

Order versus Chaos

Software Organisation in a Larger Project

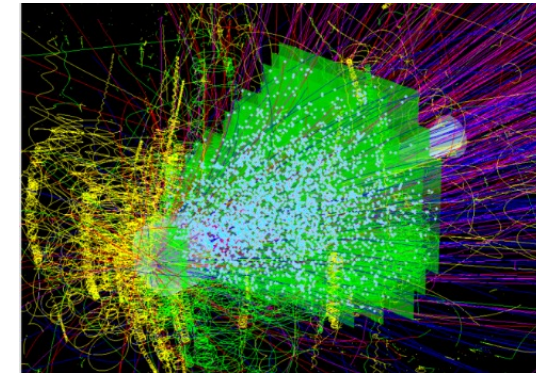
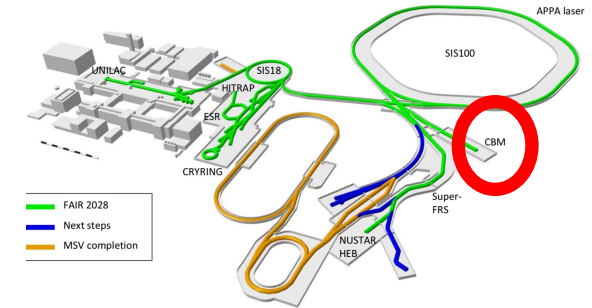
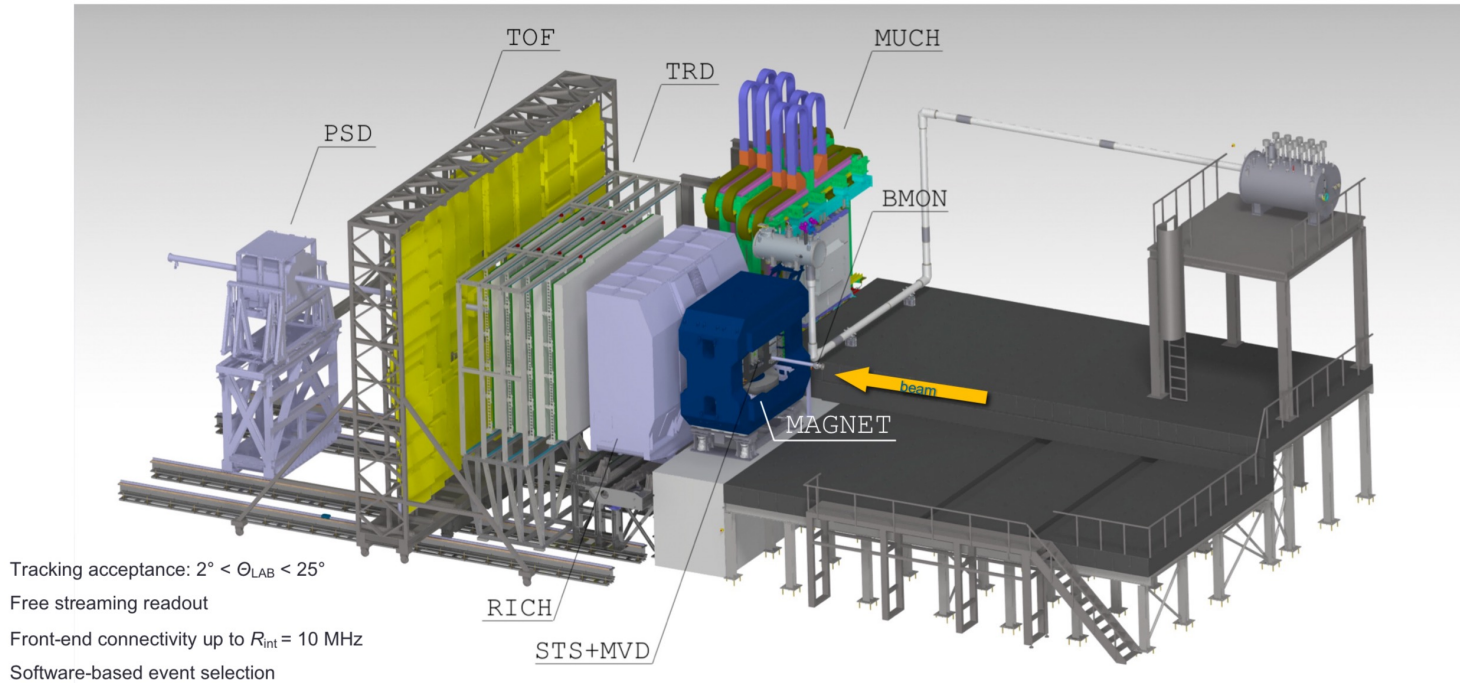
V. Frieze
GSI C++ User Group
11 June 2025

Disclaimer

This is not about to tell you how to organise larger software projects.

This is to ask for advise and discuss and learn best practices how to organise larger software projects - on the example of cbmroot.

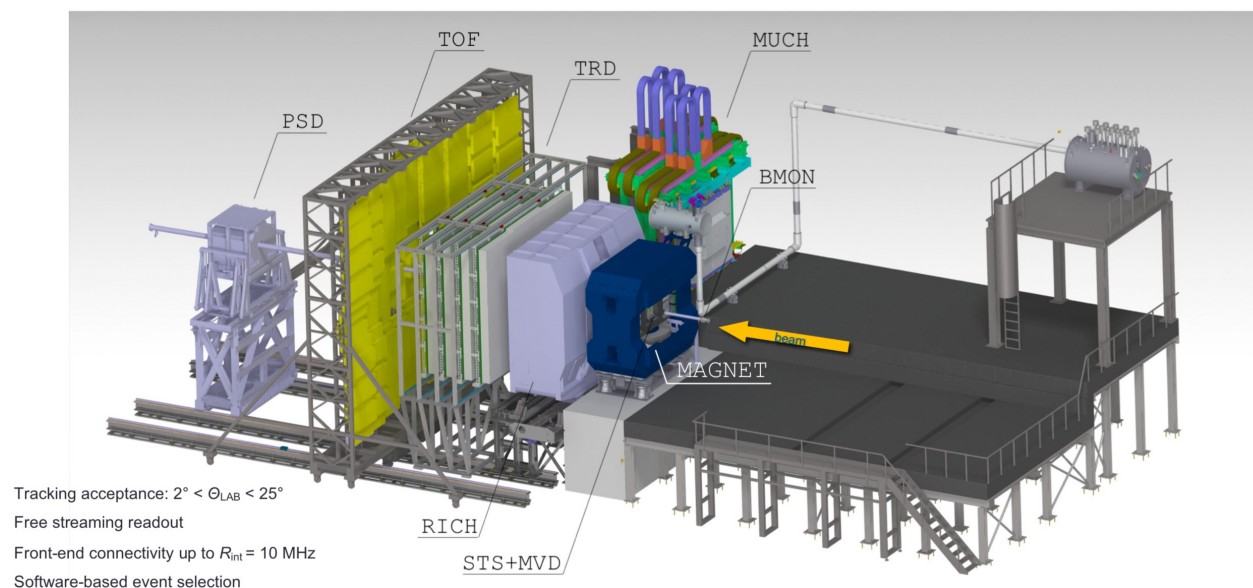
Setting the Stage: The CBM Experiment



- A heavy-ion experiment at FAIR, to operate from 2028 on with beams from SIS-100
- Fixed-target; high interaction rates; no hardware trigger
- Needs online data selection in software.

Mission

Establish the software needed to run the experiment and extract physics results from its data.



Many facets:

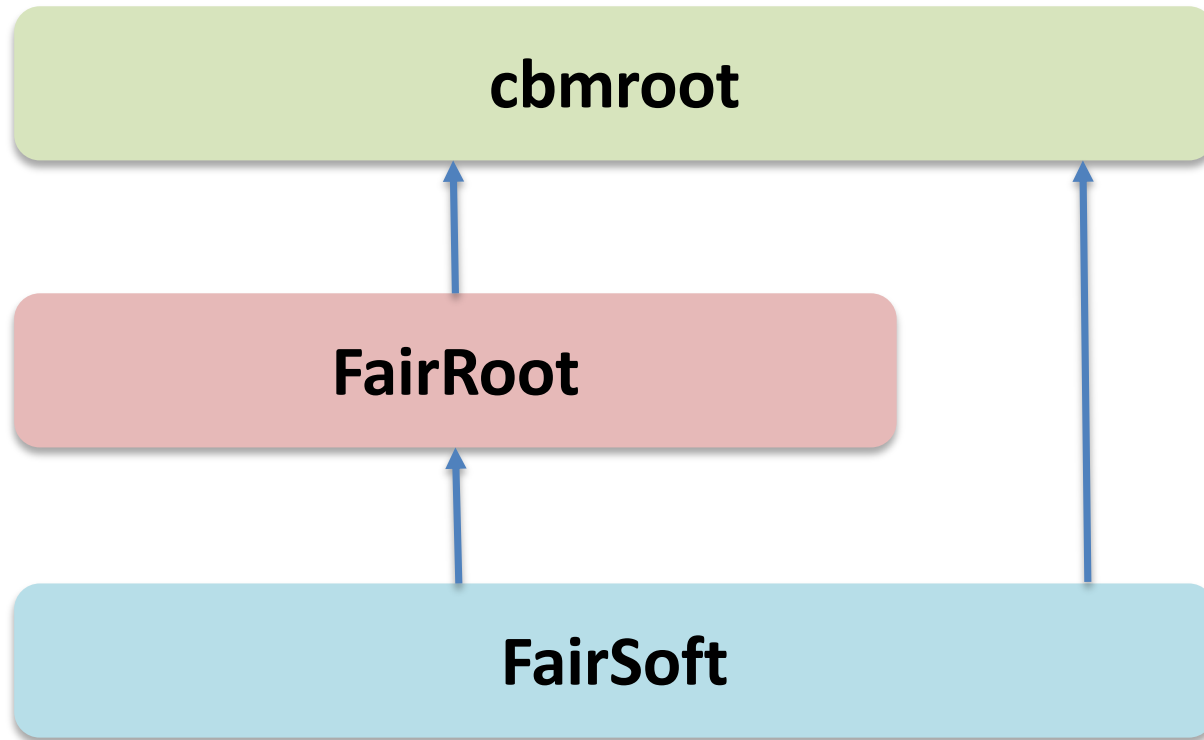
- Simulate the experiment
- Configure and control the experiment
- Digest raw data in real time to the level of physics-based data selection.
- Reconstruct events and tracks from the free-streaming data.
- Implement the offline physics analysis software.

About 500 GB/s raw data from experiment

Several GB/s archival rate

O(10) PB / year archived data volume

cbmroot



- A long history of about two decades
- Starting point was simulation and offline computing.
- Builds on top of FairRoot, using ROOT as a platform
- A number of external packages (FairSoft): ROOT, GEANT3/4, BOOST,
- Online data processing (HPC) in the focus since about 5 years.
- Using mCBM as test bed.


Cbmroot: Current Organisation

The screenshot displays the GitLab repository page for 'cbmroot'. The interface includes a sidebar with navigation options (Project, Pinned, Merge requests, Manage, Plan, Code, Build, Secure, Deploy, Operate, Monitor, Analyze, Settings) and a main content area showing the repository details. The repository name 'cbmroot' is prominently displayed, along with a search bar and a table of files. The table lists files with their names, last commit messages, and last update times. The right sidebar provides project information, including the number of commits (3,984), branches (14), tags (239), project storage (8.8 GiB), and releases (4). It also includes links to the README, GNU General Public License, CI/CD configuration, GitLab Pages, Add Kubernetes cluster, Add Wiki, and Configure Integrations.

Name	Last commit	Last update
.ci/online	ci: Bump online container to FairSoft 22...	3 months ago
MQ	Remove non ASCII characters	2 months ago
algo	BugFix: OnlineRecoCanProcessTimeslic...	4 days ago
analysis	Apply clang-format	1 month ago
cmake	Add hal directory to paths.	2 months ago
core	Merging CbmYamlInterface with CbmUt...	6 days ago
docs	Jupyter kernelspec files	10 months ago
doxygen	Code cosmetics. Add newline at end of ...	4 years ago
external	Bring in rich_v25a	1 day ago
files	TOF AFCK pulser monitors: fixes and n...	2 weeks ago
macro	Bring in rich_v25a	1 day ago
reco	Merging CbmYamlInterface with CbmUt...	6 days ago
scripts	[CI] Edit checklibs.sh to catch all HAL li...	2 months ago
services	cbm-online-par-dump: replacing cbm:...	1 month ago
sim	update Geometry of Rich detector to ve...	3 weeks ago
tutorials	Remove unneeded target dependency	1 year ago
.clang-format	algo: Rename config: to yaml:.	1 year ago
.clang-tidy	Add clang-tidy configuration	1 year ago
.dockerignore	Build and deploy containers for online r...	1 year ago
.gitignore	Add git rebase original file backups to ...	7 months ago
.gitlab-ci.yml	[CI] Disable installation test for vae25 d...	19 hours ago
CMakeLists.txt	Fix several CMake warnings	2 months ago

- One single project / repository
- Hosted in gitlab (CI/CD)
- Almost completely C++ code
- 300 k LOC, about 100 shared libraries
- Comprises a variety of functionality (online / offline)
- Long history, many contributors: strong variation in level of coding and C++ standards
- Constantly evolving: started as typical offline framework, now more and more online and performance aspects introduced.

Cbmroot: Current Organisation

<div> CbmRoot</div> <div>< PREV CURRENT NEXT ></div> <div>DashboardCalendarProject</div>												
Continuous 136 builds [view timeline]												
Site	Build Name	Update	Configure		Build		Test				Start Time	Labels
		Revision	Error	Warn	Error	Warn	Not Run	Fail	Pass	Time		
FairRoots-MacBook-Pro	macOS1015-darwin-x86_64-gcc12.0.0-fairsoft_apr21p2-fairroot_v18.8.7_MR-2115	a758dd	0	0	0 ₋₂	63	0	0	195 ⁺¹⁹⁵	16m 1s	15 hours ago	(none)
FairRoot-macosx12	macOS12-darwin-x86_64-gcc14.0.0-fairsoft_nov22p1-fairroot_v18.8.0_MR-2115	a758dd	0	0	0 ₋₂	71 ⁺² ₋₂	0	0	195 ⁺¹⁹⁵	17m 59s	15 hours ago	(none)
FairRoots-iMac-Pro	macOS13-darwin-x86_64-gcc14.0.0-fairsoft_nov22p1-fairroot_v18.8.0_MR-2115	a758dd	0	0	0 ₋₂	71	0	0	195 ⁺¹⁹⁵	16m 59s	15 hours ago	(none)
FairRoots-iMac-Pro	macOS11-darwin-x86_64-gcc13.0.0-fairsoft_nov22p1-fairroot_v18.8.0_MR-2115	a758dd	0	0	0 ₋₂	68	0	0	195 ⁺¹⁹⁵	17m 45s	15 hours ago	(none)
FairRoots-iMac-Pro	macOS11-darwin-x86_64-gcc13.0.0-fairsoft_apr21p2-fairroot_v18.8.7_MR-2115	a758dd	0	0	0 ₋₂	67 ⁺¹ ₋₁	0	0	195 ⁺¹⁹⁵	14m 38s	15 hours ago	(none)
FairRoot-macosx12	macOS12-darwin-x86_64-gcc13.1.6-fairsoft_apr21p2-fairroot_v18.8.7_MR-2115	a758dd	0	0	0 ₋₂	79	0	0	195 ⁺¹⁹⁵	14m 27s	15 hours ago	(none)
kn12	vae25-linux-x86_64-gcc12-fairsoft_jan24p1-fairroot_v18.8.2_MR-2115	a758dd	0	2	0	74	0	0	195	16m 53s	15 hours ago	(none)
kn12	vae25-linux-x86_64-gcc12-fairsoft_nov22p1-fairroot_v18.8.0_MR-2115	a758dd	0	2	0	16	0	0	195	17m 25s	15 hours ago	(none)
kn12	vae24-linux-x86_64-gcc10-fairsoft_jan24p1-fairroot_v18.8.2_MR-2115	a758dd	0	3	0	74	0	0	195	13m 50s	15 hours ago	(none)
kn12	vae24-linux-x86_64-gcc10-fairsoft_nov22p3-fairroot_v18.8.0_MR-2115	a758dd	0	3	0	16	0	0	195	15m 35s	15 hours ago	(none)

FirstPrevious12345NextLast

- Code quality management in CDash
- About 200 tests (continuous / nightly / weekly)
- Mostly checking code integrity (flawless compilation and execution)
- Automated check of results started, but is in infancy yet.
- Test coverage about 25 %.

Cbmroot: Current Organisation

CbmRoot Suche:

[+ Übersicht](#) [Aktivität](#) [Roadmap](#) **[Tickets](#)** [Aufgewendete Zeit](#) [Gantt-Diagramm](#) [Kalender](#) [News](#) [Dokumente](#) [Wiki](#) [Foren](#) [Datelen](#) [Repository](#) [Konfiguration](#)

Tickets Neues Ticket

Filter Filter hinzufügen

☒ Status offen

[Anwenden](#) [Zurücksetzen](#) [abfrage speichern](#)

<input type="checkbox"/>	#	Tracker	Status	Priorität	Thema	Zugewiesen an	Abgabedatum	Aktualisiert	% erledigt
<input type="checkbox"/>	3647	Bug	New	Normal	Remove ClassImp macros from source code	Volker Friese		11.06.2025 09:41	...
<input type="checkbox"/>	3646	Bug	New	Normal	Lustre/Virgo Account	Florian Uhlig		10.06.2025 12:25	...
<input type="checkbox"/>	3643	Bug	New	Normal	TRD ASIC and GBTX board material	Philipp Kähler		03.06.2025 17:45	...
<input type="checkbox"/>	3642	Feature	New	Normal	BBA issues	Sergey Gorbunov		10.06.2025 12:54	...
<input type="checkbox"/>	3641	Bug	In Progress	Normal	kf_kparticle.C : unable to load	Florian Uhlig		31.05.2025 19:49	...
<input type="checkbox"/>	3637	Support	New	High	simulation does not work for CBMROOT installed on /cvm/cbm.gsi.de/debian11/	Florian Uhlig		22.05.2025 16:45	...
<input type="checkbox"/>	3636	Support	New	Normal	RICH CFV v23b creation code	Simon Neuhaus		10.06.2025 10:59	...
<input type="checkbox"/>	3635	Bug	New	Normal	Conflict of external/flesnet/lib/monitoring /System.hpp and external/flesnet/lib/fles_ipc /System.hpp	Jan de Cuveland		22.05.2025 01:32	...
<input type="checkbox"/>	3629	Feature	New	Normal	Adaption of Unigen UrQMD converter	Krzysztof Piasecki	19.05.2025	28.05.2025 23:55	...
<input type="checkbox"/>	3626	Sim-Development	New	Urgent	Secondary vertex finding issue (pp collisions)	Sergey Gorbunov		14.05.2025 09:32	...
<input type="checkbox"/>	3625	Bug	In Progress	High	Rotation of the beam is wrongly implented	Radim Dvorak		10.06.2025 18:23	...
<input type="checkbox"/>	3624	Sim-Development	New	Urgent	Few tracks secondary vertexing	Sergey Gorbunov		19.05.2025 11:32	...
<input type="checkbox"/>	3623	Bug	Resolved	Normal	run_info binary not installed	Sergei Zharko		13.05.2025 16:15	...
<input type="checkbox"/>	3620	Feature	New	Normal	Too slow mCBM reconstruction with TRD2D	Alexandru Bercuci		09.05.2025 09:41	...
<input type="checkbox"/>	3618	Feature	New	Normal	mCBM from master does not work on vae24 (Debian11)	Sergei Zharko		30.04.2025 10:34	...
<input type="checkbox"/>	3613	Development	New	Normal	Use official weekly production for comparison tool	Souvik Chattopadhyay		24.04.2025 11:21	...
<input type="checkbox"/>	3612	Bug	New	Normal	Something strange with the BuildOnlineContainerMR CI stage	Florian Uhlig		23.04.2025 14:23	...
<input type="checkbox"/>	3611	Bug	New	Normal	some discrepancy in the mcbm transport file mcbm_transport.C and mcbm_transport_beam.C	Christian Sturm		24.04.2025 10:45	...
<input type="checkbox"/>	3610	Bug	New	High	[Common productions] Crash in transport	Frédéric Julian Linz		28.04.2025 20:19	...
<input type="checkbox"/>	3609	Feature	New	Normal	Strange features in CbmBbaAlignmentTask	Sergey Gorbunov		24.04.2025 09:36	...
<input type="checkbox"/>	3607	Feature	New	Normal	Software issues with time-based digitization	Frédéric Julian Linz		29.04.2025 14:36	...
<input type="checkbox"/>	3605	Infrastructure	Resolved	Normal	Clang-format issue	Florian Uhlig		07.04.2025 15:19	...
<input type="checkbox"/>	3604	Development	In Progress	Immediate	Cbmroot JUL24_patches installation for common productions	Florian Uhlig	07.04.2025	22.04.2025 12:54	...
<input type="checkbox"/>	3603	Feature	New	Normal	Hypernuclei definition for transport simulation	Frédéric Julian Linz		04.04.2025 09:22	...
<input type="checkbox"/>	3597	Feature	New	Normal	Online Data Challenge	Volker Friese	21.08.2025	14.05.2025 13:46	...

« Zurück **1** 2 3 ... 28 Weiter » (1-25/695) Pro Seite: 25, 50, 100

- Using Redmine as communication platform for developers and users
- Tickets / Planning / Documentation (Wiki) / Fora

Experiences / Issues / Problems

- Hosting and maintain such large a system becomes increasingly difficult and expensive with size.
- The necessity to support legacy code impedes development with modern coding standards.
- Large variety of coding skills among the contributors.
- Difficult to introduce / enforce coding standards and conventions.
- Library dependences become hard to manage (e.g, avoid circular dependences)
- Dependence from external packages introduces maintenance efforts (e.g., when updating).
- Available workforce is < 50% of what you would like to have.

Disclaimer (2)

So much to what we have.

Now come the questions: how shall we proceed? Recommendations, best practices, etc...

Goal is to have a sustainable code base comprising the needs for operating the experiment in some three years from now.

Q1: A Single Repository or Multiple Ones?

Pro single:

- Synergy in maintain the development infrastructure (gitlab, CDash, ...)
- Easier system integration and control of coherence.
- Easier to see interfaces in the IDE.
- No need to define dependencies (interfaces, versions) between projects.

Pro multiple:

- Better separation and organisation of development.
- Easier to introduce different levels of coding standards / conventions.
- Granular decision which external packages to link to.
- Developments in active areas are not decelerated by those in less active areas.
- Better overview and management.

Q2: How to Define Libraries

- When does it make sense to introduce a new, separate library?
- Limiting cases:
 - One single library for everything
 - One library for each compilation unit

Q3: How to Organise Code

- One directory per library?
- Should the directory of a library have subdirectories?
- Should the installation be flat (single include directory) or have the same hierarchy structure as the code base?
- Connected to that: Is it advisable to have non-unique file names? (i.e., the path is needed to qualify the class)
- Example:
 - Non-unique: cbmroot/algo/detectors/sts/Hitfinder.cxx
 - Unique: Cbmroot/reco/detector/sts/CbmStsAlgoFindHits.cxx

Q4: Namespaces

- When does it make sense to introduce a (sub-)namespace?
- Should a namespace be connected to a library? Or a directory?

Q4: Namespaces

- When does it make sense to introduce a (sub-)namespace?
- Should a namespace be connected to a library? Or a directory?

Q5: Coding Conventions

- How much sense do coding conventions make?
- Don't they solely rely on taste? E.g., CbmStsAlgoAnaCluster vs cbm_sts_algo_ana_cluster
- If I choose to have some: How do I enforce them?
- Too strict regulations decelerate the development process.
- Taste changes with time.... But adopting to new standards is a huge enterprise...

Q6: Anything Else...

Do you have more aspects / recommendation which I did not touch here?

