CMS O&C
Offline Software and Computing

Experience with AlmaLinux

J. Letts (UCSD), D. Piparo (CERN) - Linux Future Committee Meeting - January 16, 2023

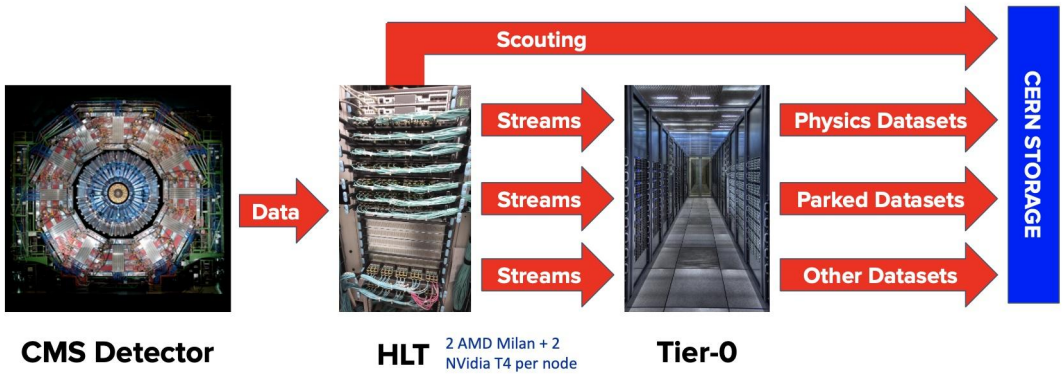
Caveat

- Some of the statements made in these slides about the CMS computing infrastructure and data processing are ***approximate***
- The objective is to give ***the right context for the audience***
- The focus is to talk about the experience with AlmaLinux!
- Feel free to ask questions during and at the end of the talk (and even after!)

Introduction

Data Flow from the Experiment to the Tier-0 from 10 Km

Data from the detector flows to the High-Level Trigger farm (HLT), which for Run 3 is equipped with NVidia T4 GPUs. The HLT sends data files of data separated according to the various trigger categories.



Source: sciencenode

The Tier-0 processes the data files and outputs RAW and other specialized streams data to CERN storage, divided into “Primary Datasets” (PDs), e.g. dimuon, e-gamma, JetMET.

- One archival copy of the RAW physics data is made at CERN.
- A second, working copy is transferred to one of the CMS Tier-1 sites (except for “parked” and “scouting” data set types e.g.)

Data processing in a Nutshell: Spotlight on Linux

Central Processing

- The CMS application, cmsRun, is executed in a variety of contexts with very different goals
  - CMSSW is highly configurable, dynamic library load: **1 executable for all central use cases**
- Data **reconstruction**: runs on events in RAW format acquired with CMS or simulated datasets
- **High Level Trigger**: fast reconstruction, 300 ms VS seconds, runs at the HLT farm at P5
- **Simulation** (and generation): runs from a random seed and produces datasets in RAW format *as if* they had been acquired by CMS
- Simulation, reconstruction and HLT (therefore online and offline): **same software stack, same or very similar releases, compiled consistently with the same compiler for the same operating system**

User workflows

- **Rich set of use cases**: analysis, development, debug, profiling, tests…

All the use cases above need a solid environment, made of software built consistently on a **stable, reliable, secure, operating system**

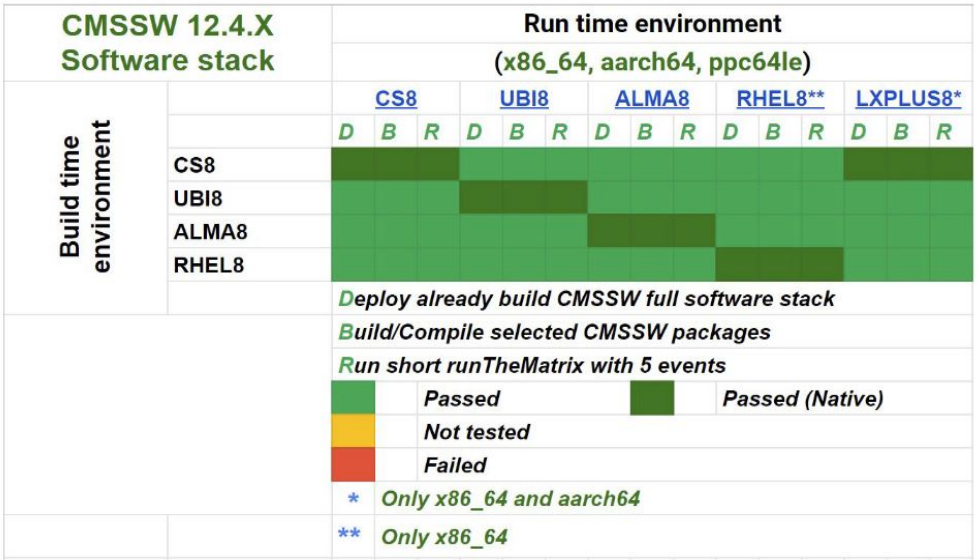
Elements of Software Distribution and Operating System

- **CMS distributes its software via CVMFS**, an aggressively cached distributed POSIX read-only file system
- **CMS software runs in containers with Singularity** offline (CERN, HPCs, Tier-1,2,3, clouds)
  - Container images are also distributed with CVMFS
- Flexibility to decide which OS to adopt
- All software, operating system and build mechanism (compiler and flags), CPU architecture **need to produce physics results validated**: a process involving experts comparing physics outputs
  - One production platform (OS+CPU arch) had to be identified for Run 3
- It was crucial for CMS to choose a distribution that:
  - **Supported modern and future accelerators** (e.g. GPUs at HLT and HPCs)
  - Can be **deployed identical online (HLT) and offline** (stability requirements online are strict)
  - Supported **Power, ARM and x86 CPUs**
  - Was **free and open source**
- CentOS Stream was and is a good distribution. Unfortunately it did not meet CMS requirements
- We started in depth investigations of alternatives for our containers…

A Taste of the Investigation Effort

- We picked a few options to start test our software
- Thorough exercise, substantial effort: we had to gain confidence
- Check that sw built on distro_1, could also run on distro_2: build the “CMSSW Linux Compatibility Matrix”

This work was highly not trivial: done during Q1 2021!



It was reassuring to see ample compatibility among all different distributions, for all platforms.

Side rememark: clear demonstration of the flexibility of CMS software and building infrastructure!

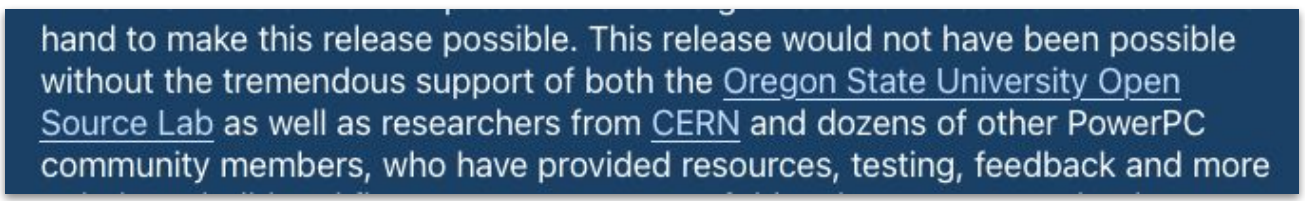
CMSSW: 5MLOC detector specific code + 400 3rd party packages

Work driven by Shahzad Muzaffar (EP-SFT and CMS)

Experience with Alma

CMS Solution

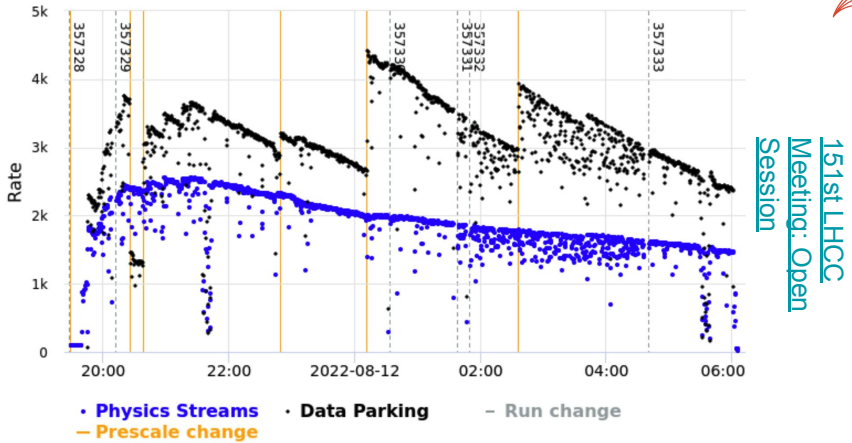
- **Adopt AlmaLinux8 for Offline** processing, building and integration
- **Adopt RHEL8 at P5**, including the HLT nodes
- **Deploy the same release offline and online**
- Since Alma and RHEL are identical, call the platform in our builds EL8
  - E.g. *el8_aarch64_gcc11*, *el8_amd64_gcc11*, *el8_ppc64le_gcc11*
  - Fore more info: *ls /cvmfs/cms.cern.ch/*
- Scientists can develop, test and debug on *8 (remember the Linux Matrix - anyway all changes are tested by the central CI as well)
  - Can switch to AlmaLinux8 container with a simple CMS script



https://almalinux.org/blog/almalinux-for-powerpc-85-stable-now-available/

… And it worked!

Successful prompt processing @ > 5 kHz



151st LHCC Meeting: Open Session

CMS-PAS-TOP-2022-12

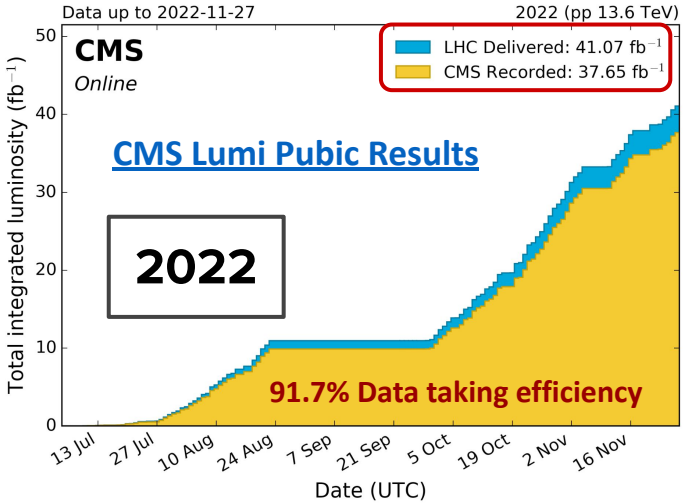
Contact: cms-pag-conveners-top@cern.ch 2022/09/21

First measurement of the top quark pair production cross section in proton-proton collisions at $\sqrt{s} = 13.6\,\text{TeV}$

The CMS Collaboration

151st LHCC Meeting: Open Session

- HLT highlights
  - GPU offload 40% of evt. processing
  - Yields 70% increase in throughput
  - Consistent results CPU vs GPU
  - Graph Neural Network for jet tagging



Summary of the Experience so far

- **AlmaLinux8 worked very well for offline production and data taking**
  - The combination with RHEL8 deployed at the HLT worked very well too
- **The AlmaLinux team is very responsive and helpful**
  - Their support was crucial when it came to the support of the Power architecture
  - The interactions were very rewarding!
- We could verify a **high degree of compatibility with other Linux flavours**, e.g. Stream, UBI or Rocky
- If AlmaLinux is not available but CVMFS is, CMS scientists can switch with one command to the CMS AlmaLinux container
  - For development, testing, profiling or debugging
  - **We wish Alma8 portals on LXPLUS stay available to us at least until the end of LS3 to make the access to the OS even easier at CERN and elsewhere**