Supporting multiple hardware architectures at CMS: the integration and validation of POWER9

Tommaso Boccali[1], Saqib Haleem[2], Dirk Hufnagel[3], Alan Malta Rodrigues[4], Jordan Martins[5], Marco Mascheroni[6], Hasan Ozturk[7], Antonio Pérez-Calero Yzquierdo[8,9], Kirill Skovpen[10], Todor Trendafilov Ivanov[4], Daniele Spiga[11], Christoph Wissing[12]

[1]INFN Sezione di Pisa (Italy), [2]National Centre for Physics (PK), [3]Fermi National Accelerator Laboratory (USA), [4]University of Notre Dame (US), [5]Universidade do Estado do Rio de Janeiro (BR), [6]Univ. of California San Diego (US), [7]CERN (Switzerland), [8]CIEMAT (Spain), [9]Port d'Informaciò Cientifica,PIC, (Spain), [10]Ghent University (BE), [11]INFN Sezione di Perugia (Italy), [12]DESY (Germany)

Computing resources in the Worldwide LHC Computing Grid (WLCG) have been based entirely on the x86 architecture for more than two decades. **Heterogeneous non-x86 resources, such as ARM, POWER and Risc-V, will become available also thanks to their presence in existing and planned world-class HPC installations.**

The CMS experiment has started to prepare for this situation with the CMS software stack (CMSSW) already compiled for multiple architectures.

The ultimate objective of the CMS experiment is to enable the possibility to define a list of supported architectures for any injected workflow and let the underlying computing system transparently handle the technicalities: from job description to the resources selection and site exploitation, where the system detects which architecture the job has landed on and adapts automatically.

Excerpt of a CMS Workflow description

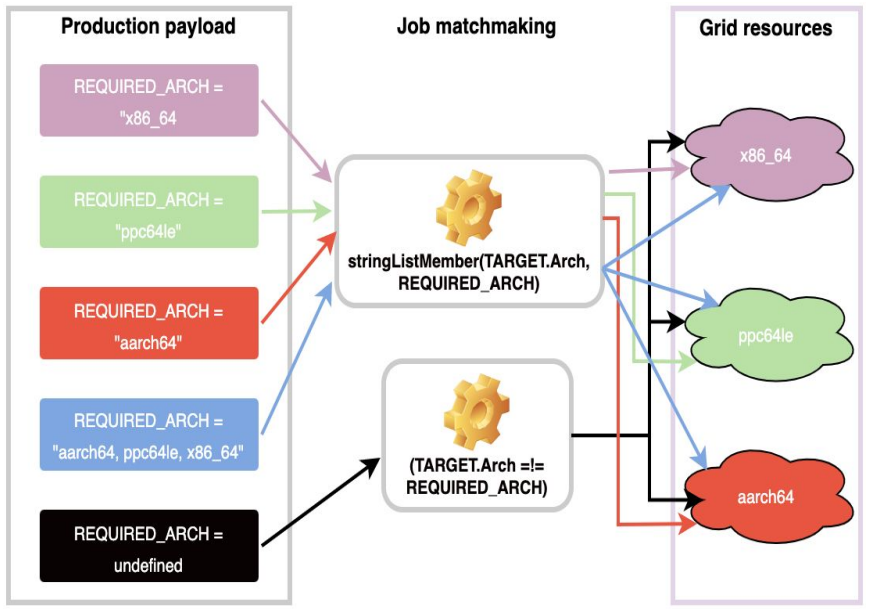
```
{
  "_id": "spiga_TC_SLC7_Marconi_TB_CMS_Marconi_220729_194721_8502",
  "PrepID": "TEST-CMSSW_11_2_0__fullsim_noPU_2021_14TeV-1608392371-ZMM_14",
  "RequestString": "TC_SLC7_Marconi_TB_CMS_Marconi",
  "Comments": {
    "CheckList": "",
    "WorkFlowDesc": ""
  },
  "ProcessingString": {
    "Digi_2021": "Digi_2021_TC_SLC7_Marconi_CMS_Marconi_TBv20210430_test",
    "Reco_2021": "Reco_2021_TC_SLC7_Marconi_CMS_Marconi_TBv20210430_test",
    "ALCA_2021": "ALCA_2021_TC_SLC7_Marconi_CMS_Marconi_TBv20210430_test",
    "ZMM_14TeV_TuneCP5_2021_GenSim": "ZMM_14TeV_TuneCP5_2021_GenSim_TC_SLC7_Marconi_CMS_Marconi_TBv20210430_test"
  },
  "ScramArch": [
    "slc7_ppc64le_gcc9",
    "slc7_amd64_gcc900"
  ],
  "SizePerEvent": 1234,
  "Memory": 3000,
```

Translates into the following HTCondor ClassAd

Requirements = **(stringListMember(TARGET.Arch,"ppc64le,X86_64"))** && (TARGET.OpSys == "LINUX") && (TARGET.Disk >= RequestDisk) && (TARGET.Memory >= RequestMemory) && (TARGET.Cpus >= RequestCpus) && (TARGET.HasFileTransfer)

Once a workflow is injected Requestors define which architectures are supported

WMAgent propagates the architecture requirements from the workflow to the job description, through HTCondor job classads.

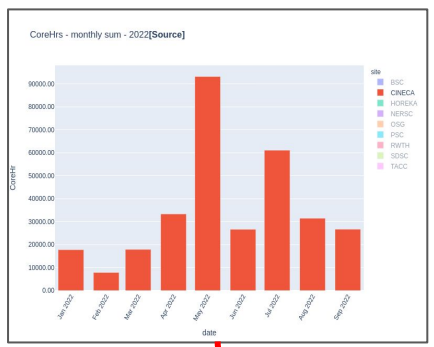


GlideinWMS makes the resource provisioning based on the production payload requirements.

Once computing resources are available, there is a job matchmaking between the grid host architecture and the payloads idle in the condor Schedd.

- Some jobs do not really require any specific architecture

The validation journey@M100



Validation of physics observables in simulated events

Profiting from the opportunity to exploit the first sizable IBM Power9 allocation available on Marconi100 HPC system at CINECA, CMS developed a validation process. The strategy has been to run Release Validation MonteCarlo workflows both with and without Pileup in order to perform comparisons with the very same samples produced on x86, using the same CMSSW release running on infrastructures at CERN.

The validation has been performed using the regular production system. The operations team introduced a new feature to identify such workflows and assigned them to the site where the PPC machines are available. Thanks to using the production pipeline, we took advantage of existing features such as automatic input and output placements.

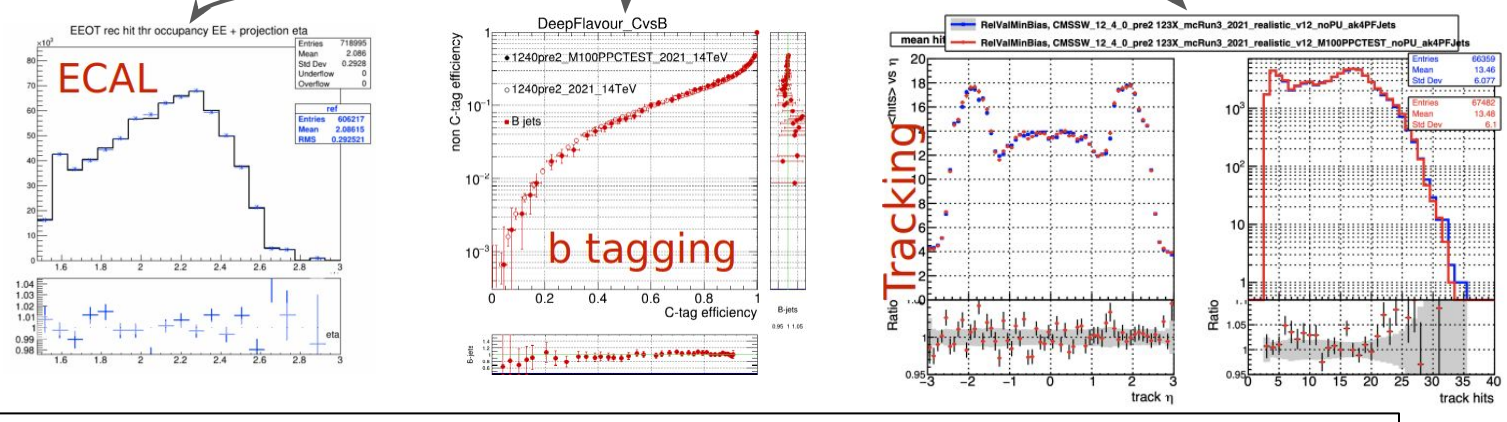
FullSim

Release Name	Tracker	Ecal	HGcal	Hcal	CASTOR	DT	CSC	RPC	GEM	MTD	PPS	L1	Tracking	Electron	Photon	Muon	Jet	MET	bTag	Tau	PF	Info	RelMon
12_4_0_pre2_M100PPC	✓	✓	–	✓	–	✓	✓	✓	–	–	–	–	✓	✓	✓	✓	✓	✓	✓	✓	✓	i	✗

Various physics processes generated: **top quark pair production, Drell-Yan, W+jets, SUSY, etc.**

Compare distributions of physics observables between the two productions

Results of the comparisons are reported by various subsystem experts in detector and physics object reconstruction.



The POWER CPU architecture can be now used by CMS for generation and processing of physics data

- **Marconi 100@CINECA and OLCF Summit will be fully exploited**

The computing infrastructure of the experiment has been enabled **to transparently handle heterogeneous non-x86_68 resources**. The support of multiple architectures, in principle any, has been fully validated

- **Currently the CMS core software supports x86_68, ARM and POWER**