# Automatized Compilation of Large Packages Example of mpdroot

#### Ján Buša Jr.

Mescheryakov Laboratory of Information Technologies, JINR, Dubna, Russia

Institute of Experimental Physics, SAS, Košice, Slovakia

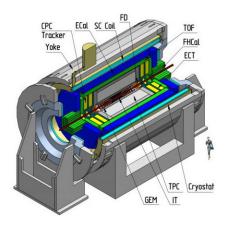
Dubna 15. 11. 2022





## Multi-Purpose Detector

- part of Nuclotron-based Ion Collider fAcility
- Multi Purpose Detector main detector under costruction
- mpdroot library for reconstruction of tracks and physical analysis (http://mpdroot.jinr.ru/)



## Just Install FairRoot and Build mpdroot...

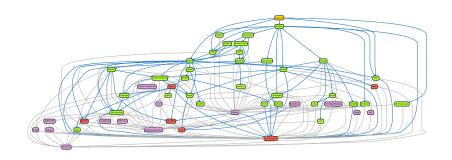
How to install MPDroot (each user had to build own version):

- install all basic packages (about 70);
- download, patch, and build FairRoot;
- download and build mpdroot from sources. In-source build highly recomended.

It's really "easy":

- basic packages are often not available in compatible versions
- patching FairRoot can be tedious (during last major version change more than 100 lines of code)
- building FairRoot takes about 2-4 hours on a regular computer
- in-source build guides users to use improper paths when linking own code
- users tend to use same version of mpdroot for long time

# The mpdroot Dependencies



### Centralized Build of MPDroot

#### Requirements:

- we want to deliver latest version of MPDroot to the users
- all dependencies have to be statisfied and no version clashes are allowed
- results need to be repeatable

#### Solution:

- CVMFS robust distribution of software not only between clusters/supercomuters
- aliBuild set of tools for building software together with its dependencies
- toolbox rootless wrapper for podman (docker). Supported on most linux-based systems. On MacOS and Windows one can use docker
- EnvironmentModules environment enabling loading multiple versions of the same software with its dependencies

# CernVM-FS https://cernvm.cern.ch/fs/

- CMVMFS is a read-only file system
- anyone can mount the shared directories to its computer
- extensively uses caching of the files on local computer
- only used (accessed) files are downloaded. That makes first load slow but following are as fast as local usage
- since it is centralized, changes on server are directly available to all users

# Toolbox https://github.com/containers/toolbox

- wrapper for podman (rootless docker alternative) ordinary user can use it
- allows user to "work" inside virtual machine no connection to the users system
- we support CentOS7, but to test our software we build also CentOS Stream 8 and 9 versions
- basic system image is fixed without any updates and therefore all results are reproducible

# Environment Modules https://modules.readthedocs.io/

- started as a project to separate various versions of same software
- simple "module files" contain paths to the software together with list of dependencies
- to use some software one has to load it first via command module add mpdroot/latest. Paths to executables as well as libraries is than added in the commandline. All necessary dependencies are loaded as well.

### Example of module file (excerpt):

```
module load BASE/1.0 GEANT4/$GEANT4_VERSION-$GEANT4_REVISION module load ROOT/$ROOT_VERSION-$ROOT_REVISION setenv MPDROOT \$PKG_ROOT prepend-path ROOT_INCLUDE_PATH \$PKG_ROOT/include prepend-path LD_LIBRARY_PATH \$PKG_ROOT/include
```

# aliBuild https://alisw.github.io/alibuild/

- created as a solution to the ALICE experiment software distribution problem
- automatic build of software and all its dependencies
- reusing already built packages
- if some package was changed, all dependants are rebuild
- apropriate modules are created and numbered automatically
- can get messy if packages change frequently and are not cleared
- based on recipes guides how to build some package

## mpdroot Recipe

```
package: mpdroot
version: "%(commit_hash)s"
tag: "dev"
source: https://git.jinr.ru/nica/mpdroot.git
requires:
  - FairRoot
 - eigen3
 - nlohmann ison
  - VMC
  - GSI.
unset SIMPATH
export MPDROOT=$INSTALLROOT
cmake $SOURCEDIR
      ${BOOST_ROOT:+-DBOOST_ROOT=$BOOST_ROOT}
      ${EIGEN3 ROOT:+-DEIGEN3 ROOT=$EIGEN3 ROOT}
      ${LIBXML2_ROOT:+-DPC_LIBXML_LIBDIR=$LIBXML2_ROOT/lib}
      ${NLOHMANN_JSON_ROOT:+-DNLOHMANN_JSON_ROOT=$NLOHMANN_JSON_ROOT}
      -DCMAKE INSTALL PREFIX=$MPDROOT
cmake --build . -j$JOBS install
# Modulefile
MODULEDIR="$INSTALLROOT/etc/modulefiles"
MODULEFILE="$MODULEDIR/$PKGNAME"
mkdir -p "$MODULEDIR"
alibuild-generate-module --bin --lib > $MODULEFILE
cat >> "$MODULEFILE" <<EoF
setenv MPDROOT \$PKG ROOT
seteny MPDROOT MACROS \$PKG ROOT/macros
prepend-path ROOT_INCLUDE PATH $::env(BASEDIR)/GSL/$GSL_VERSION-$GSL_REVISION/include
EoF
```

# **Changing Defaults**

```
default:
  - &MPDROOT TAG v22.09.22
  - &GCC_TOOLCHAIN_TAG v10.2.0-alice2
package: defaults-release
version: v1
disable:
  - arrow
env:
  CMAKE BUILD TYPE: RELWITHDEBINFO
  CXXFLAGS: -fPIC -02 --std=c++17
  ENABLE VMC: 'ON'
overrides:
  mpdroot:
    tag: *MPDROOT_TAG
    version: *MPDROOT TAG
  ROOT:
    tag: "v6-26-08"
    requires:
      - GSI.
      - libxml2
  GCC-Toolchain:
    tag: *GCC TOOLCHAIN TAG
    version: *GCC TOOLCHAIN TAG
```

## Wrapping Up

#### Our view:

- software is stored in git
- on each push to master or creation of tag, command aliBuild --build --no-deps --defaults releaseName is called and built package is placed into CVMFS. Software is available no later than 1h after being sucessfully built.

### Developer's view:

- toolbox enter c7-nica
- source /cvmfs/nica.jinr.ru/sw/login/login.sh
- module add mpddev/latest
- export MPDROOT=/path/to/install
- git clone path\_to\_mpdroot && mkdir build && cd build && cmake .. && make install

## Wrapping-Up (Continued)

#### User's view:

- toolbox enter c7-nica
- source /cvmfs/nica.jinr.ru/sw/login/login.sh
- module add mpdroot/latest

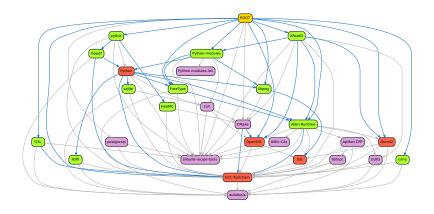
#### or concrete version

module add mpdroot/v22.09.22-1

### Conclusions

- for ordinary users build of mpdroot takes no time
- latest version of software is instantly available to everyone
- it is easy to return to previous versions (including dependencies)
- all results are repeatable since we store basic system image as well as all recipes
- most of the packages we build from scratch (including low level, s.a., CMake, GCC, OpenSSL)
- with recipes new dependencies are distributed among developers without necessity for them to install anything (it is done automatically)

## All Great Except...



## It can't be that bad, can it?

			refer time outer DEDNIES (42.47	v06-04		
ts/latest		FairRoot/v18.6.8-5	generators/latest	libxml2/v2.9.3-2	protobuf/v3.15.8-3	SMASH/latest
s/latest-release		FairRoot/v18.6.8-6	generators/latest-release	libsml2/v2.9.3-3	protobuf/v3.15.8-4	SMASH/latest-release
	004hep/latest					
	004hep/latest-release				pythia/v8.307-1	
build-recipe-tools/0.2.3-1						
ouild-recipe-tools/0.2.3-2						
build-recipe-tools/latest	004hep/v01-22-3	flatbuffers/latest			pythia/v8243-alicela-3	
ouild-recipe-tools/latest-release		flatbuffers/latest-release			pythia/v8243-alice1a-4	
			googletest/1.8.0-1			
		flatbuffers/v1.12.0-2	googletest/1.12.1-1			
		flatbuffers/v1.12.0-3	googletest/latest		pythiaG/428-alice2-1	
		flatbuffers/v1.12.0-4	googletest/latest-release			
	EDM4hep/latest	flatbuffers/v1.12.1-1				
	EDM4hep/latest-release					
	EDM4hap/v80-86-1					
n-Runtime/v2-19-le-3	EDM4hep/v80-06-2		OSL/latest	mpddev/v22.09.22-3		
	EDM4hep/v00-05-3					
			HepMC/HEPMC_02_06_10-1			
			HepMC/HEPMC_02_06_10-2			
			HepMC/HEPMC_02_06_10-3			
			HepMC/HEPMC_02_06_10-4			
	EnvironmentModules/latest		HepMC/latest			
	EnvironmentModules/latest-release		HepMC/latest-release			
	EnvironmentModules/v5.0.1-1		HepMC3/3.2.5-1			
	EnvironmentModules/v5.0.1-2			NICA_scheduler/22.88.0-2	Python-modules/latest	
	EnvironmentModules/v5.1.1-1			NICA_scheduler/latest	Python-modules/latest-release	
	FairCMakeModules/latest		HepMC3/3.2.5-4			
	FairCMakeModules/latest-release		HepMC3/latest	ninja/fortran-v1.8.2.g3b-1		
	FairCMakeModules/v0.2.0-1		HepMC3/latest-release			
	FairCMakeModules/v0.2.0-2					
	FairCMakeModules/vl.0.0-1				Python/v3.6.10-3	
	FairCMakeModules/v1.0.0-2					
st/v1.75.0-1				nlohmann_json/latest		
it/v1.75.0-2						
			lhapdf/latest	nlohmann_json/v3.11.2-1		XRootD/v5.3.2-1
	FairLogger/v1.10.1-2		lhapdf/latest-release			XRootD/v5.3.2-2
	FairLogger/vl.11.0-1		lhapdf/v6.2.1-alice2-1			XRootD/v5.3.2-3
	FairLogger/v1.11.0-2		lhapdf/v6.2.1-alice2-2		ROOT/latest-release	XRootD/v5.4.2-alice
	FairLogger/vl.11.1-1		lhapdf/v6.2.1-alice2-3		R00T/v6-24-06-3	XRootD/v5.4.2-alice
			lhapdf/v6.2.1-alice2-4			XRootD/v5.4.2-alice
	FairMQ/latest		lhapdf/v6.2.1-alice2-5			yaml-cpp/latest
	FairMQ/latest-release					yaml-cpp/latest-rel
						yaml-cpp/yaml-cpp-0
				OpenSSL/v1.0.2o-1	R00T/v6-26-06-1	
				OpenSSL/v1.0.20-2	R00T/v6-26-08-1	
				OpenSSL/v1.1.1m-1		
				OpenSSL/v1.1.1m-2	simulation/latest-release	ZeroMQ/latest
		GEANT4_VMC/latest-release	libpng/latest		simulation/v1.0-1	ZeroMQ/latest-relea
/latest-release	FairMQ/v1.4.55-1	GEANT4_VMC/v5-3-1	libpng/latest-release	podio/latest-release	simulation/v1.0-2	ZeroMQ/v4.3.3-1
g/latest	FairMQ/v1.4.55-2	GEANT4_VMC/v5-3-2	libpng/v1.6.34-1	podio/v00-15-1	simulation/v1.0-3	ZeroMQ/v4.3.3+2
g/latest-release	FairRoot/latest	GEANT4_VMC/v5-3-3	libpng/v1.6.34-2	podio/v00-15-2	simulation/v1.0-4	ZeroMQ/v4.3.3-3
E/v13.0.0-4	FairRoot/latest-release	GEANT4_VMC/v5-3-4	libpng/v1.6.34-3	podio/v80-16-1	simulation/v1.0-5	ZeroMQ/v4.3.3-4
E/v13.0.0-5	FairRoot/v18.6.4-1	GEANT4_VMC/v6-1-1	libong/v1.6.34-4	postgresql/latest	simulation/v1.0-6	zlib/latest
ie/latest	FairRoot/v18.6.4-2	GEANT4 VMC/v6-1-2	libtirpc/latest	postgresql/latest-release	simulation/v1.0-7	zlib/latest-release
e/latest-release	FairRoot/v18.6.4-3	GEANT4 VMC/v6-1-3	libtirpc/latest-release	postgresol/REL 14 2-1	simulation/vi.0-8	zl1b/v1.2.8-1
le/v3.19.2-1	FairRoot/v18.6.4-4	GEANT4 VMC/v6-1-4	libtirpc/libtiruc-1-1-4-1		simulation/v1.0-0	zlib/v1.2.8-2
se/v3.19.2-2	FairRoot/v18.6.5-1	GEANT4 VMC/v6-1-5	libtirpc/libtirpc-1-1-4-2	protobuf/latest	simulation/v1.0-10	zlib/v1.2.8-3
se/v1.23.1-1	FairRoot/v18.6.5-2	GEANT4 VMC/vG-1-6	libxml2/latest	protobuf/latest-release	simulation/v1.0-11	
ke/v1.21.1-2	FairRoot/v18.6.8-3	GEANTA VMC/vG-1-7	libxml2/latest-release	protobuf/v3.15.8-1	simulation/v1.0-12	
1/7.78.9-1	FairRoot/v18.6.8-4	GEANT4_VMC/v6-1-8	libxml2/v2.9.3-1	protobuf/v3.15.8-2	SMASH/2.2.1-1	

## Help Needed

- (semi)automatic cleaning of obsolete packages needs to be added
- GUI for cleaning of packages would be nice to have
- accounting which packages were used and what is the distribution of their usage in time
- distribution of mpdroot (and dependencies) via rpms
- protection of closed-source software
- sometimes even new recipes are not easy to add

## Bachelor/Master/...

What I need to know:

nothing – kind of

What will I get:

diploma

What will I learn (part of/maybe all):

- git (not only for storing files)
- bash, CMake, aliBuild/aliDist
- Docker/Toolboxes, RHEL based packaging
- creating web-based application
- PLEX