# Belle II Software Roadmap

Frank Meier

2025 Belle II Summer Workshop 23 June 2025





#### Introduction

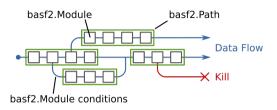
- ▶ Belle II Analysis Software Framework (basf2) divided into 35 packages
  - One for each subdetector: arich, cdc, ecl klm, pxd, svd, top, vxd
  - Core packages: framework, reconstruction, tracking
  - Data taking: daq, hlt, rawdata, trg
  - Data quality: alignment, calibration, dqm, validation
  - Data storage tables: mdst, skim
  - MC: decfiles, generators, geometry, simulation, structure
  - Background: background, beast, ir
  - Offline analysis: analysis, b2bii, mva
  - Documentation / outreach: display, masterclass, online\_book
- Each package has one or two librarians (total number of librarians: 40)
- ► Code written in C++
- One shared library created per package and installed in top-level lib directory
- Build system based on SCons

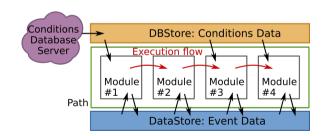
		ecceeeccee		
	0000		000000	
0000	0000		000000	
0000	6666		eeeee	
0000	6666		00000	
6666	6666		00000	
000000				
	000000000000000000000000000000000000000			
88 88 8888888 88 88 88 88	0000 11 00 00 11 0000000 11		e 22 22 e 22 22 22 22 e 22 22 e 22 22	
Copyright(C) (See "basf2 Re Ve	license" fe lease light- rsion light-	mbers of the or more info 2505-deimos 2505-deimos	e Belle II ormation.)	
BELLE2 LOCAL	DIR:			/el9/releases/light-2505-dei
BELLE2_SUBDIR BELLE2_EXTERN BELLE2_ARCH:	ALS_VERSION:	Linux x86 (	64	
Default globa Kernel versio Python versio	l tags: n:	('main_2024 5.14.0-427 3.11.9	4-12-20',) .42.1.el9_	4.x86_64



#### Modular structure

- ► Linear arrangement of C++ modules in a path
- Core functions of modules
  - initialize
  - beginRun
  - event
  - endRun
  - terminate
- Python steering script to set up path

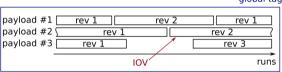






# Conditions Database

- Storage place of additional data needed to interpret and analyze the data that can change over time, *e.g.*, detector configuration or calibration constants
- ▶ Payloads: binary objects (usually ROOT files) identified by name and revision number
- Each payload has defined intervals of validity (iov), *i.e.*, experiment and run range
- ▶ Globaltag: collection of payloads and iovs for a certain dataset, identified by unique name





- Once prepared globaltag is immutable and cannot be modified any further to ensure reproducibility of analyses
- Different processing iterations use different globaltags
- Globaltag of reconstruction stored in metadata and automatically applied
- Chain of globaltags possible

Frank Meier (Duke University)



## Data-taking

- basf2 runs on 12 high-level trigger nodes (14 from fall 2025)
- ZeroMQ
  - ► Acts like concurrency framework while looking like an embeddable networking library
  - Sockets that carry atomic messages across various transports like in-process, inter-process, TCP, and multicast
  - Fast
  - Asynchronous I/O model  $\rightarrow$  scalable to multi-core operation



## Reconstruction chain

- RootInputModule
- Geometry
- Clustering of calorimeter
- Clustering of pixel and silicon vertex detectors
- ► Track finding
- Track fitting
- Track extrapolation
- ► Track-cluster matching
- ► Software trigger
- Post-filter tracking
- PID
- RootOutputModule



# Tracking

- ► Pattern recognition / track finding
  - Finding hits belonging to the same charged particle
  - SectorMaps

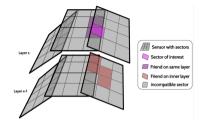
```
Segment Network
```

Cellular Automaton to find longest paths

- SVD-only pattern recognition and CDC-only pattern recognition
- Inter-detector track finding via Combinatorial Kalman Filter

► Track fit

- Extracting track parameters from fit to collection of hits
- Deterministic Annealing Filter (DAF)
- Currently use GenFit package (DOI:10.5281/zenodo.10301439)
- Track refining
  - Flip and Refit
- More details in sphinx documentation







# Tracking

- ► Pattern recognition / track finding
  - Finding hits belonging to the same charged particle
  - SectorMaps

Segment Network

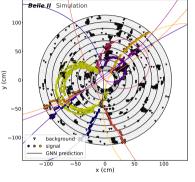
 $\downarrow$ 

Cellular Automaton to find longest paths

- SVD-only pattern recognition and CDC-only pattern recognition
- Inter-detector track finding via Combinatorial Kalman Filter

Track fit

- Extracting track parameters from fit to collection of hits
- Deterministic Annealing Filter (DAF)
- Currently use GenFit package (DOI:10.5281/zenodo.10301439)
- Track refining
  - Flip and Refit
- More details in sphinx documentation
- ► End-to-End Multi-Track Reconstruction using Graph Neural Networks at Belle II Comput. Softw. Big Sci. 9, 6 (2025)



CAT Finder

#### Unit-tests

- ► First layer of software validation
- ▶ Run full test suite for each commit of open merge requests and each merge into main or release branch
- Unit-tests intended to catch non-trivial dependencies and implications of code changes
- Currently about 1000 unit-tests of C++ code using GoogleTest
  - Check basic functionality of modules, return values of functions and variables
- About 350 additional python tests
  - Make sure that standard scripts do not crash
  - Compare output of certain scripts with reference expectation, e.g., for mdst backward compatibility
- Running all tests (in 16 parallel processes) takes 15-20 min







# Nightly validation

- Run once per day (night)
- Workflow of nightly validation
  - 1. Generate smallish samples
  - 2. Run validation scripts
  - 3. Create output histograms
  - 4. Comparison with reference
    - Calculate p value of histogram compatibility
    - Calculate performance numbers, *e.g.*, width of distribution
- Plots of various software releases uploaded to web server
- Email notifications sent out to assigned contacts

# Nightly validation



Belle II Validation		
Validation	Package: klm	
Revisions	r uokuge. kini	
pular/prebuilt/default:	Script Files	
+latest rel/night v	(nightly:2025-06-20)(release-09-00-01)	
reference	Failed Scripts	
upgrade-2023-08-15 3-11-29 00:58 JST   ec00bd1bc	No failed scripts	
hightly-2025-06-20 5-06-20 11:03 JST   2208808473	Finished Scripts	
nightly-2025-06-19 5-06-19 11:09 JST   154cf70501	Skipped Scripts	
nightly-2025-06-18 5-06-18 13:35 JST   d6ae8d7453		
elease-09-00-01	Result File: BKLMMuon	
5-05-14 00:57 JST   7cb950ad5	Downloads: nightly-2025-06-20 release-09-00-01	
elease-09-00-00	INTERNATIONAL DESCRIPTION DESCRIPTION DESCRIPTION DESCRIPTION DESCRIPTION DESCRIPTION DESCRIPTION DESCRIPTION DE	259/p to HLANDER
5-01-18 18:08 JST   9174ede41		1.
5-06-03 15:29 JST   1a4402ad7		un ha
elease-08-03-00		
-02-28 23:58 JST   ddcba7343		
lease-08-02-06		And a second
-01-28 09:25 JST   70689!24e	IN the of 2 Addition of Cold Television Cold T	
-04-10 21:39 JST   3495/9e/2		
lease-08-01-10		
-09-23 21:04 JST   5a1f5864b		
elease-08-00-10		
-08-09 15:16 JST   fa7445a93		
elease-06-02-00		
3-12-28 06:27 JST   e2aa77080		
elease-06-01-16 4-08-07 03:24 JST   ff75dc742		



## Monitoring

- Nightly build run with different configurations (debug, intel, clang)
- Many resource checks (memory consumption, execution time, output file size)
  - Summarize build warnings, cppcheck, doxygen check, dependency check, geometry overlap check
- History plots of warning and error counters as well as resource usage

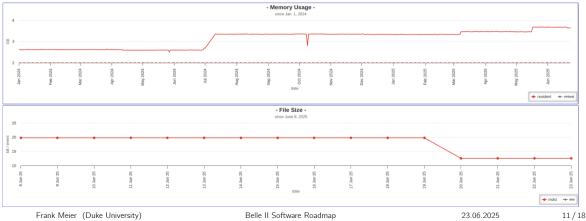
heregenetikset kinge	ator Bullion Marring Charles Code Analysis Mathematics	BUTTON BATTON									
Personal of do	wispnant build										_
Printy Access, part	A REAL OF PERSON AND A REAL PROPERTY OF THE REAL PR										002
Package data	in .										
Participa -	Line la	A TAK BATE	A Drive Bure	A 10/044	A Tang Tank	A Garmen	O Talkada	y bunny	a treas	and the second	4.747
signed	Table Mile	4.	4.0	4-	4	4-		4 -	A memory of	4-	4-
analysis.		4.00	40	4.1	A North	4.00	O 141208 MAY 20	4 11	A Territor 2 D Marry 7	4.00	4.00
#101	1.49 9416	4.0	4.0	4 -	A 1011	4.0	a	¥ =	4.0	4.0	4.0
10.000		4-	4-	4-	4-	4-	4.000	4 -	4-	4-	4-
Independ	Nets Data	4.00	40	4.00	4.11	40	4.000	4.00	A Merings 3	40	40
and a	President Co.	4.0	4.0	4.0	A 1017	4.0	a	4.0	4.0	4.0	4.0
and a second sec	Table Mile, New York, Hill, Hill Strandy	4-	4-	4-	4-	4-		4 -	Q may 1	4-	4-
-		4.00	40	4.4	A front	4.00	4.000	4.4	A Veringe A	4.4	4 .
**		A 10110-1	A 10100 B	A many N	A 1014	A 10100 10	4	4 -	A Terring I Q Horn MR	4.0	10
ALC: 1		4-	4-	4-	4-	-	A	4 -	4-	4-	4-
dially.	Casara De Parte	40	4.0	10	4.	4.4	a	4.0	40	4.0	40
-		4.0	10	4.2	4.2	4.0	4.000	4.2	A Terrar I	4.0	40
		A HANNE OF	A 10110 B	A many at	A 1011	4-	4	4-	A memory of	4-	4-
Approach 1		4.0	4.0	10	A 1011	4.0	Q 141.00 141.0	1.0	A Territor & C Victor 1	4.0	10
generators	Design Devices, Warks Test	4.0	4.0	4.2	4.0	4.0	ANDRO	4.2	A Terrar N	4.0	40
-	Party Line, Spinster,	4-	4-	4-	4-	4-	1	4 -	10	4-	4-
and the second s		4.0	4.0	10	1.0	4-	A 1011 0111	10	A Manage J	4.0	4-
	Faile Tables, Pola Tage Vitible-B	4.0	10	4.2	4.2	4.0	A 10111011	4.2	A Terrar I Q Terra I	4.0	10
	THE ACTOR	4 -	4.0	4 -	4 -	4.0	4.000	4 -	10	4.0	4.0
-		2.	4-	4-	A No. 7	4-	A 111 MIL	4-	A Territor & Q Honey &	4-	4-
materia	factures.	10	4.0	4.2	12	10	a	4.0	10		4.0
100		4.0	10	4.0	4	4.0	A 100 0 100	4.0	10	4.0	10
-	Male has from the	2.	4-	4 -	1 -	4-		4 -	A Territor &	4.0	4-
artist look	Prints Deves		10	4.2	4.8	10		4.0	10	4.0	10
84	Inde Sectors, Orster Reput	4	10	4.2	4	4.0	A 1014 BO T	4.2	A manage of the manage of	4.0	10
1000	Mar Intel	2	4-	4-	4-	4-	A	4-	4-	4-	4-
	Casara Da Paris, Res Das	10	4.0	4.0	10	10		4.0	10		10
and the second se		4.0	10	4.2	4.5	4.0		4.5	10	4.0	10
-	7008.00	4-	4-	4-	4-	4-	4	4-	Q margin	4-	4-
-	Deca Laste, Sales Briauly	4 .	4.0	4.4	4.4	4.4	0.001	4.	4.0	4.00	4.00
and an and an	Las Pillonia	10	10	4.2	12	4.0		4.2	10		10
10	Lapters, the secar	4-	4-	-	4-	4-		4-	4-	-	1-
	tare for	4	4.4	4.4	4.	44	4.000	4 -	A Terring 2	4.4	4.00
Telling .	Orable Brand, Kills Fan, Ballers Ballers	4 1	10	4.2	4.5	4.0	A NOTE AND	4.2	The rate of the late of the la	4.0	10

Frank Meier (Duke University)



# Monitoring

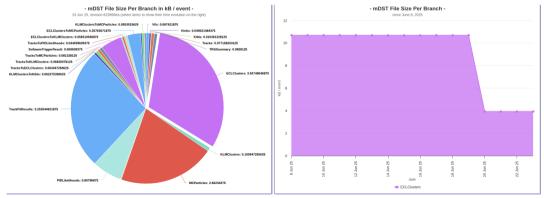
- ▶ Nightly build run with different configurations (debug, intel, clang)
- ► Many resource checks (memory consumption, execution time, output file size)
  - Summarize build warnings, cppcheck, doxygen check, dependency check, geometry overlap check
- ► History plots of warning and error counters as well as resource usage





## Monitoring

- ▶ Nightly build run with different configurations (debug, intel, clang)
- Many resource checks (memory consumption, execution time, output file size)
  - Summarize build warnings, cppcheck, doxygen check, dependency check, geometry overlap check
- History plots of warning and error counters as well as resource usage



Frank Meier (Duke University)



#### Documentation

- Good documentation crucial to (recruiting) process of software development and maintenance
- Documentation publicly available at https://software.belle2.org/
- Sphinx
  - Use reStructuredText
  - Use sphinx's autodoc feature to conveniently create documentation based on python's docstrings
- Doxygen for C++ documentation
- Tests for (almost) all packages to ensure that everything is documented

i I	=	₹ C <b>C</b>
$\mathcal{B}$	3. Beginners' Tutorials	
≩basf2€	This online textbook aims to help new Belle II members to get started with the software by following series of hands-on lessons.	through a
ech CTT1 + K	If you need to cite this online book, please use the following citation: Belle II Collaboration, The Belle II Software Online Book. [Online]. Available: <your_urt_lor_pkis.< td=""><td>page&gt;.</td></your_urt_lor_pkis.<>	page>.
it's New Nation and Setup	[∲ тр	
mend Line Tools ¥	Just as there are many versions of the Befel II software, there are many versions of this documentation to match it. After all, if a new feature is added in our software, we also want to h documentation for it. You can charge your version by clicking on Other versions.	
al Core Modules	If you are a new to all of this, we recommend you to select the recommended release version	
nis 🗸 🗸	(release-xx-xx-(recommended)) in the above list).	
· · · ·	You can also take a sneak peek at the most recent version of the documentation by selecting th	
ground Package	You can also sake a sneak peek at the most recent version or the documentation by selecting the development version. However not all of the code examples might work for you yet.	
oration 🗸	Contraspondent and the second of the cost of an open market of you per	
iy Files 🗸 🗸	The earliest release version which contains this online book is release-05-01-12.	
Delle II Evert Display		
t Generators 🗸 🗸	A Warning	
s for Validation of the	If you change the version to an earlier version than the current one, some pages (also this page	
Trigger	might not exist. If you read the documentation of a newer version than the software that you use	
(K <sup>®</sup> and Muon →	particular the development version), you might not yet be able to use some of the features show	
package	🕈 Tip	
v v	If you get stuck or have any guestions to the online book material, the #starterkit-workshop cha	
oretruction	our chat is full of nice people who can provide help. However this is not the place for specific or	
Asten v	detailed questions about your own analysis.	
	The first lesson of this book - Collaborative Tools, - will show several other places where you ca	
• ×	help later on. It also includes a quick tips section about the chat: Asking a question in #starterki	
ř	workshop.	
dng 🗸 🖌		
	3.1. Welcome!	
s for Physics Validation of ware	3.1.1. Collaborative Tools.	
	3.2. Fundamentals	
ng training 🗸 🗸	3.2.1. Introduction	
mare development V	3.2.2. Data Taking	
r to document your v th Sphinx	3.2.2. Data Taking 3.2.3. Simulation: The Monte Carlo	
ware publications		
and poloications	3.2.4. Reconstruction	
	3.2.5. Analysis	
	23.06.2025	12/



## Release policy

- Major releases
  - Once a year (feature freeze for release-10 in one week)
  - Very thorough validation
  - Contains all software changes that are merged to the main branch
- Minor releases
  - Frequency: one to two per major release
  - Limited amount of new features, usually for specific purpose
- Patch releases
  - Mostly for bug fixes, especially for data-taking and calibration
  - During data-taking synchronized with maintenance days
- ► Light releases
  - Every two months
  - For introduction of new offline data analysis features
  - Contain only framework, mdst, mva, analysis, skim, geometry, online\_book, and b2bii packages
  - No unpacking or digitization  $\Rightarrow$  only mdst and udst can be processed
  - Currently named after celestial objects: aldebaran, betelgeuse, ceres, deimos

Frank Meier (Duke University)



## Supported environments

- ▶ basf2 meant to work on any recent 64bit Linux system but only tested and binaries provided for
  - Enterprise Linux 8 or CentOS 8
  - Enterprise Linux 9 or AlmaLinux 9
  - Ubuntu 22.04
  - Ubuntu 24.04
- basf2 distributed via cvmfs
- ARM version under development
- Central Buildbot instance connected via GitLab webhooks to code changes
  ⇒ triggers builds on various workers





#### Externals

- ▶ Versioned set of external software packages used and linked against the Belle II software
- Dependency between packages considered and compatibility guaranteed
- C++ packages like
  - ROOT, XRootD
  - gcc, clang, gdb, cmake, Python
  - boost, Eigen, gsl
  - EvtGen, Geant4, clhep, PYTHIA
  - git, cppcheck, doxygen
- Includes patches
- Python packages like pandas, matplotlib, torch, tensorflow, jupyter, ...
- Source files uploaded to web server to never lose availability





#### Tools

- ► Collection of scripts to prepare environment for execution of Belle II software
- b2setup
  - Setting environment variables
- b2code-create, b2code-style-check, b2code-style-fix, b2code-clean
  - Creating local directory for core software development and fixing style issues
- b2install-prepare, b2install-release, b2install-externals, b2install-data
  - ▶ Installing pre-compiled software versions or example data on local machine
- b2analysis-create, b2analysis-get, b2analysis-update
  - Creating local directory for development of analysis code including preparation of build system and addition of repository to git



## Links & License

- ► basf2 source code
  - Internal GitLab repository at https://gitlab.desy.de/belle2/software/basf2
- basf2 links against defined set of third-party libraries called externals
  - Internal GitLab repository at https://gitlab.desy.de/belle2/software/externals
- Repository with scripts to install and set up basf2 called tools
  - Internal GitLab repository at https://gitlab.desy.de/belle2/software/tools
- Repository with script for version managing (recommended releases and global tags)
  - Internal GitLab repository at https://gitlab.desy.de/belle2/software/versioning
- ▶ LGPL (GNU Lesser General Public License) version 3 or later
  - ► Header in each file:

basf2 (Belle II Analysis Software Framework) Author: The Belle II Collaboration

See git log for contributors and copyright holders. This file is licensed under LGPL-3.0, see LICENSE.md.



## Links & License

- ► basf2 source code
  - Publicly available at https://github.com/belle2/basf2
- basf2 links against defined set of third-party libraries called externals
  - Publicly available at https://github.com/belle2/externals
- Repository with scripts to install and set up basf2 called tools
  - Publicly available at https://github.com/belle2/tools
- ▶ Repository with script for version managing (recommended releases and global tags)
  - Publicly available at https://github.com/belle2/versioning
- ▶ LGPL (GNU Lesser General Public License) version 3 or later
  - ► Header in each file:

basf2 (Belle II Analysis Software Framework) Author: The Belle II Collaboration

See git log for contributors and copyright holders. This file is licensed under LGPL-3.0, see LICENSE.md.

Frank Meier (Duke University)



#### Conclusion

- Belle II software = C++ code with python interface
- Serial execution of dynamically loaded modules to process collection of events
- Conditions Database stores settings and calibration constants
- Unit-test ensure stability of software
- Nightly validation
- Documentation via sphinx and doxygen
- ▶ Please cite Comput. Softw. Big Sci. 3, 1 (2019) and DOI:10.5281/zenodo.5574115 in Belle II papers



#### Conclusion

- ► Belle II software = C++ code with python interface
- Serial execution of dynamically loaded modules to process collection of events
- Conditions Database stores settings and calibration constants
- Unit-test ensure stability of software
- Nightly validation
- Documentation via sphinx and doxygen
- ▶ Please cite Comput. Softw. Big Sci. 3, 1 (2019) and DOI:10.5281/zenodo.5574115 in Belle II papers

# Thanks for your attention!