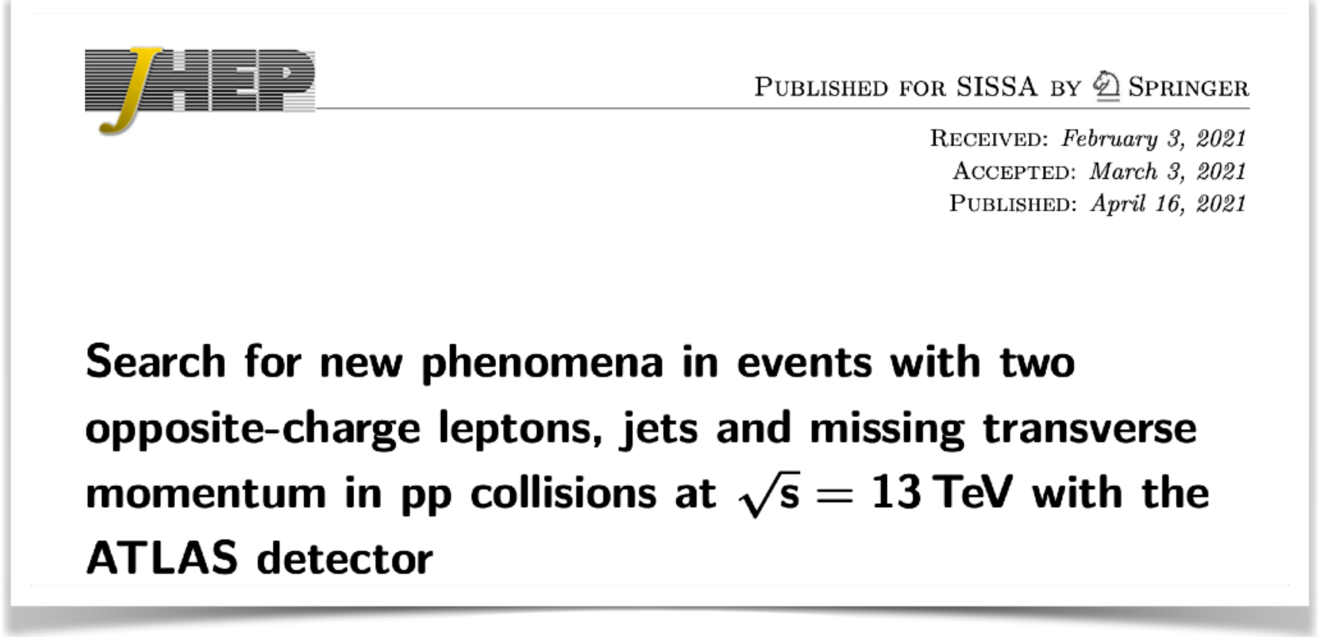
- Motivations
- Use cases tested:
  - ATLAS: stop to 4-body SUSY analysis
    - ➡ collaboration with INFN Lecce
  - CMS: top quark+MET analysis
    - ➡ collaboration with INFN Perugia

Motivations

- Most of the LHC searches/measurements rely on locally developed scripts that process the datasets, with parallel tasks and on an asynchronous batch system
- **Challenges of HL-LHC and future** colliders are pushing to **re-think the HEP computing models**
  - Impact on several aspects, from software to the computing infrastructure
- From the software perspective, **interactive/quasi interactive analysis** is a promising paradigm
  - User-friendly environment
  - The implementation is simplified by adopting open-source industry standards: *Dask, Jupyter Notebooks* and *HTCondor*
  - Validating new frameworks (e.g. *ROOT RDataFrame* with multi-threading)
- **Preliminary feasibility studies** have been pursued exploiting **INFN Napoli** high rate platform
  - Distributed infrastructure which leverages *Dask*

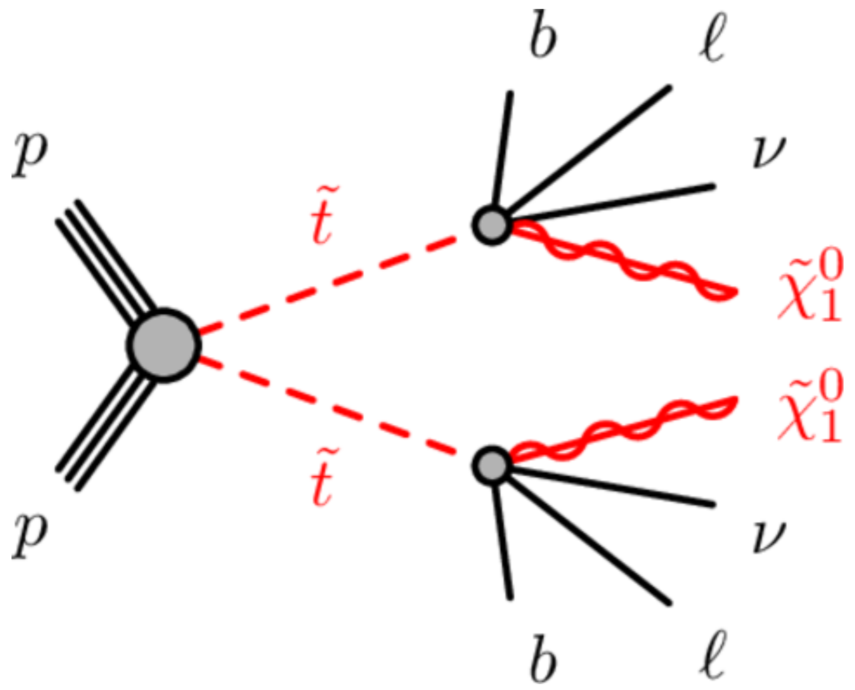
ATLAS use-case

SUperSYmmetry: Beyond Standard Model theory


JHEP

PUBLISHED FOR SISSA BY SPRINGER
RECEIVED: February 3, 2021
ACCEPTED: March 3, 2021
PUBLISHED: April 16, 2021

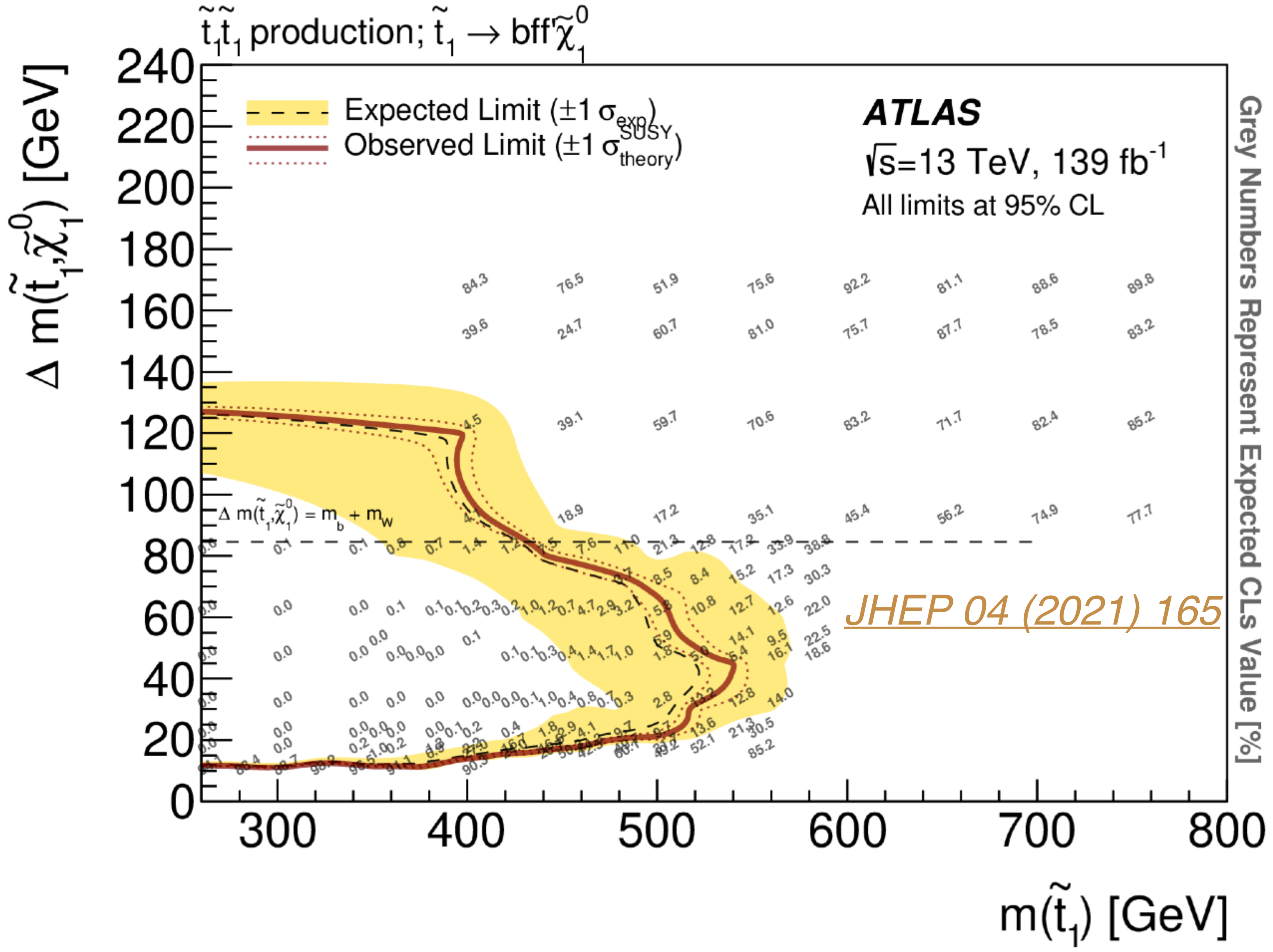
Search for new phenomena in events with two opposite-charge leptons, jets and missing transverse momentum in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector



Soft leptons coming from a virtual W* boson decay

Compressed mass spectra:
$\Delta m < m_W + m_b$

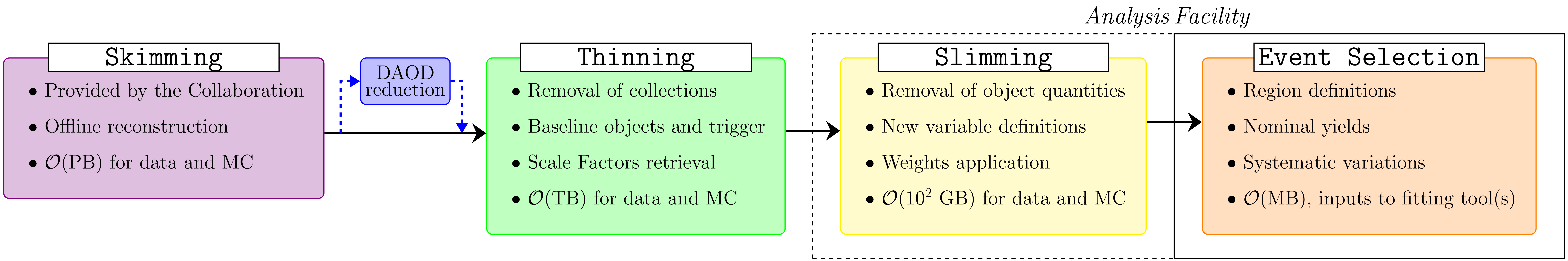
- Three different analysis in the ***Run 2 paper***, already published, according to mass splitting between *stop* ($\tilde{t}_1$) and *neutralino* ($\tilde{\chi}^0_1$), allowing different decay modes:
  - 2 body → $\Delta m > m_t$
  - 3 body → $m_W + m_b < \Delta m < m_t$
  - 4 body, the one picked up → $\Delta m < m_W + m_b$
- Common final state signature: 2 OS leptons (electrons/muons), jets and missing transverse energy
- Cut & Count based approach



JHEP 04 (2021) 165

Working with INFN Lecce

4-body search workflow



Slimming

ATLAS slimming code already in RDataFrame, but entirely written and compiled in C++ —> NO dask distributed approach

Event Selection

- Event selection for fitting tools
- RDataFrame + Dask applied to Wt bkg sample ~ 1.8 GB copied to the INFN Na workspace
- Playing with syst. variations
- Code ready to play with other backgrounds

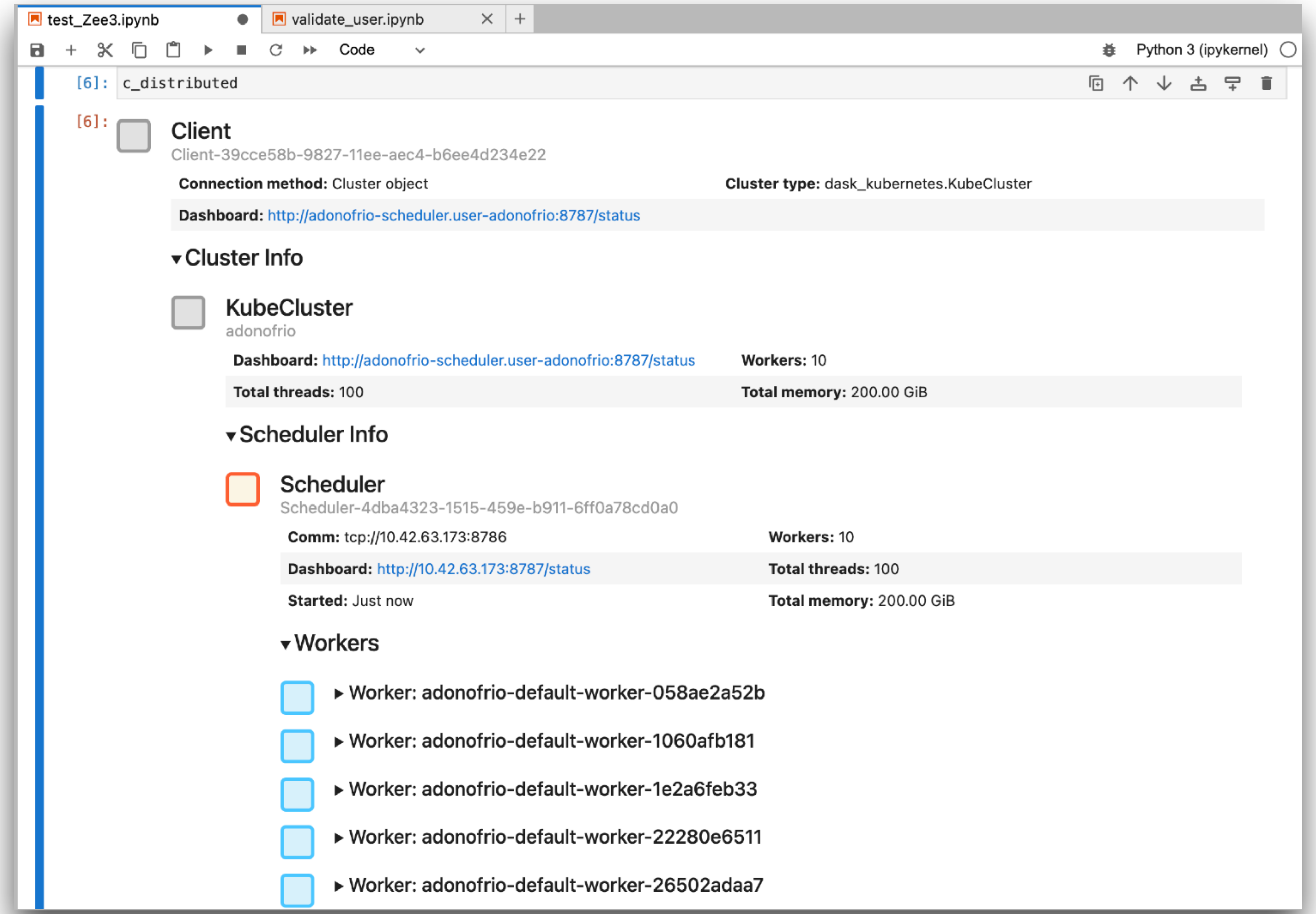
dask

jupyter

Infrastructure and kubecluster

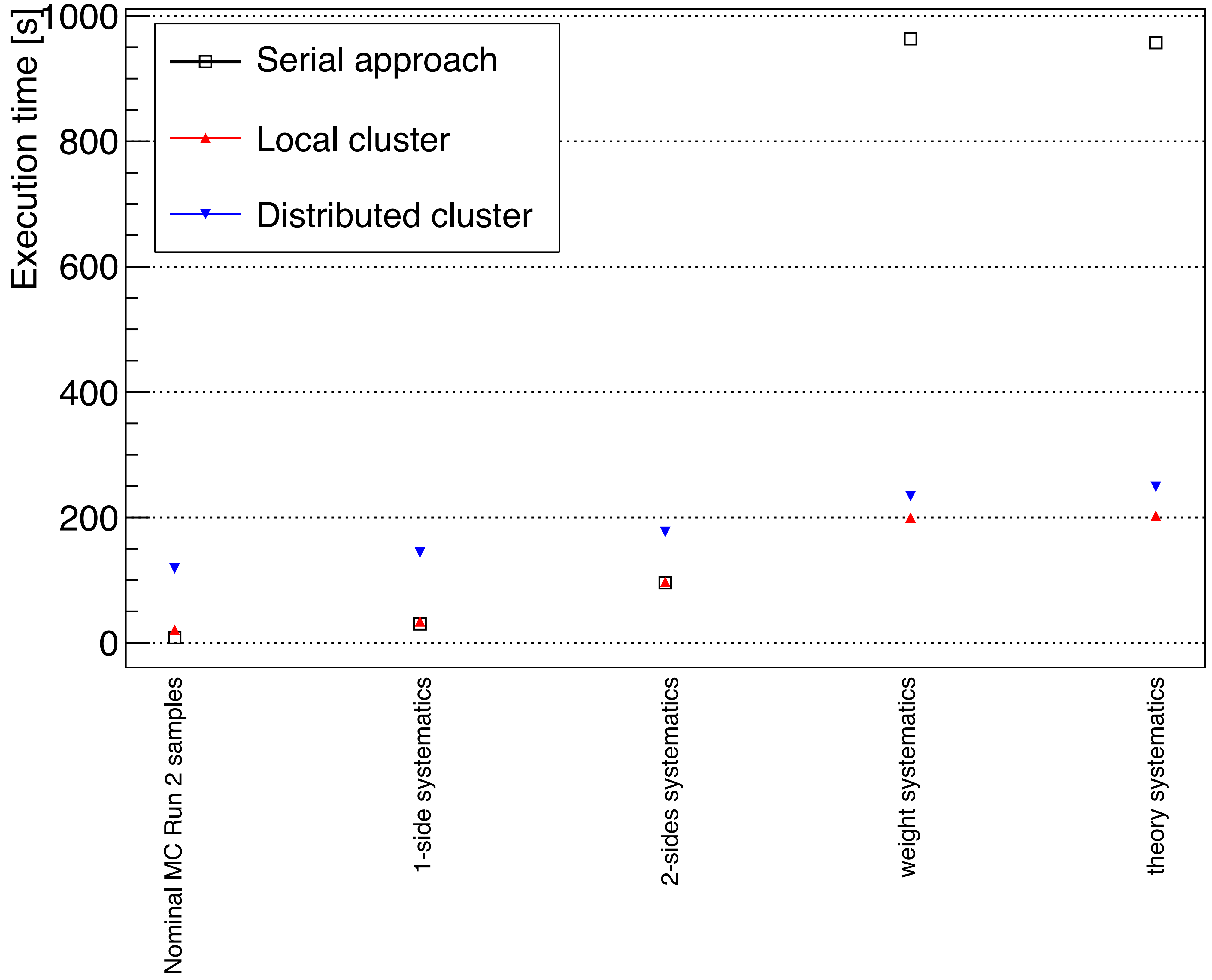
- The local (INFN Naples) deployment is based on the *Open-Stack IaaS* paradigm
- Starting from the already existing *I.Bi.S.CO* installation, several updates were performed
- The physical cluster is made up of 2 identical virtual machines, each equipped with 1CPU quadCore and 8GB RAM, currently expanded up to 12 cores and 64GB
- Rocky Linux 8.6 is the operating system
- 2 nodes are equipped with **Docker** (20.10) for containerisation and **Kubernetes** (1.26.3) for the orchestration
  - One node plays as controlplane, etcf & worker; the other node acts as a plain worker
- The cluster is equipped with **JupyterHub** & **JupyterLAB** where the user can play with **Python, ROOT & Dask** libraries

- Kubernetes infrastructure: 5+1 virtual machines
- 32 cores & 64 GiB in each compute
- If more users are connected, scheduling jobs priority
- Clusters are dynamic objects

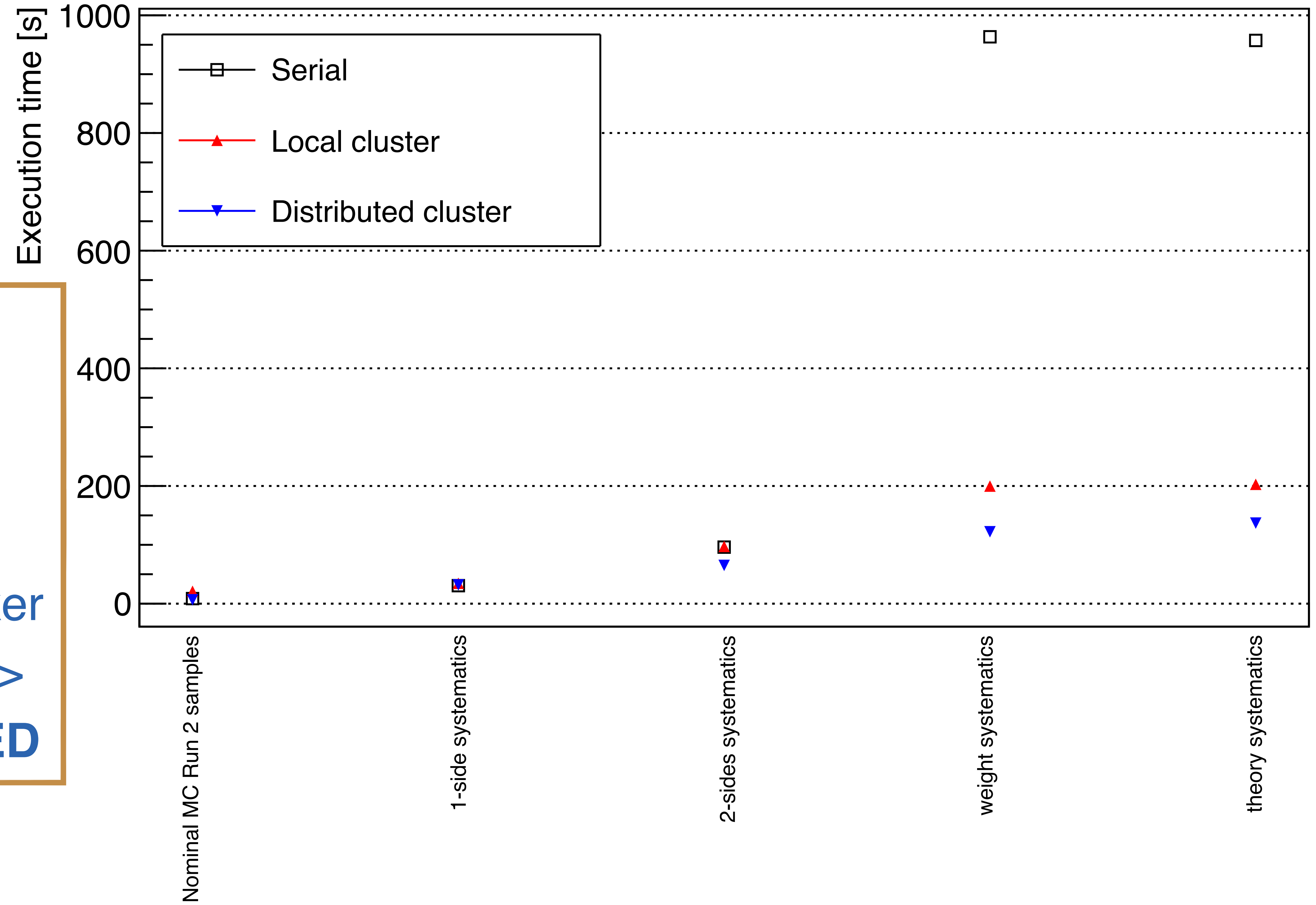


Preliminary results

Defined Metric	
Overall execution time	Time elapsed from the start of the execution (execution triggered) to the end of execution



Limiting factor for the distributed approach: time spent to copy the inputs to the worker nodes (~120 s)—> **TO BE IMPROVED**



- Exploiting the distributed approach, the execution time improves *wrt* the standard/serial approach if we iterate over a significative number of systematic variations (each step in the x-axis includes previous contributions)

Scheduler and Working Nodes Reports

Distributed approach

Dask Performance Report

Select different tabs on the top for additional information

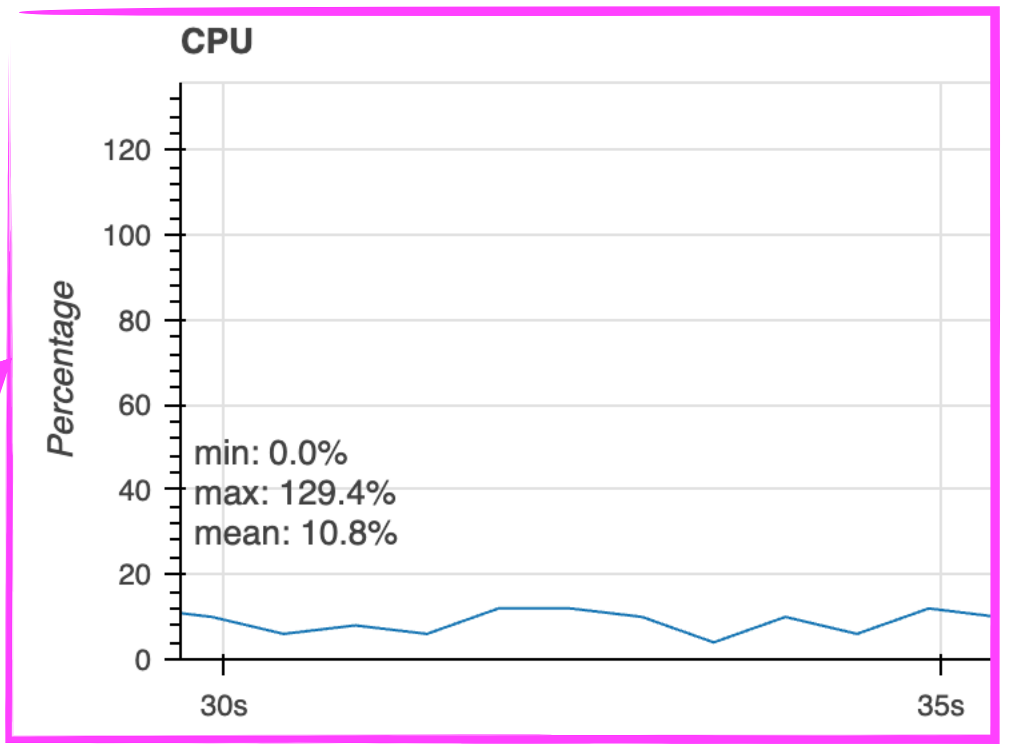
Duration: 252.87 s

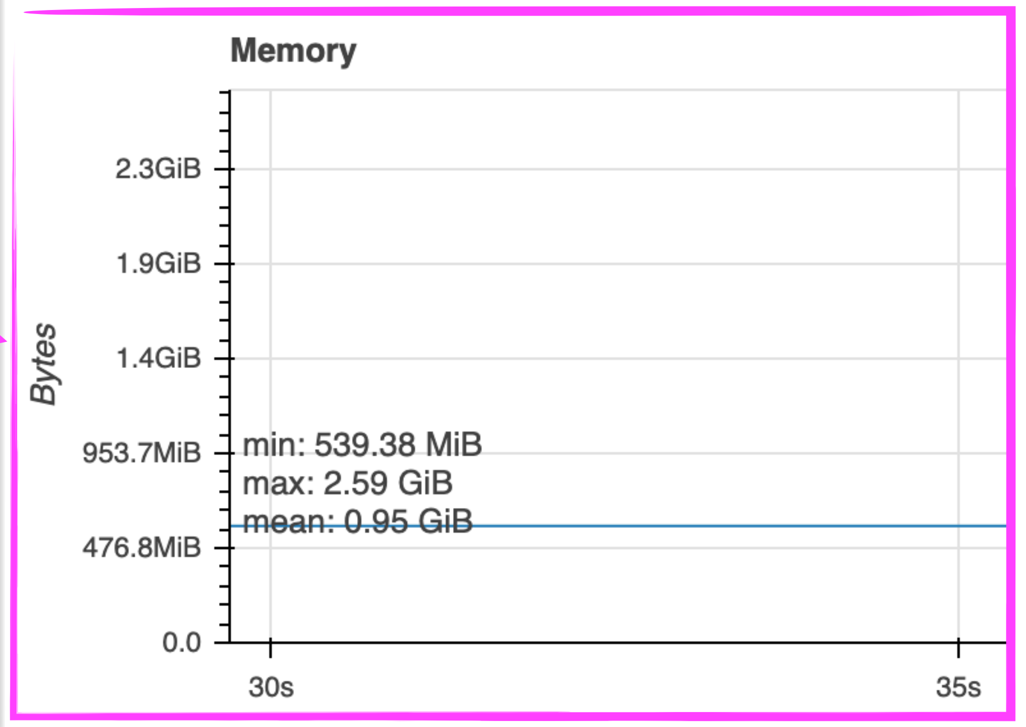
Tasks Information

- number of tasks: 621
- compute time: 118.06 s
- deserialize time: 2.39 s

Scheduler Information

- Address: tcp://127.0.0.1:43821
- Workers: 2
- Threads: 2
- Memory: 4.39 GiB
- Dask Version: 2022.11.0
- Dask.Distributed Version: 2022.11.0





/home/jovyan/work/pnrr-atlas-analysis/dask-report.html

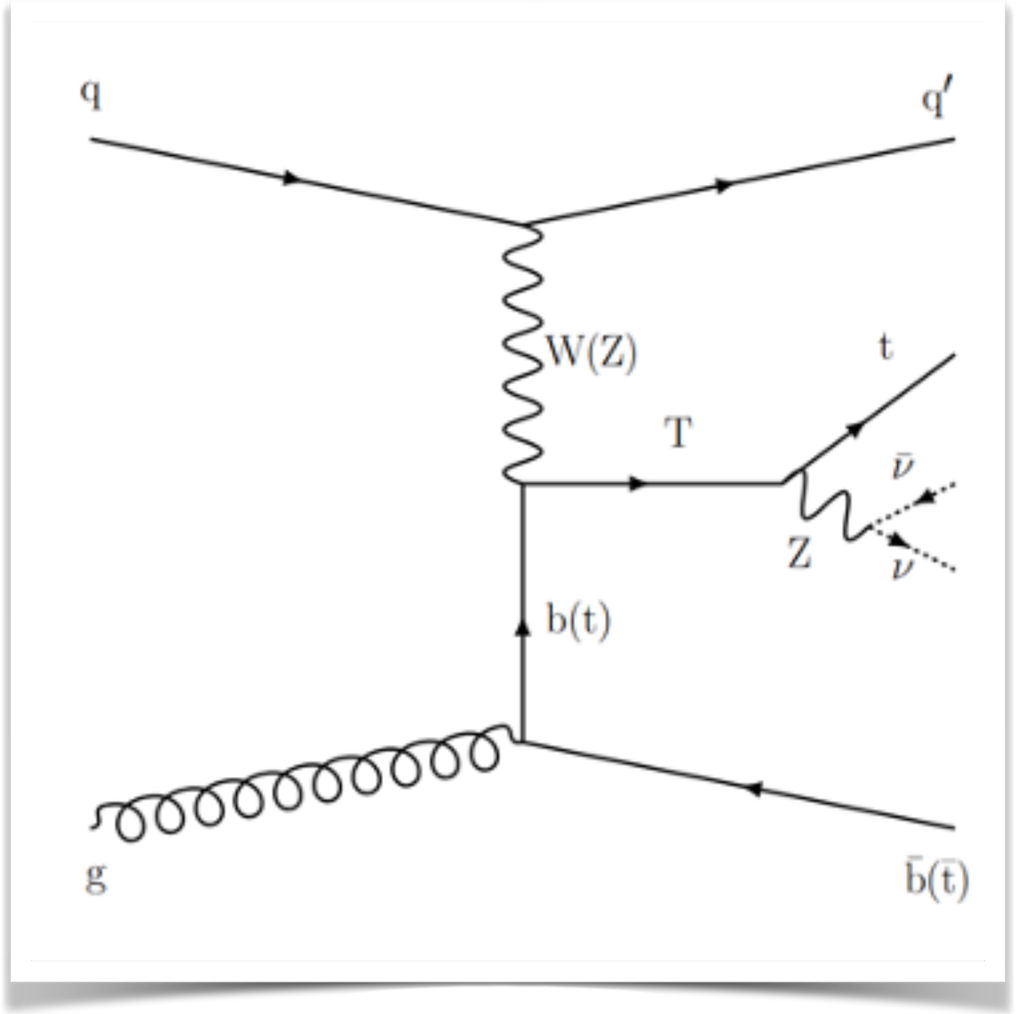
Connecting to working nodes

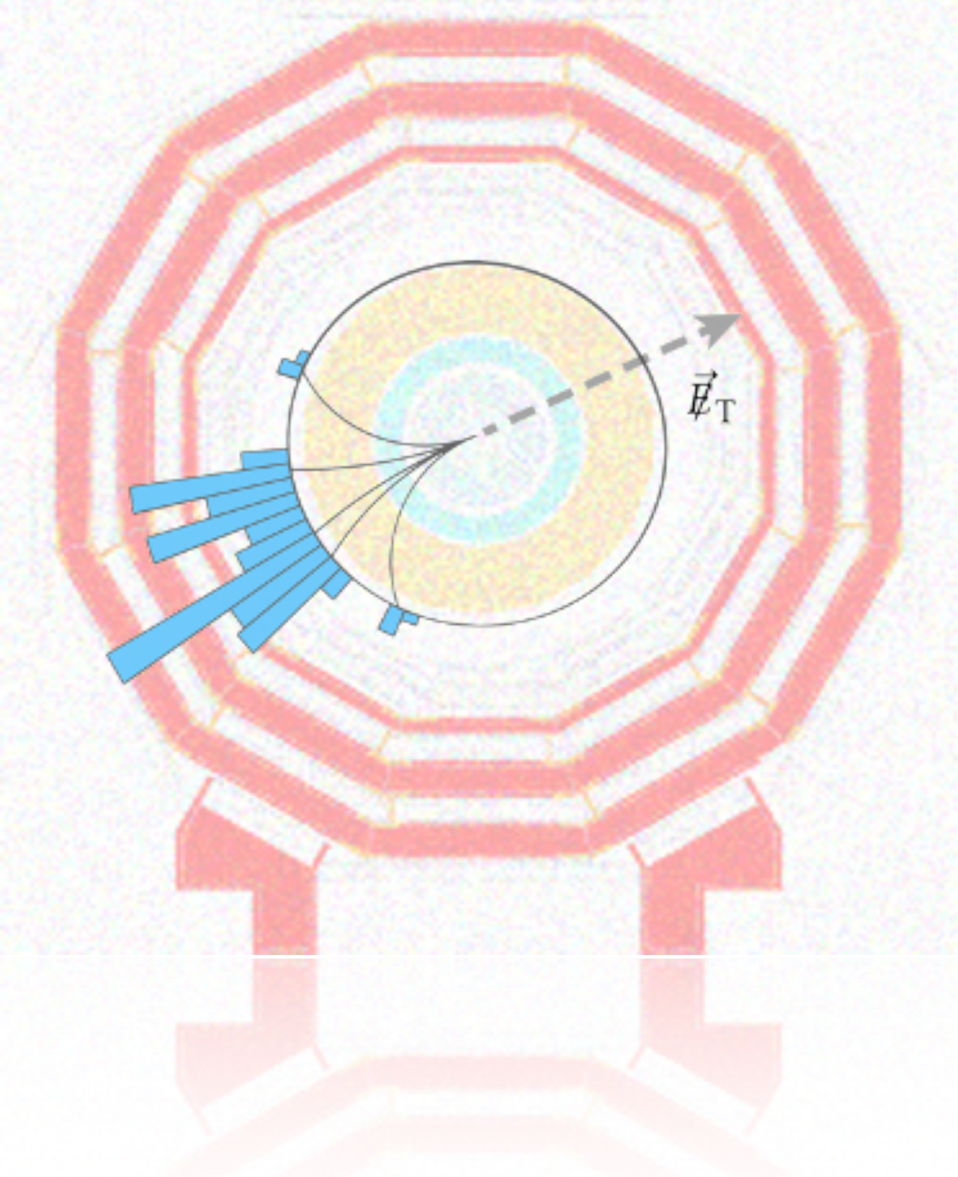
- Out of 9 worker nodes, we get up to 91% CPU occupancy on one node and up to 5% on the other nodes

```
[almalinux@kuber-node-2 ~]$ top

top - 10:47:40 up 17 days, 22:40,  1 user,  load average: 0.47, 0.28, 0.20
Tasks: 504 total,   1 running, 503 sleeping,   0 stopped,   0 zombie
%Cpu(s):  3.1 us,  0.4 sy,  0.0 ni, 96.2 id,  0.2 wa,  0.1 hi,  0.0 si,  0.0 st
MiB Mem :  64298.3 total,  52125.8 free,   6999.9 used,   5863.5 buff/cache
MiB Swap:      0.0 total,      0.0 free,      0.0 used.  57298.4 avail Mem

    PID USER      PR  NI    VIRT    RES    SHR S  %CPU  %MEM     TIME+ COMMAND
3467064 almalin+  20   0 1077988 627812 378248 S  91.4   1.0   0:05.15 python
3436893 almalin+  20   0 6841396 553904 170460 S   3.0   0.8   0:15.16 python
2163103 almalin+  20   0  544308 118628  29288 S   2.3   0.2  94:53.01 python3.10
1244708 almalin+  20   0 6485488 390772  27588 S   2.0   0.6 337:41.39 python
   1574 root      20   0 4011788 148624  72036 S   1.7   0.2 755:31.40 kubelet
1342269 almalin+  20   0  544320 116988  29364 S   1.7   0.2 250:13.90 python3.10
1245200 almalin+  20   0 1165092 378416  19604 S   1.3   0.6 279:15.23 python
3462166 almalin+  20   0 2453528   1.9g  29908 S   1.3   3.1   0:13.91 python3.10
3467062 almalin+  20   0  438548  87068  27772 S   1.3   0.1   0:02.15 python
1245202 almalin+  20   0 1189084 437148  19584 S   1.0   0.7 260:06.77 python
3461916 almalin+  20   0 4722168  61600  19928 S   1.0   0.1   0:06.10 dask-worker
   1173 root      20   0 3428412  75068  32896 S   0.7   0.1  63:20.86 containerd
2163062 almalin+  20   0 4797108  65568  19992 S   0.7   0.1  53:46.23 dask-worker
3467329 almalin+  20   0   16408   7092   5576 R   0.7   0.0   0:00.17 top
3947642 almalin+  20   0  811908 148688  18232 S   0.7   0.2   5:05.35 jupyterhub-sing
   1294 root      20   0 4272748 138224  58384 S   0.3   0.2 286:48.93 dockerd
```

CMS use-case

- Early Run 3 analysis (2022-2023 data taking)
- Beyond Standard Model searches
- Vector-Like Quark T in T→ tZ channel
- Final state: hadronic Top quark and Z (νν)
- Development of the already published full Run 2 analysis *JHEP05(2022)093*, with the idea to extend the results interpretation to more models predicting the same final state
  - Dark Matter production in association with a Top quark
  - Technicolor models *The Radiative Flavor Template at the LHC*





24/05 WP2.5 presentation *link*

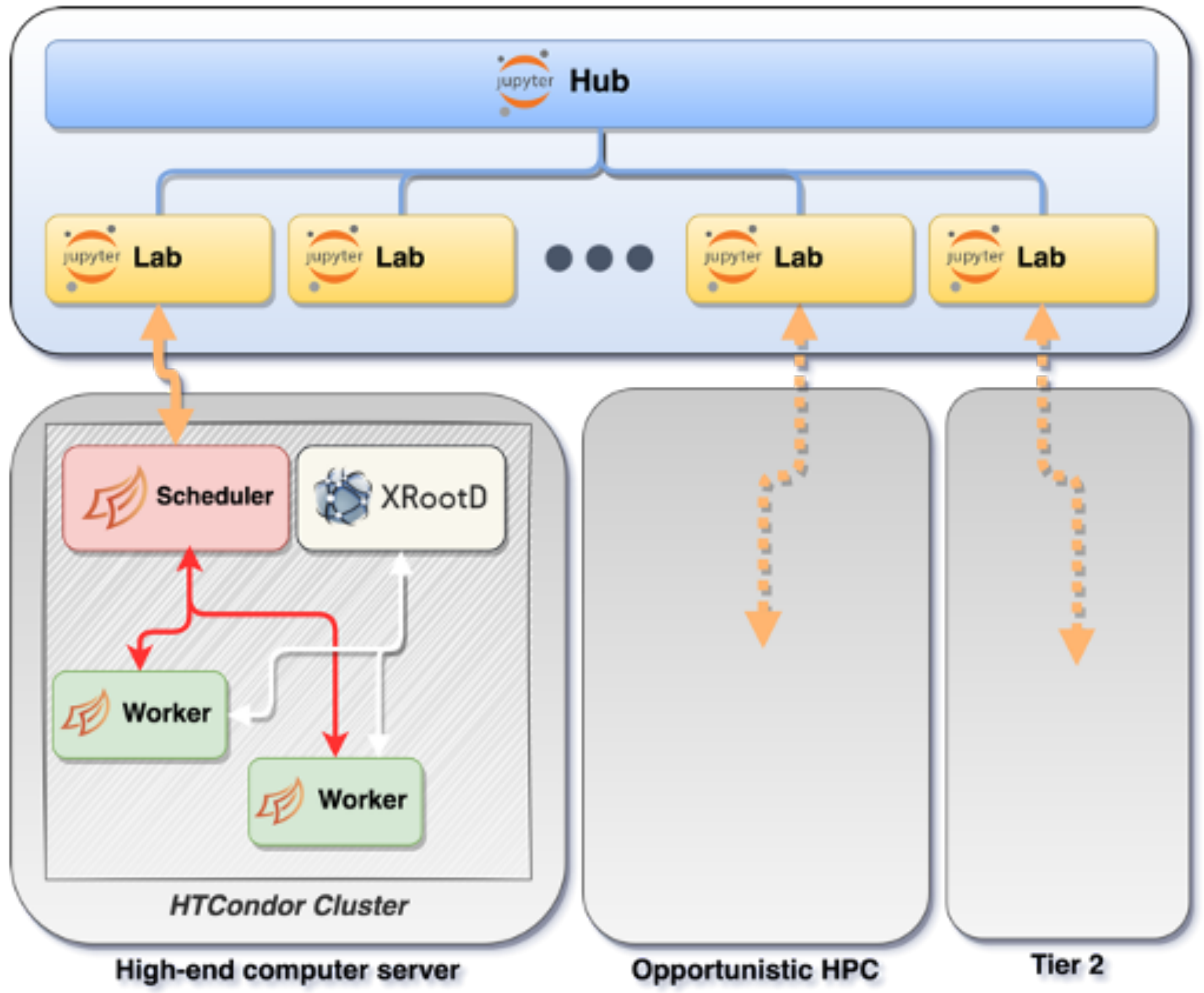
State of the art

Workflow in 2 steps:

- Data preprocessing, evaluation through ML model.
  Using CMS NanoAOD tools (pyROOT-based) and CRAB.
- Skimming and selection using Interactive Analysis
  - Input: ntuple from the 1st step
  - Selection + variables calculation through RDataFrame
  - Distribution of the process using Dask
  - Output: TH1D easy to manage, also possible to store snapshot using remote storages

Working on Perugia's analysis facility

CRAB Cms Remote Analysis Builder



- Analysis still far from the end, more processes will have to be added that will slow it down
- Currently the results are very promising
  Time reduced from ~1d to ~3h and there is still room for development

Conclusions & Next Steps

- Interactive analyses feasibility studies on the INFN Naples infrastructure succeeded
- Three use cases tested (ATLAS, CMS, FCCee), in different scenarios: different experiments and analyses
- Towards an INFN national cloud infrastructure with a datalake model to facilitate future analyses (hopefully starting from LHC Run 3)

➡ **Thoughts about ATLAS use case:**

- I/O implementation to be improved, feasible approach: including grid certificate and voms in the jupyter image
- We currently exploit a cut & count analysis as use case: this is relatively easy from a computational point of view and to fully benefit from the distributed approach we could think about workflows including for example ML techniques

➡ **Long term goals:**

- The current analysis is a successful case of migration to a new more efficient paradigm
- End-goal is to eventually test also the production steps, i.e. exploring the feasibility of a passage: WLCG ➔ Interactive Analysis
- Also in view of HL-LHC and potential future applications

Thank you!

Conferences and workshops contributions

- ECFA 2023 : talk about FCCee Zee benchmark use case —> **done**
- ICHEP2024: contribution accepted as poster (ATLAS SUSY analysis)
- CHEP2024 : contribution submitted via ATLAS computing speakers committee (ATLAS SUSY analysis)
- SIF2024 : contribution submitted (ATLAS SUSY analysis)