

State of the Union

4th EasyBuild User Meeting

January 30th 2019 - Louvain-la-Neuve (Belgium)

https://users.ugent.be/~kehoste/EasyBuild_20190130_state-of-the-union.pdf

kenneth.hoste@ugent.be

easybuild@lists.ugent.be

<https://easybuilders.github.io/easybuild>

<https://easybuild.readthedocs.io>



GHENT

UNIVERSITY

<https://www.ugent.be/hpc>

<https://www.vscentrum.be>



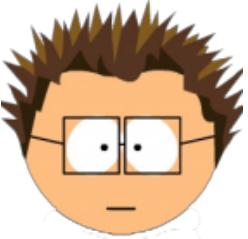
Vlaanderen
is computing

HPC-UGent



- part of central IT department of Ghent University (Belgium)
- centralised scientific computing services, training & support
- for researchers of UGent, industry & knowledge institutes
- core values:
empowerment - centralisation - automation - collaboration
- member of Flemish Supercomputer Centre (VSC)
<https://www.vscentrum.be>
- birth place of EasyBuild (summer 2009)





whoami

kenneth.hoste@ugent.be
[@boegel \(GitHub, IRC, Slack\)](https://github.com/boegel)
[@kehoste \(Twitter\)](https://twitter.com/kehoste)

- Masters & PhD in Computer Science from UGent
- joined HPC-UGent team in October 2010
- main tasks: user support & training, *software installations*
- inherited maintenance of (something that looked like) EasyBuild in 2011
- slowly also became **EasyBuild lead developer & release manager**
- likes family, beer, loud music, FOSS, helping people, dad jokes, stickers, ...
- doesn't like CMake, SCons, Bazel, setuptools, software dependencies, ...

How to make package managers cry



FOSDEM 2018

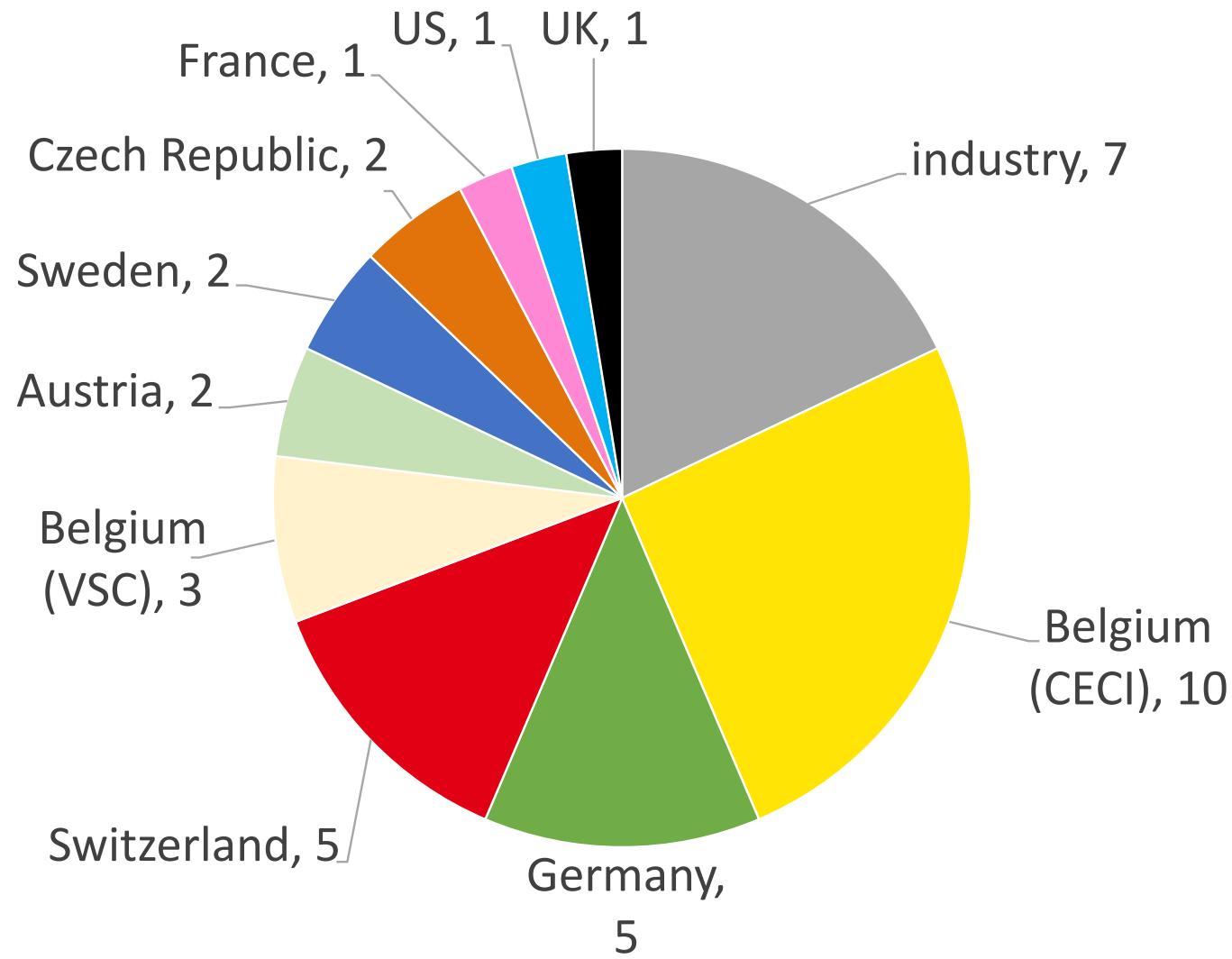
https://archive.fosdem.org/2018/schedule/event/how_to_make_package_managers_cry

<https://www.youtube.com/watch?v=NSemIYagjIU>

- me **venting ~7 years of frustration** with getting scientific software installed
- TensorFlow as main motivator
- **sarcastic tone for dramatic effect** (it worked!)
- lots of feedback (and ideas for an extended version of the talk)
- **clearly hit a nerve...**



4th EasyBuild User Meeting: 39 attendees (10 countries)



*Cenearo
diviti
Gluo
IBM
Red Hat
Sylabs*

Previous user meetings:

- 2016: 19 (HPC-UGent)
- 2017: 35 (JSC)
- 2018: 45 (SurfSARA)

EUM19 agenda (Wednesday, Jan 30th 2019)

- [10.00am] **Welcome** (prof. Bernard Nysten, UCLouvain)
- [10.15am] **EasyBuild: State of the Union** (Kenneth Hoste, HPC-UGent)
- [12.00pm] lunch break
- [1.00pm] **site presentations**: CECI, CSCS, IT4Innovations, Fred Hutch
- [3.00pm] coffee break
- [3.30pm] **Tutorial on creating/contributing easyconfig files** (Kenneth Hoste)
- [4.00pm] **Hands-on session** (+ brewery visits)

EUM19 agenda (Thursday, Jan 31st 2019) [1/2]

- [09.00am] **Collective Knowledge framework** (Grigori Fursin, diviti)
- [09.30am] **OpenHPC introduction** (Adrian Reber, Red Hat)
- [10.00am] **ReFrame, A Regression Testing and Continuous Integration Framework for HPC systems** (Victor Holanda Rusu, CSCS)
- [10.30am] Hands-on session + coffee break
- [11.20am] **Deploying Electronic-Structure Codes with EasyBuild**
(Yann Pouillon, University of Cantabria) **[remote]**
- [12.00pm] lunch break

EUM19 agenda (Thursday, Jan 31st 2019) [2/2]

- [1.00pm] **Singularity 3.0** (Eduardo Arango, Sylabs)
- [2.00pm] Hands-on session (+ coffee break)
- [3.00pm] **Cloud & local software deployment using CVMFS, EasyBuild & Lmod**
(Davide Vanzo, Vanderbilt University) **[remote]**
- [4.00pm] **Lmod/XALT update** (Robert McLay, TACC) **[remote]**
- [4.45pm] **Python at Compute Canada** (Bart Oldeman, ComputeCanada) **[remote]**
- [5.15pm] Hands-on session
- [6.30pm] **Group dinner (sponsored by HPC-UGent)**

EUM19 agenda (Friday, Feb 1st 2019)

- [9.00am] **Modules: automatic and consistent handling of module file dependencies** (Xavier Delaruelle, CEA)
- [9.30am] **Ansible module for EasyBuild** (Olivier Mattelaer, UCLouvain/CISM)
- [10.00am] **Finding a home for EasyBuild in EuroHPC** (Alan O'Cais, JSC)
- [10.30am] (coffee break +) hands-on session (+ lunch break)
- [1.00pm] **Spack: a Package Manager For Scientific Software**
(Massimiliano Culpo, EPFL)
- [1.45pm] (coffee break +) hands-on session (+ brewery visits)

Live streaming & recording of all EUM19 talks



- Live streaming of all talks via the (brand new) **EasyBuild YouTube channel**
<https://www.youtube.com/channel/UCqPyXwACj3sjtOho7m4haVA>
(or https://tiny.cc/easybuild_youtube)
- Talk recordings should be available there shortly after each talk
- Talks from all 4 previous EasyBuild User Meetings are also listed there, together with other talks on EasyBuild and related tools (see playlists)

2nd easybuild User Survey



- anonymous survey via  SurveyMonkey®
- intention was to get better insight in EasyBuild community & usage
- survey invitations sent via EasyBuild mailing list, Slack/IRC + Twitter
- 36 questions, 11 min. to fill out (on average)
- **93 responses** (up from 77 last year)
- assumed to give a fairly relevant view (but results interpreted with a grain of salt)
- (results are discussed out of order compared to order of questions in survey)

The easybuild community



group picture
3rd EasyBuild User Meeting
@ SurfSARA (Jan 2018)

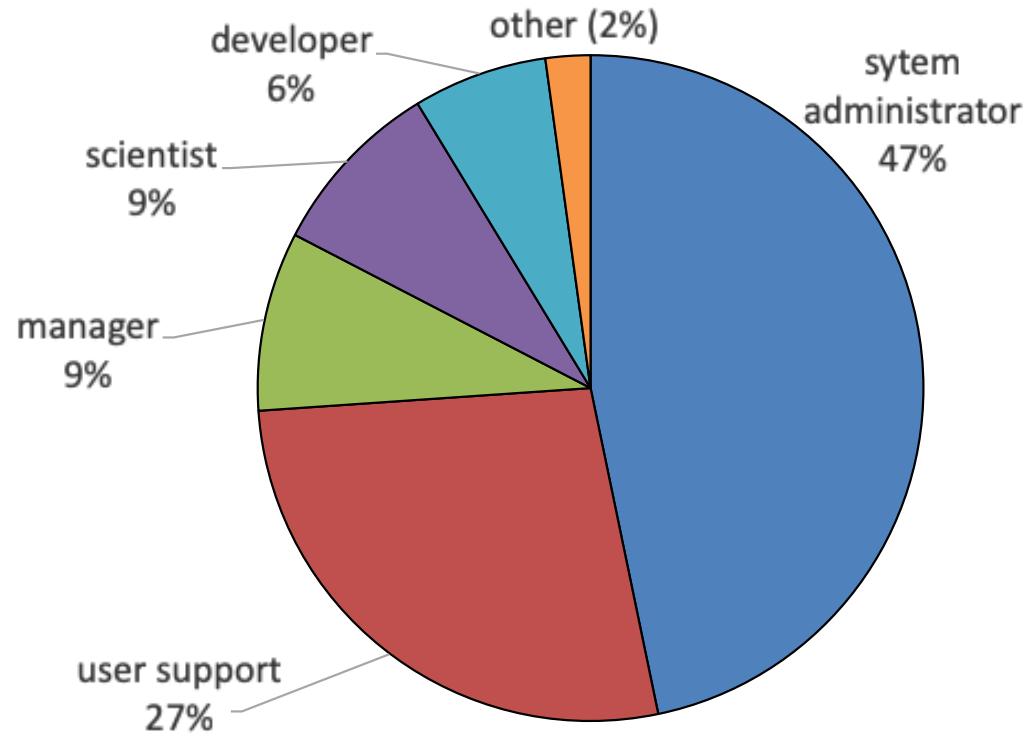


- EasyBuild community has been growing rapidly the last couple of years
- **hundreds of HPC sites and companies worldwide,
incl. JSC, CSCS, SURFsara, Pfizer, ...**
- very welcoming & supportive to newcomers

2nd easybuild User Survey

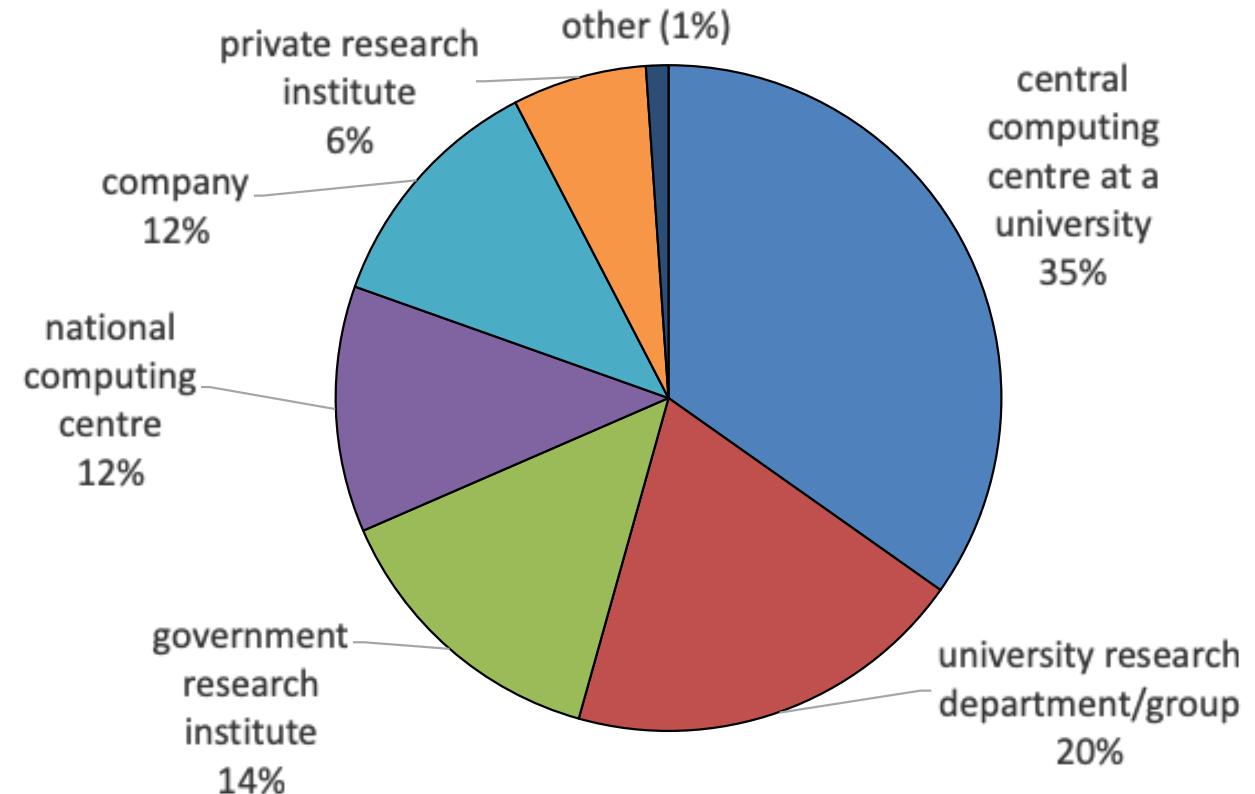


What is your primary profile?



Clear bias towards sysadmins & user support.

What type of organisation do you work for?

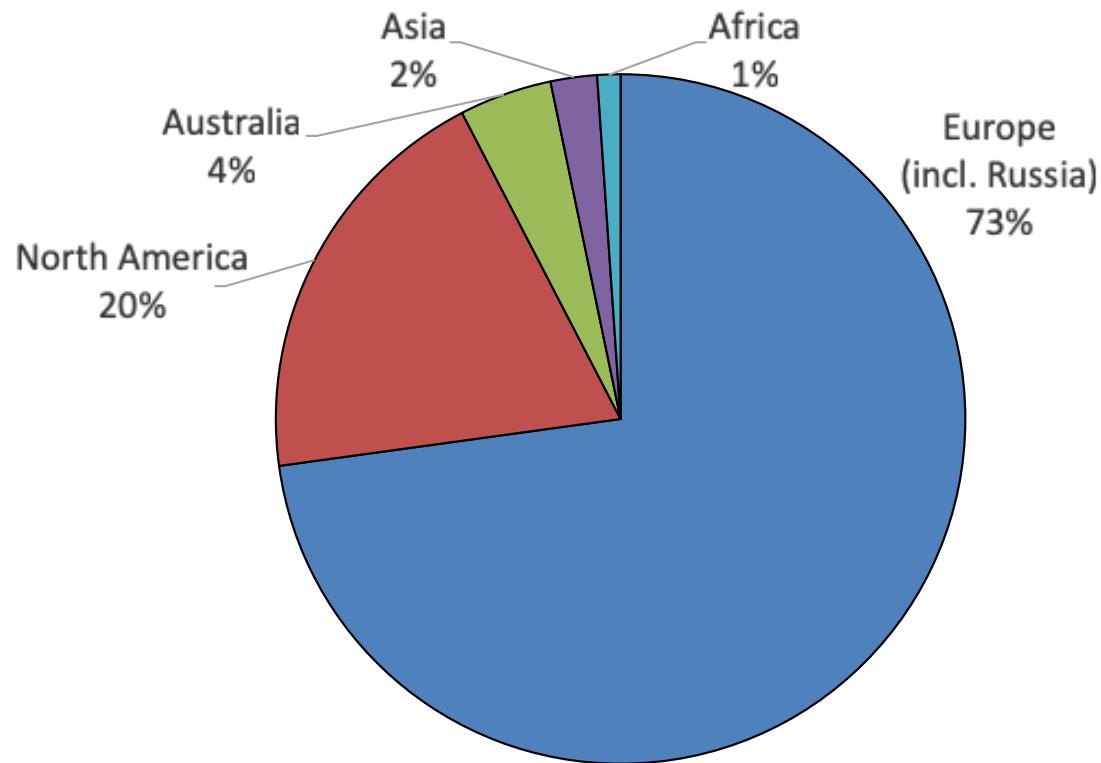


(very similar results compared to previous survey for both)

2nd easybuild User Survey



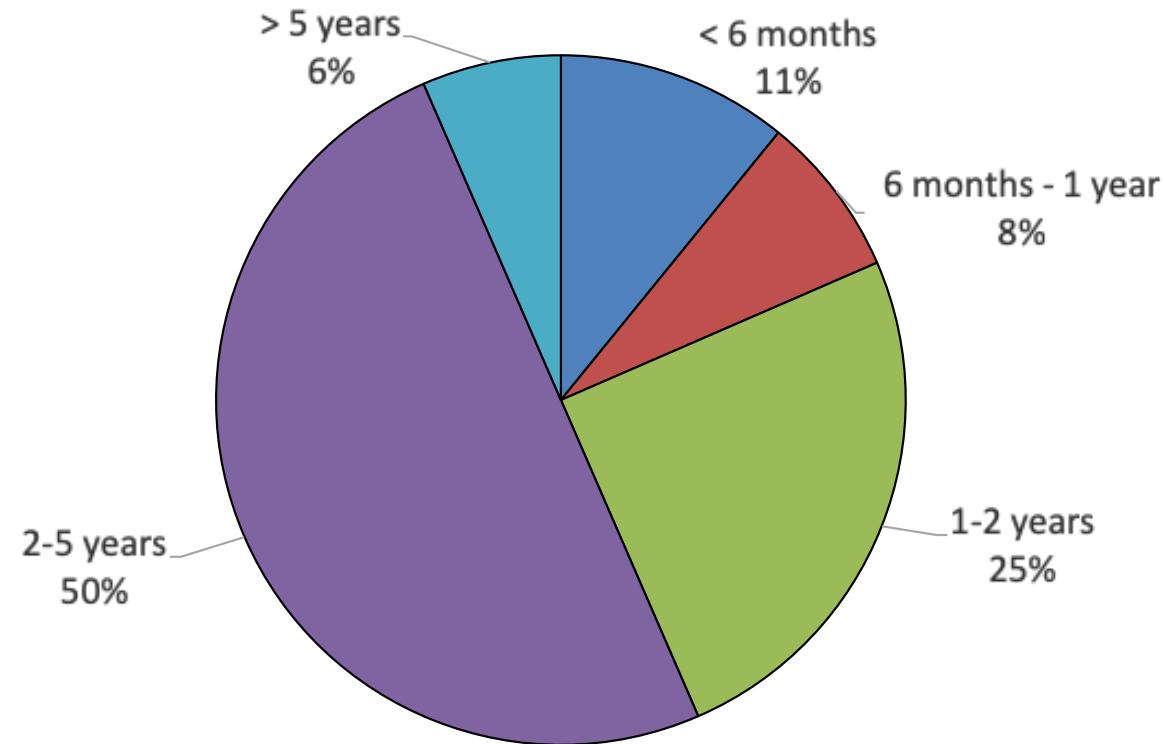
In which part of the world are you located?



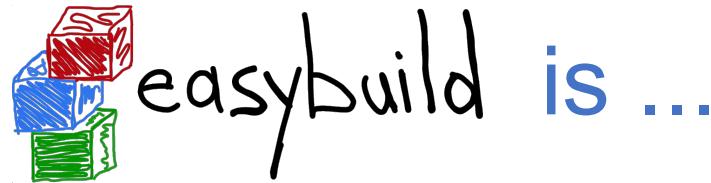
Small shift to more Europe (64% ↗ 73%),
less North America (27% ↘ 20%)

Spack effect?

How long have you been using EasyBuild?



Small shift to "long-term" (> 1 year) users,
"new" users (< 1 year) from 26% to 19%



- not YABT (Yet Another Build Tool)
it does *not* replace build tools like cmake or make; it wraps around them
- not a replacement for your favourite package manager (yum, apt-get, ...)
it leverages some tools & libraries provided by the OS (glibc, OpenSSL, libibverbs, ...)
- not a magic solution to all your (software compilation/installation) problems...
you may still run into compiler errors (if nobody else has already taken care of it)

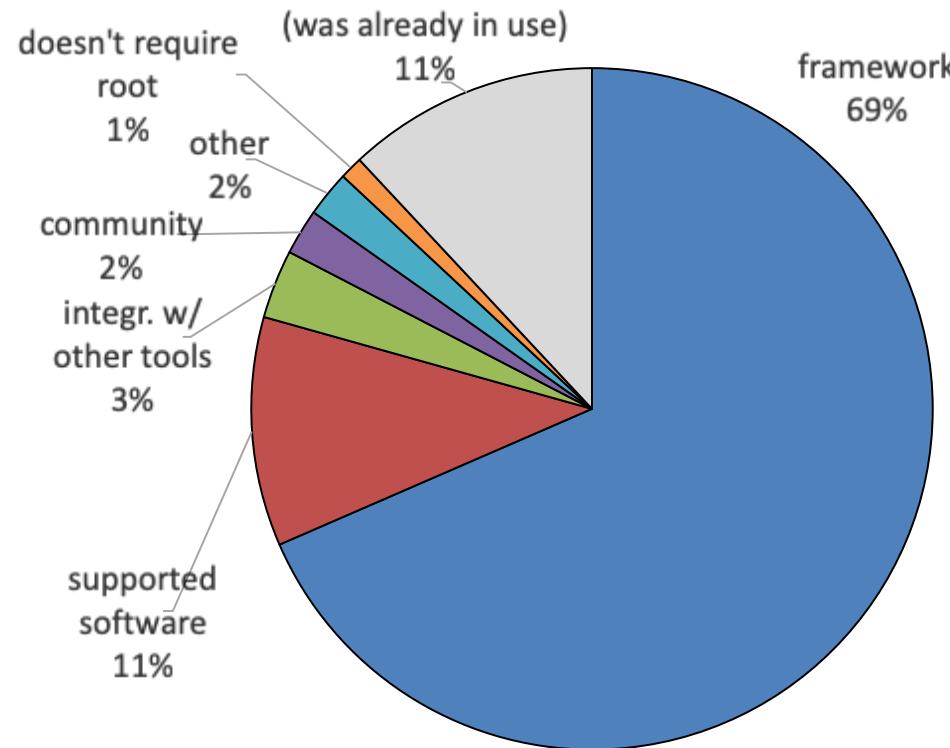


- a **uniform interface** that wraps around software installation procedures
- a huge **time-saver**, by automating tedious/boring/repetitive tasks
- a way to provide a **consistent software stack** to your users
- an **expert system** for software installation on HPC systems
- a **platform for collaboration** with HPC sites worldwide
- a way to let *users* to manage their own **software stack** on HPC systems

2nd easybuild User Survey



What is the main aspect of EasyBuild that convinced you to start using it?

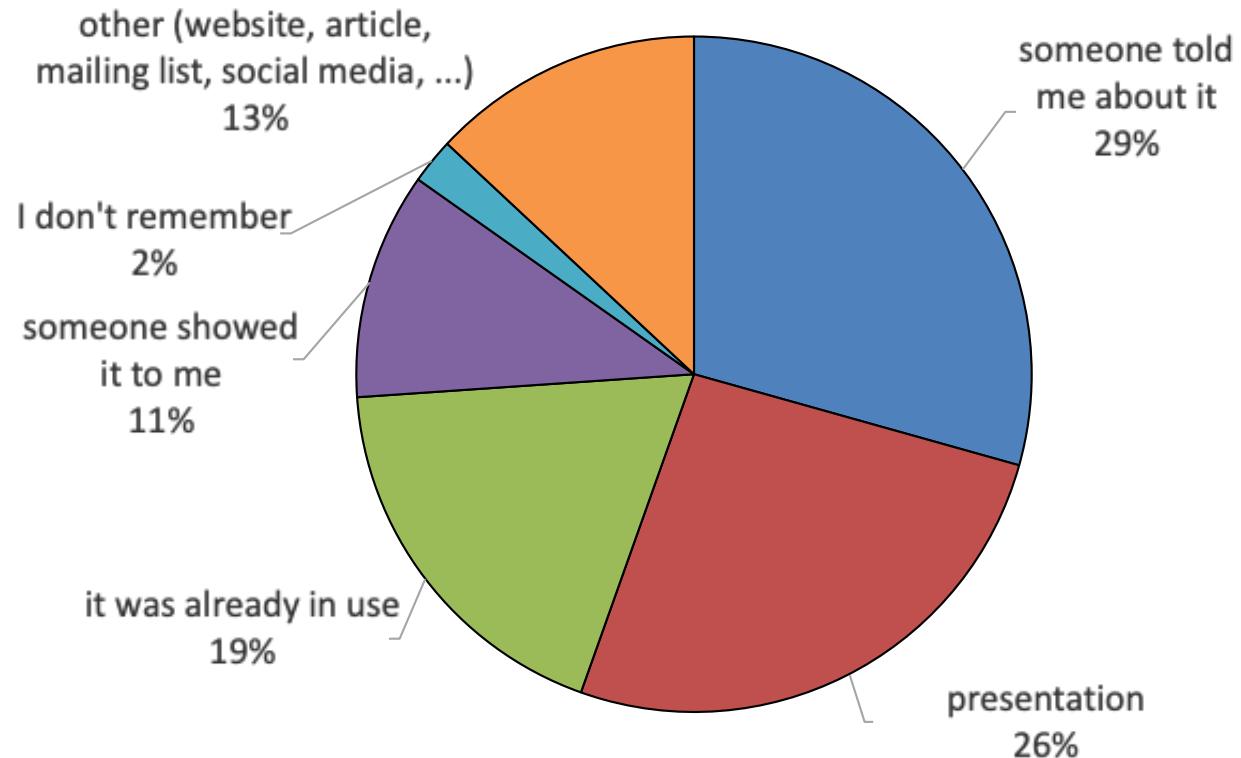


Framework functionality (67% ↪ 69%) and supported software (18% ↪ 11%) remain biggest factors.

2nd easybuild User Survey



How did you first learn about EasyBuild?



Majority is still by word-of-mouth (79% ↗ 67%),
significant increase of "already in use" (4% ↗ 19%)

6 years of (stable) easybuild

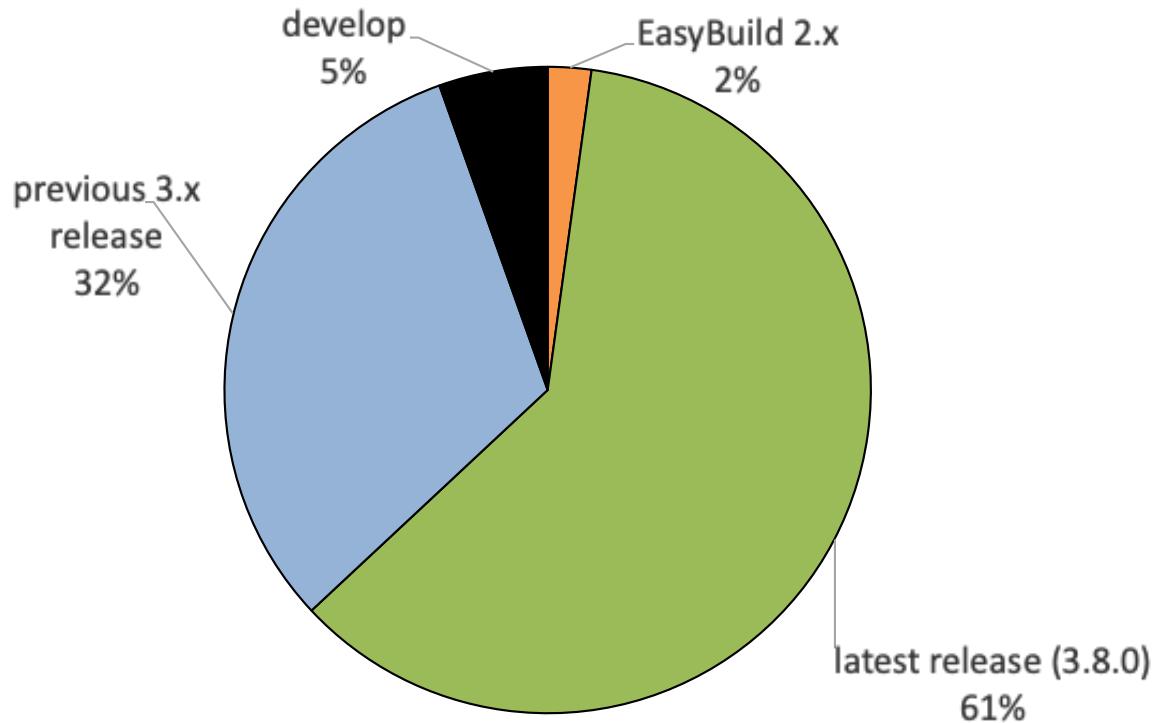
- in-house development at HPC-UGent since summer 2009
- first public release in April 2012 (EasyBuild v0.5)
- **first stable release on November 13th 2012, during SC'12 (EasyBuild v1.0)**
- intention was to get feedback, but gradually a community emerged around it...
- **frequent stable releases** since then (latest: EasyBuild v3.8.1, Jan 29th 2019)
- community-driven development: bug reports, feature requests, contributions



2nd easybuild User Survey

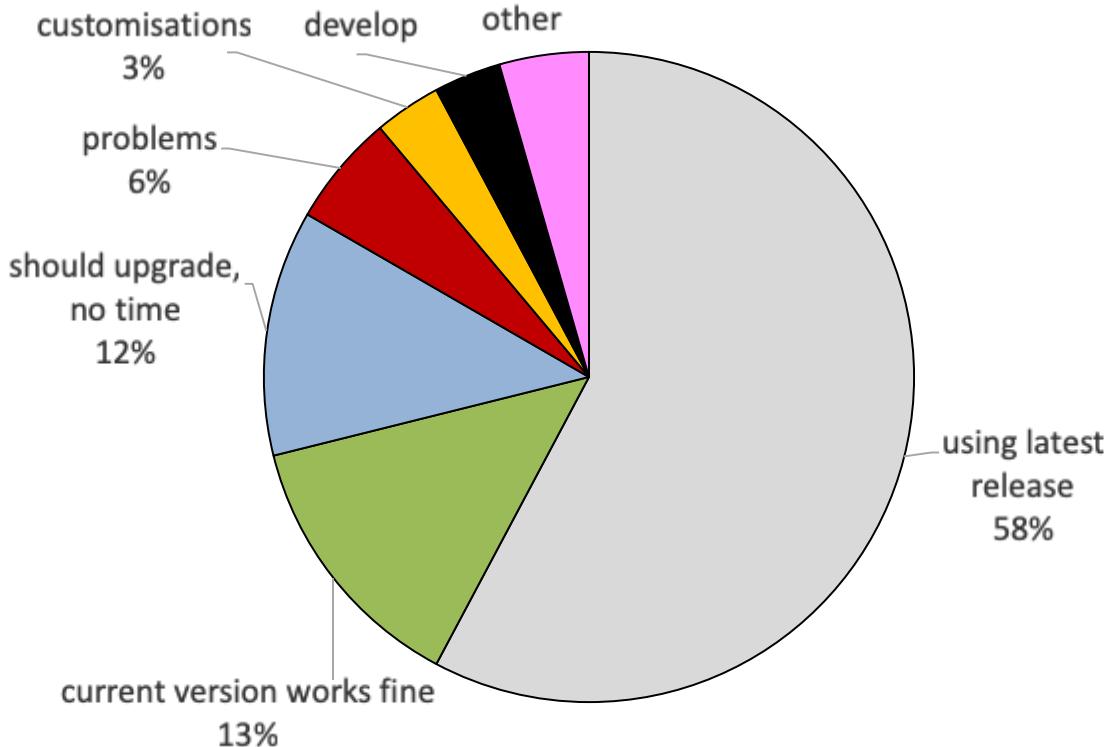


Which EasyBuild version do you use?



95% relies on EasyBuild releases,
majority is up-to-date with most recent version.
(very similar to previous survey)

If you're not using the latest release, why not?

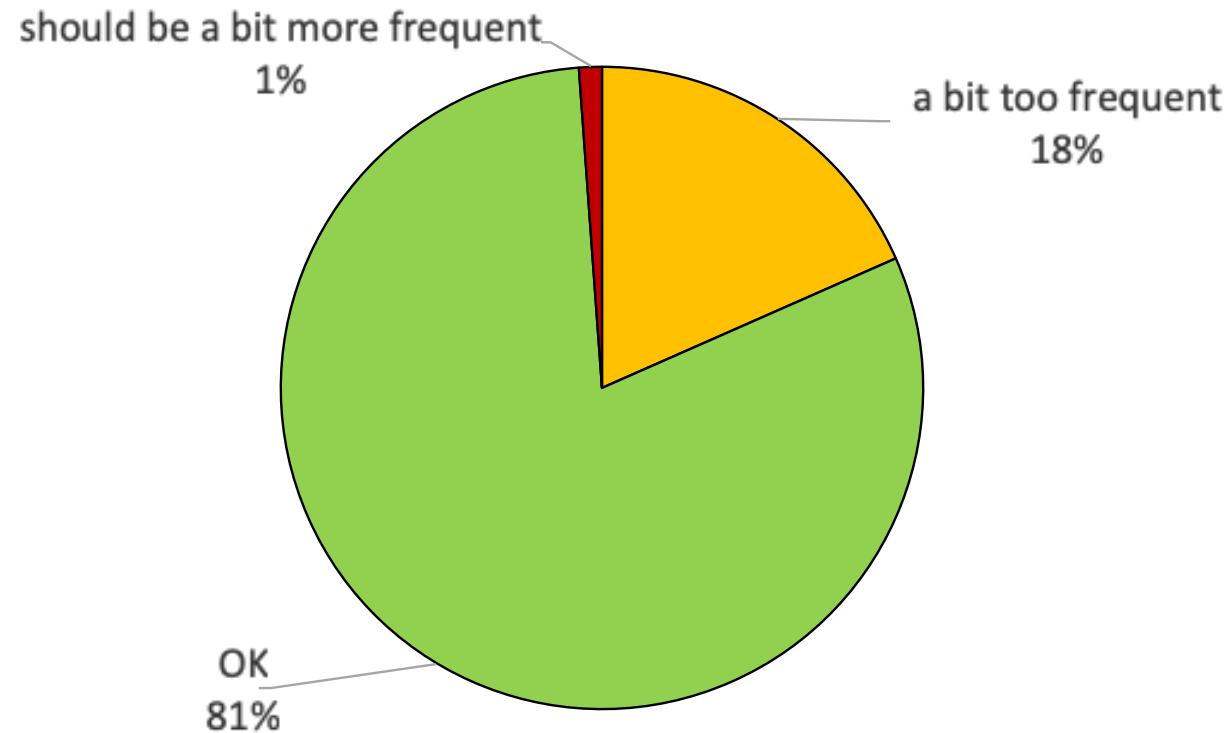


(very similar to previous survey)

2nd easybuild User Survey

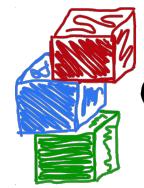


How do you like the frequency of EasyBuild releases?



significant increase in 'OK' compared to previous survey: 65% ↗ 81%

2nd

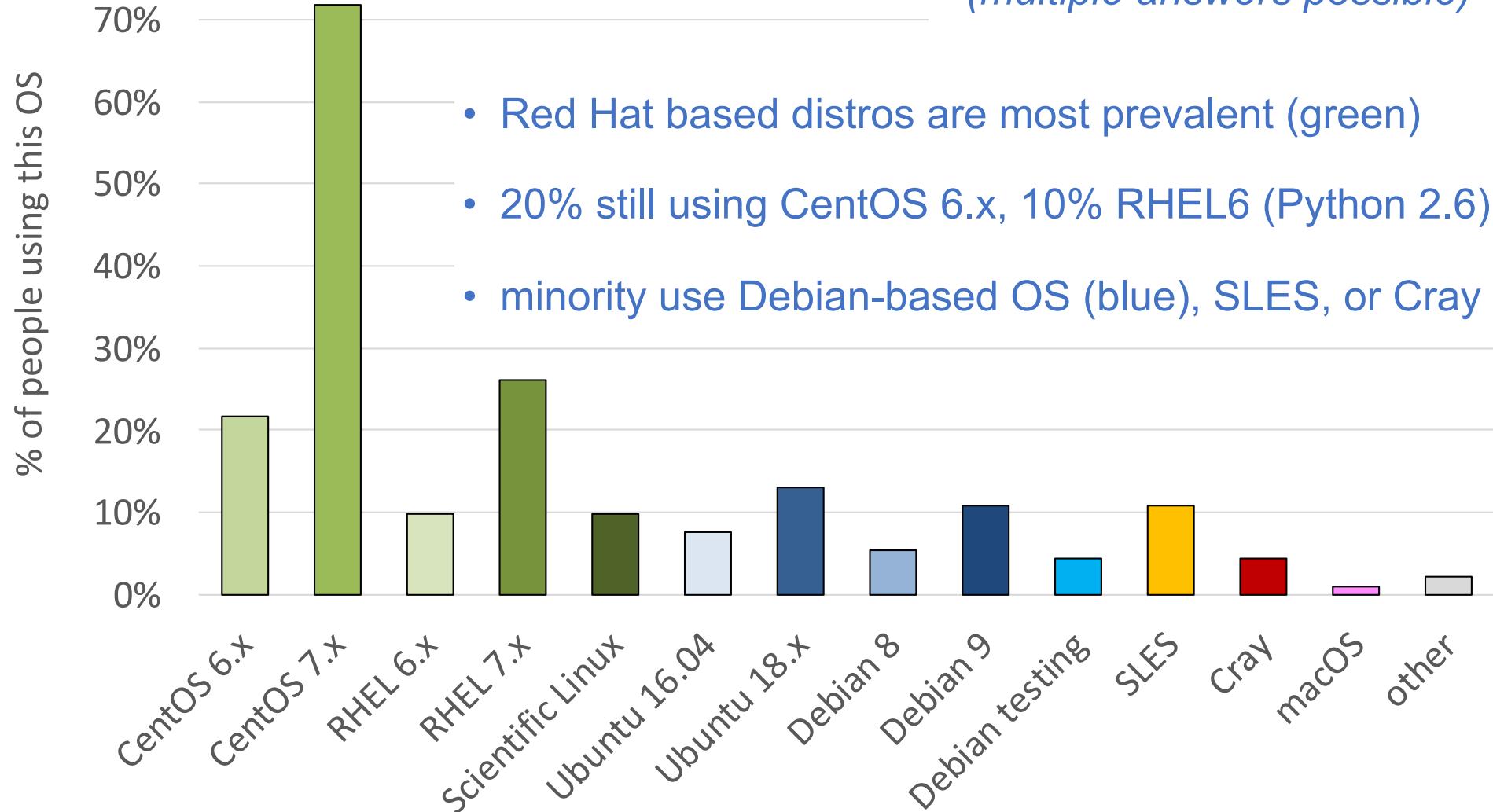


easybuild User Survey



On which operating system(s) do you use EasyBuild (most commonly)?

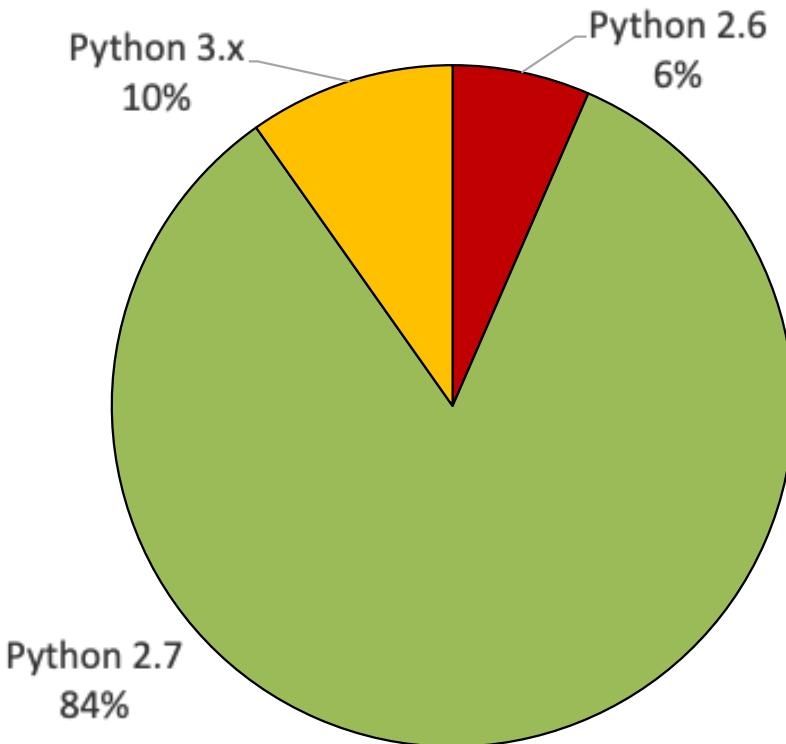
(multiple answers possible)



2nd easybuild User Survey



Which Python version do you usually use to run EasyBuild?

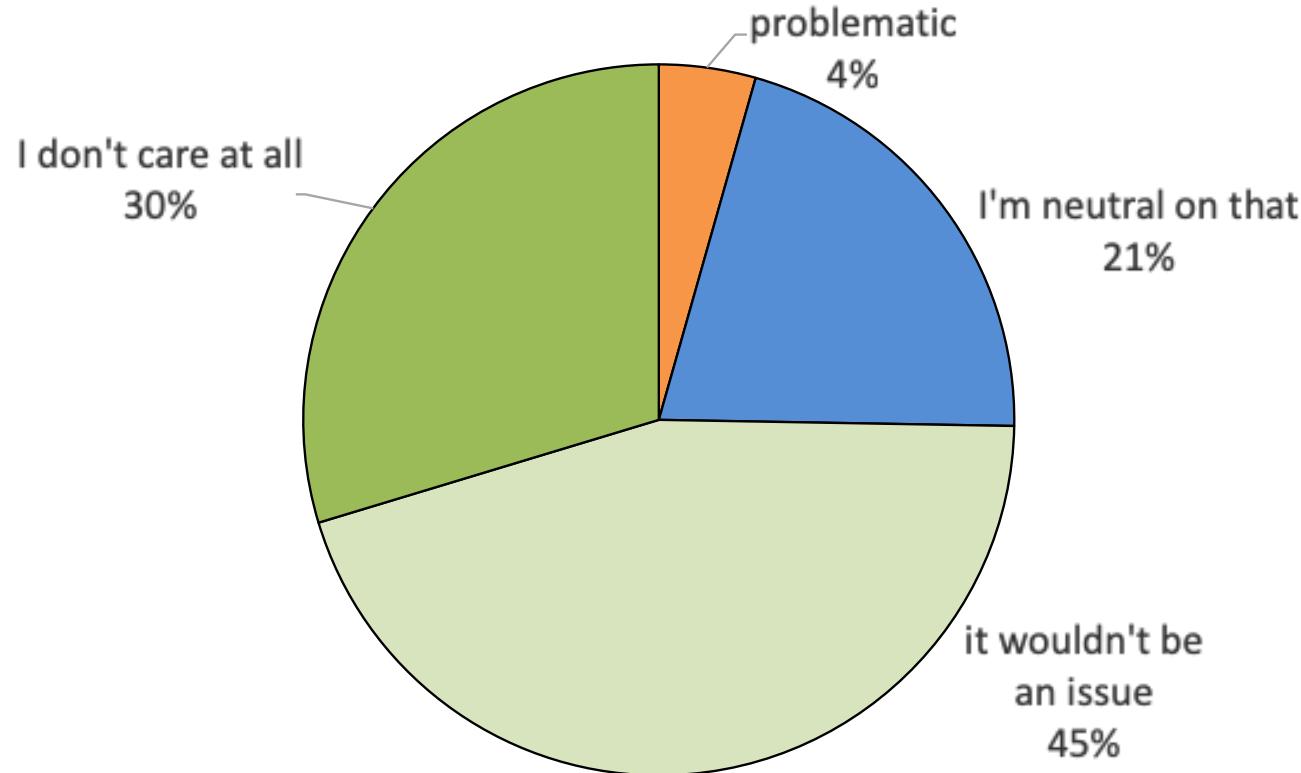


- vast majority uses Python 2.7 (83% ↗ 84%)
- significant decline for Python 2.6 (16% ↘ 6%)
- I want to talk to the people who are already using EasyBuild on top of Python 3...
(more people use Python 3.x than Python 2.6?!)

2nd easybuild User Survey



How troublesome would it be for you if EasyBuild becomes incompatible with Python 2.6?

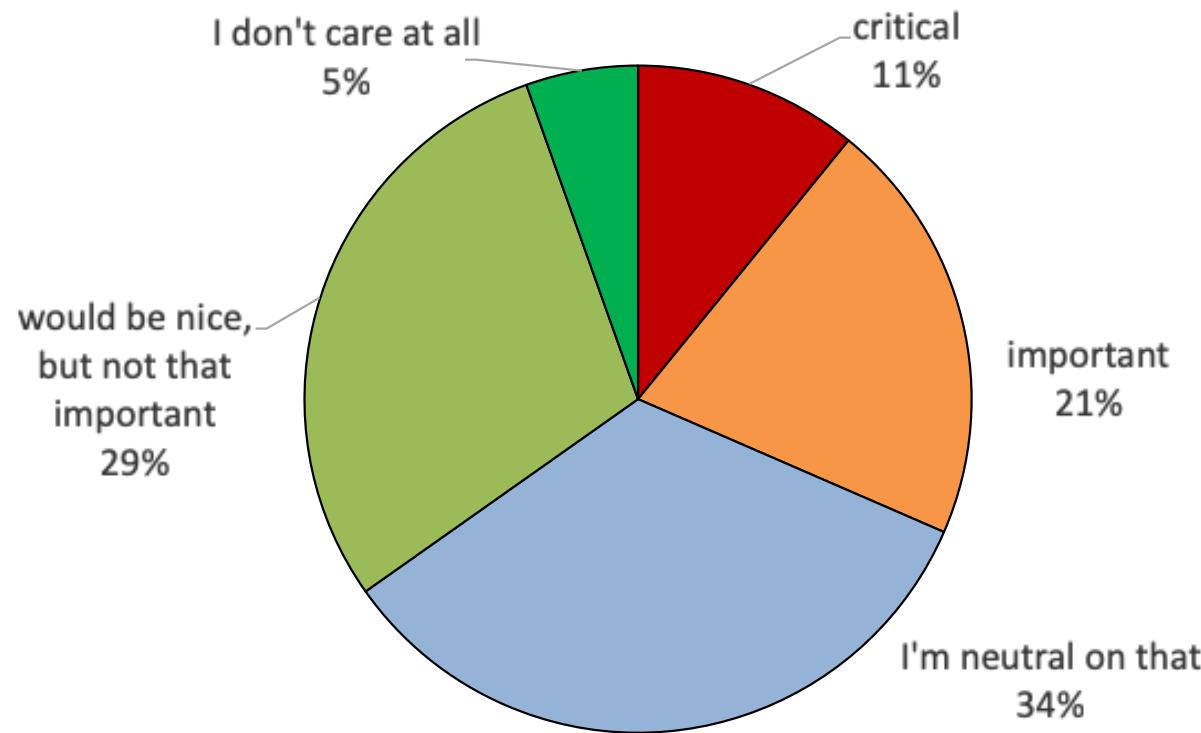


- no longer supporting Python 2.6 would probably not have a high impact
- nobody answered "disastrous"
- declining number of people care about Python 2.6 (13% ↘ 4%)
- some people stuck on CentOS 6 or RHEL 6 have installed Python 2.7

2nd easybuild User Survey



How important is it to you that EasyBuild becomes compatible with Python 3?



- results are quite similar to previous survey
- slightly more people worry about Python 3 compatibility (27% ↗ 32%)
- more people consider it critical (4% ↗ 11%)
- **Python 2 is end-of-life Jan 1st 2020...**
(<https://www.python.org/dev/peps/pep-0373>)
- ... yet RHEL8 will still include python2 (?!)

EasyBuild 4.0

see also <https://github.com/easybuilders/easybuild/issues/447>

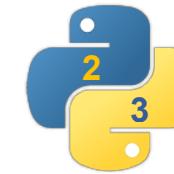


(preliminary) ETA: summer 2019

Goals:

- 1) compatibility with Python 3 (+ Python 2.7, maybe still Python 2.6) A logo consisting of two interlocking snakes, one blue and one yellow, with the number '2' on the blue snake and '3' on the yellow snake.
- 2) no (required) dependencies
 - no vsc-install, vsc-base, setuptools required
- 3) single-tarball releases on PyPI
 - rather than 3+1 releases now (framework, easyblocks, easyconfigs + 'easybuild' meta-package)
 - different GitHub repositories will stay (different rate of development)
- 4) Deprecating the dummy toolchain, adding system toolchain as replacement
- 5) custom easyblock for Open MPI
- 6) switch to using 'pip' as default installer for Python packages (?)

Python 3 support is work-in-progress



- porting effort to allow running EasyBuild with Python 3 already under way
- intention is to work on Python 3 support in parallel with EasyBuild 3.x
- see separate 4.x branch in easybuild-framework repository
 - 4.x branch is kept in sync with current develop
- current status:
 - relevant code from vsc-install + vsc-base ingested into easybuild-framework
 - all Python 3 syntax errors fixed: all Python modules in framework can be imported
 - EasyBuild configuration works
 - tests pass for `easybuild.tools.filetools` module

Python 3 support is work-in-progress



Planning:

- first all framework tests should pass on top of Python 3 (WIP, about halfway there?)
 - step by step, small PRs that are easy to review
- after framework, easyblocks can be ported (minor effort, mostly imports?)
- nothing much to do for easyconfigs?
- testing, testing, testing...
- EasyBuild 4.0.0 release with (experimental?) support for Python 3

Actively helping out with this effort is very much appreciated!

see `python3` label: <https://github.com/easybuilders/easybuild-framework/labels/python3>

Community (common) toolchains



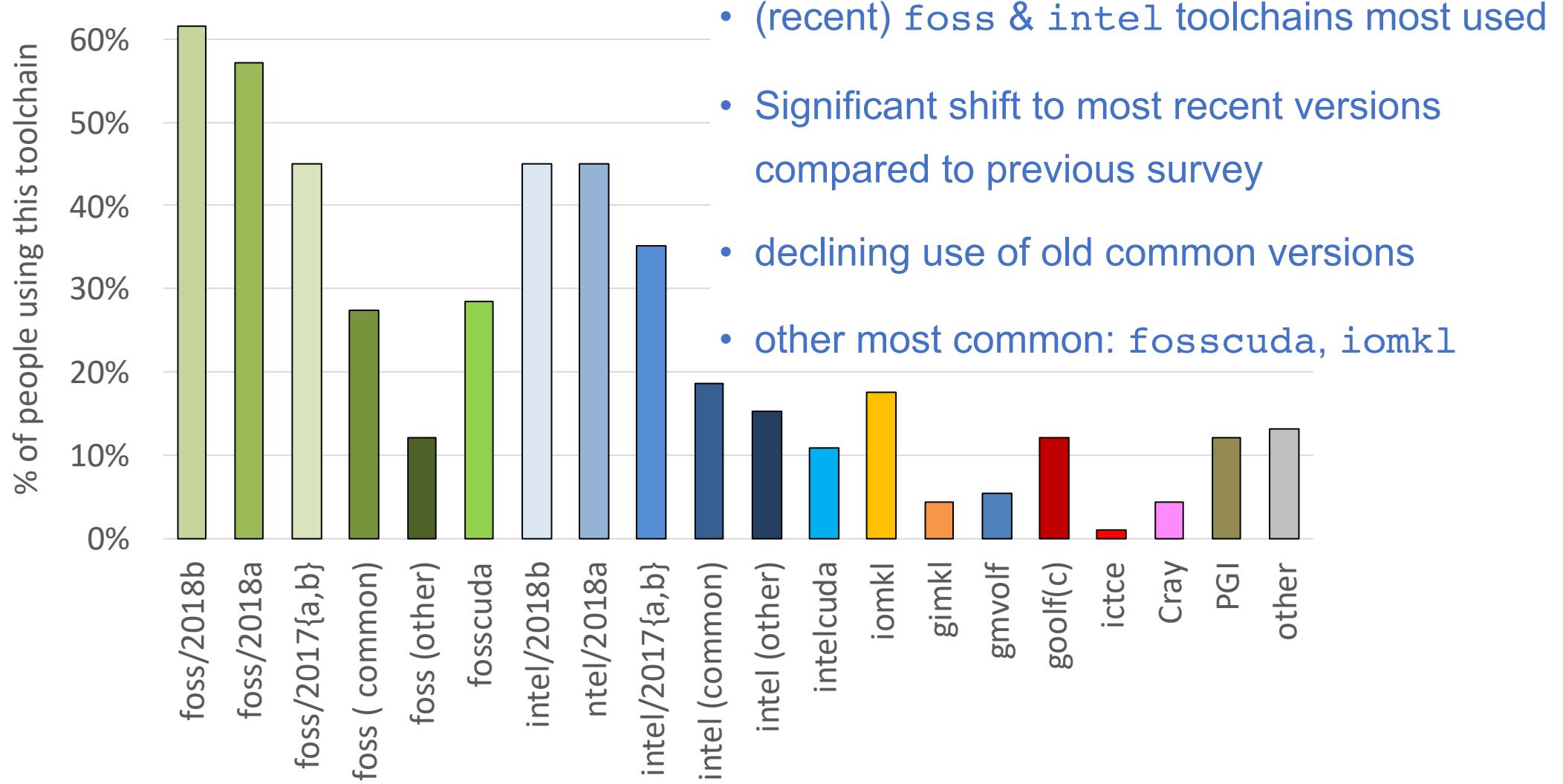
<http://easybuild.readthedocs.io/en/latest/Common-toolchains.html>

- **intel and foss¹ toolchains** are most commonly used in EasyBuild community
- helps to focus efforts of HPC sites using one or both of these toolchains
- updated twice a year, clear versioning scheme: <year>{a,b} (2017b, 2018a, ...)
- latest version:
 - **foss/2019a**
binutils 2.31.1, GCC 8.2, OpenMPI 3.1.3, OpenBLAS 0.3.5 (incl. (Sca)LAPACK), FFTW 3.3.8
 - **intel/2019b**
binutils 2.31.1 + GCC 8.2 as base
Intel compilers 2019.1.144, Intel MPI **2018.4.274**, Intel MKL 2019.1.144

2nd easybuild User Survey

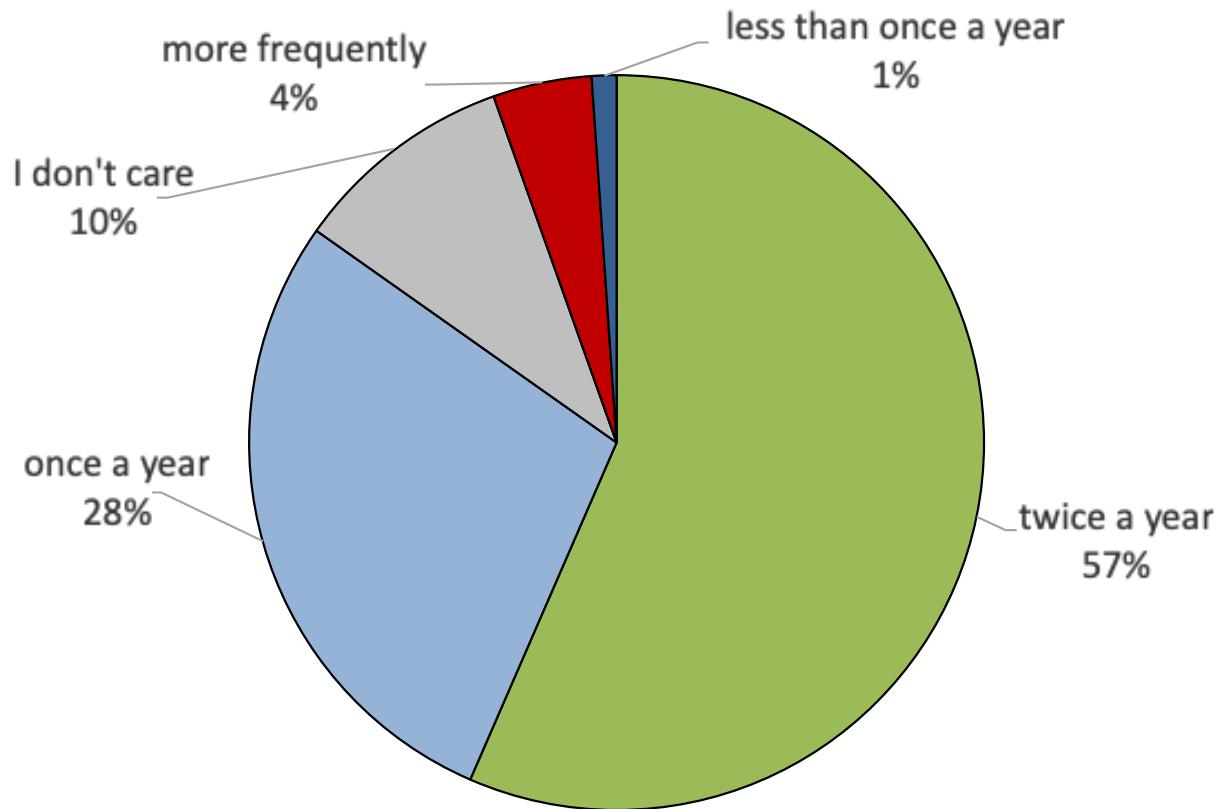


Which (full) toolchain(s) do you use? (check all that apply)



2nd easybuild User Survey

How frequently should the 'foss' and 'intel' common toolchains be updated?

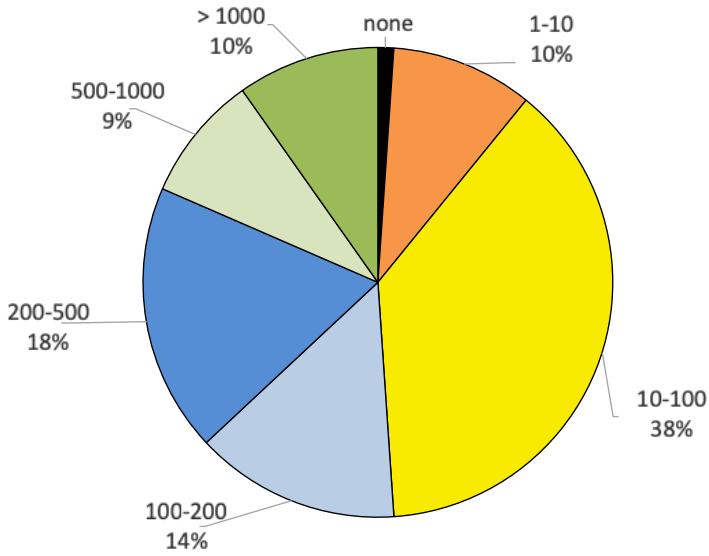


- preference for current update frequency (twice per year) has increased (47% ↗ 57%)
- slight decrease for "once a year" (37% ↘ 28%)
- no changes planned

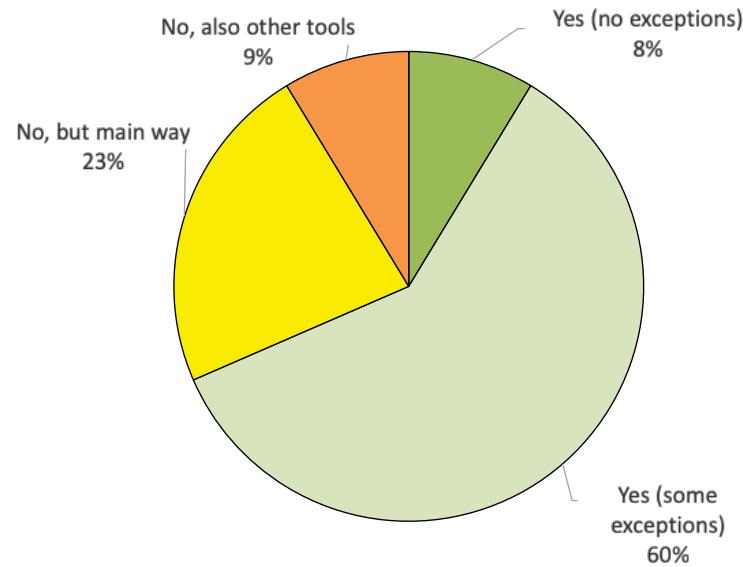
2nd easybuild User Survey



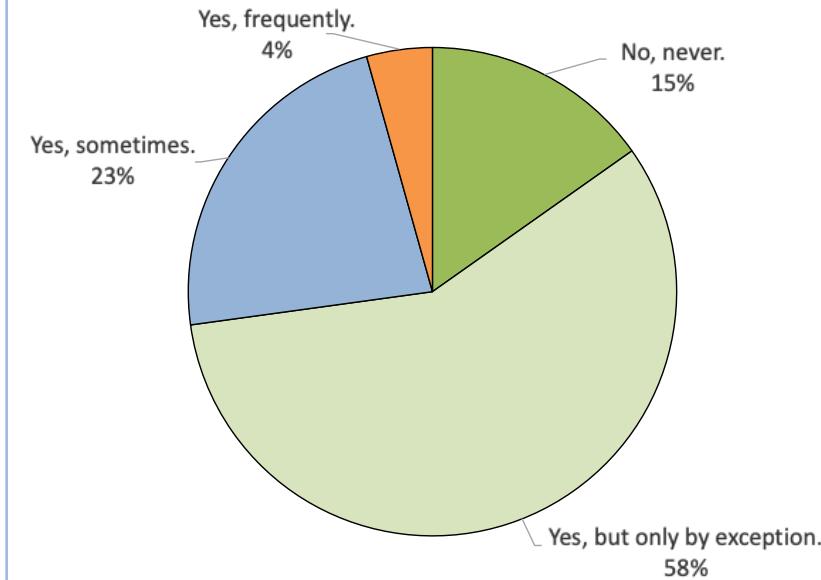
How many software installations did you perform in the last year using EasyBuild?



Is EasyBuild your only way of installing (scientific) software?



Do you still install (scientific) software manually?

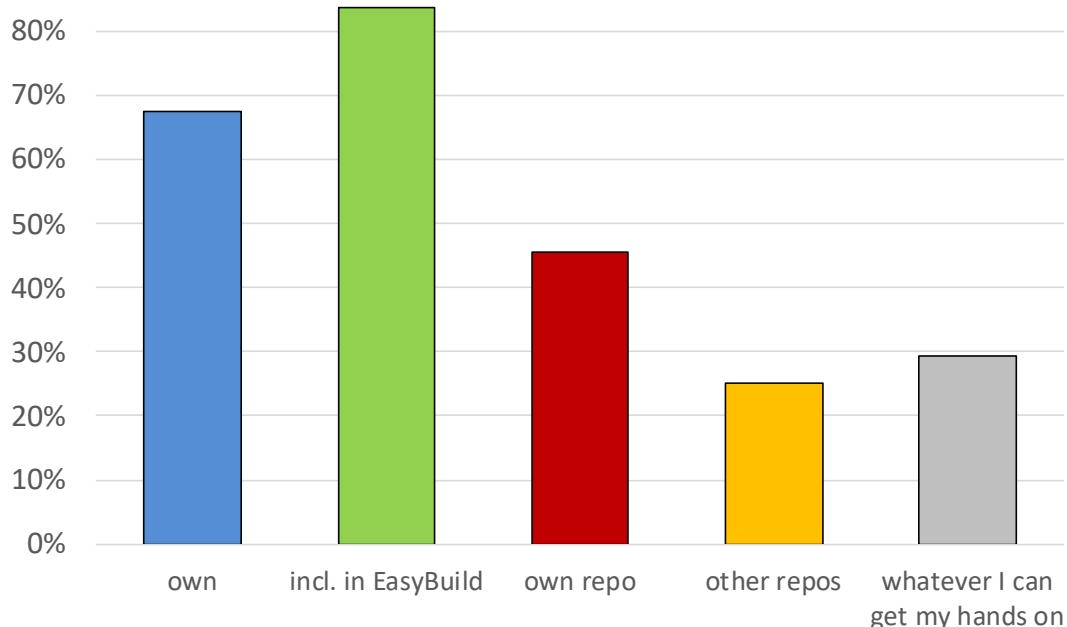


(all fairly similar to previous survey)

2nd easybuild User Survey



Which easyconfig files do you use? (check all that apply)



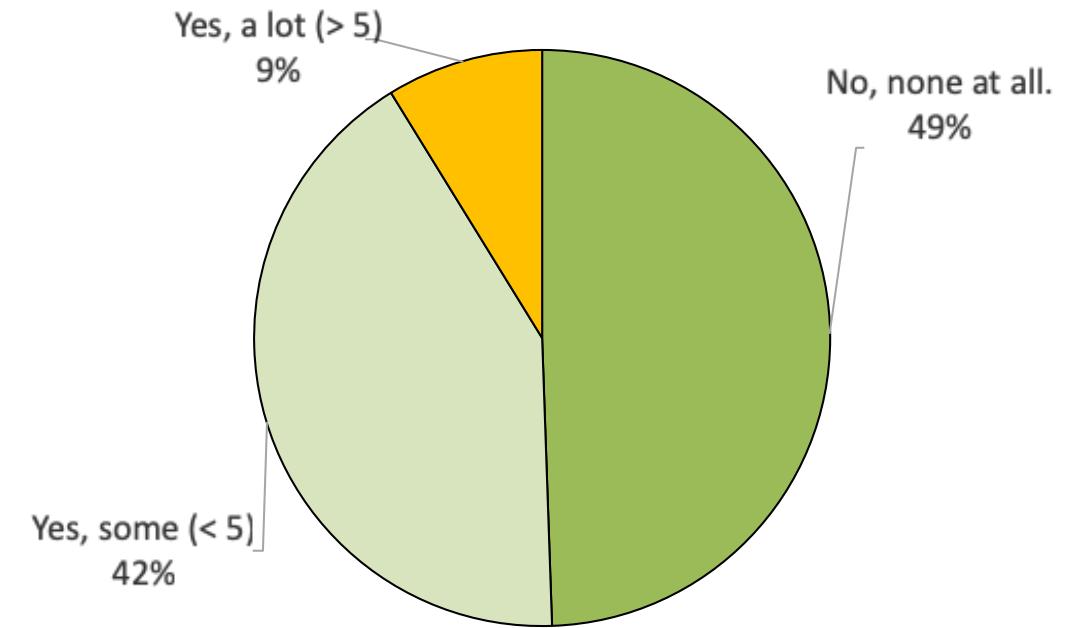
85% uses easyconfigs included with EasyBuild

67% uses own easyconfigs

45% maintains own repo

(very similar to previous survey)

Do you use any custom easyblocks?

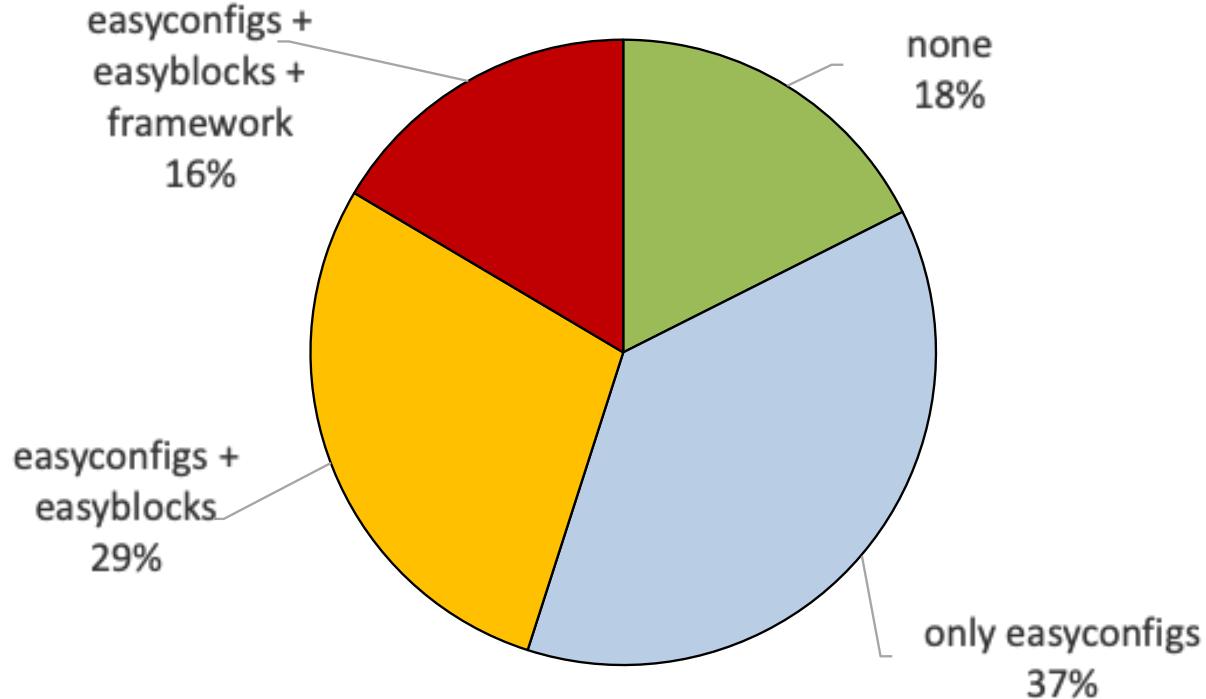


(very similar to previous survey)

2nd easybuild User Survey



Have you made any site-specific customisations to EasyBuild?

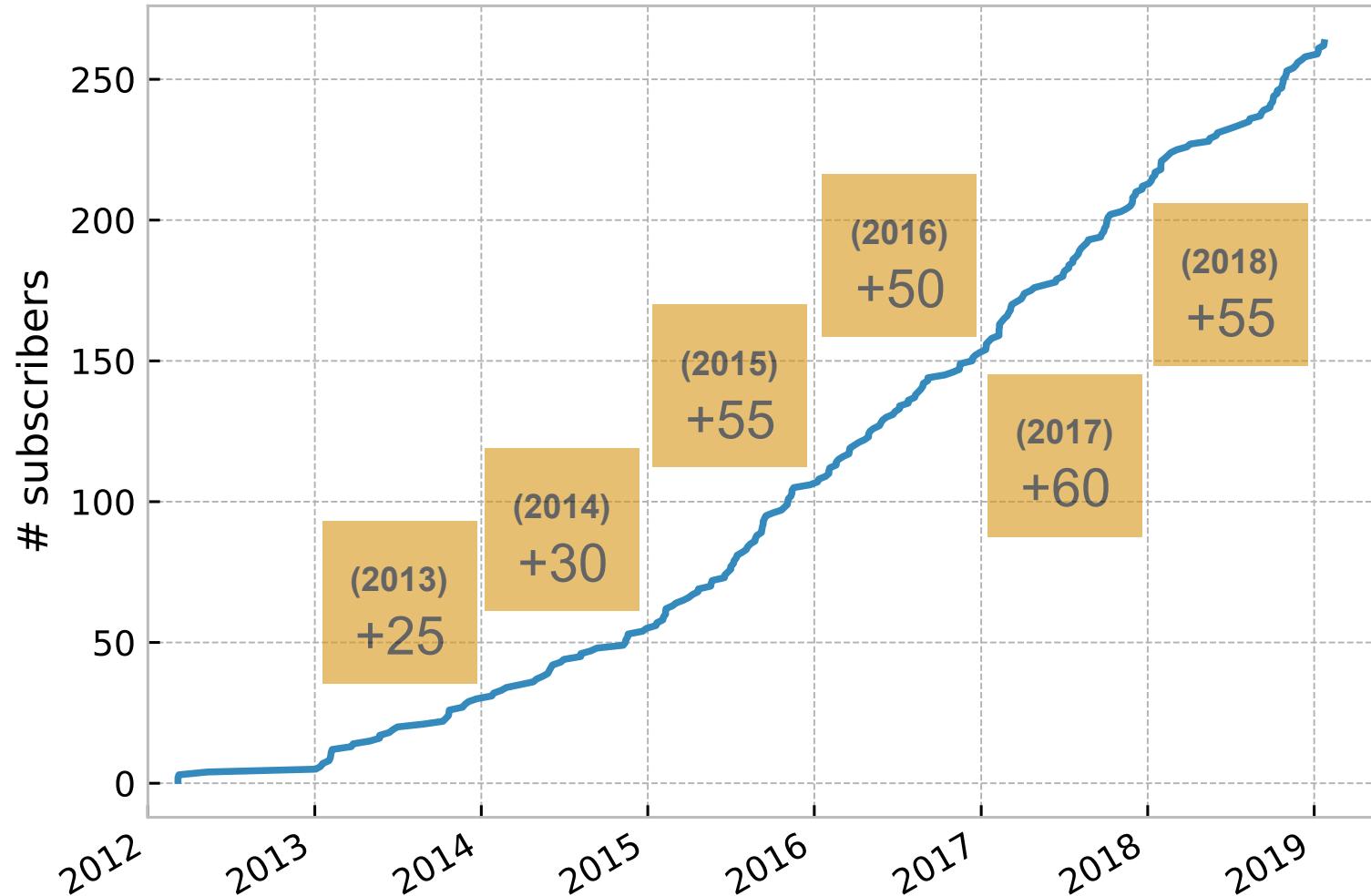


- fairly similar to previous survey
- ~45% indicated having customised code
 - reasons?
 - contributed back?



easybuild mailing list (subscribers)

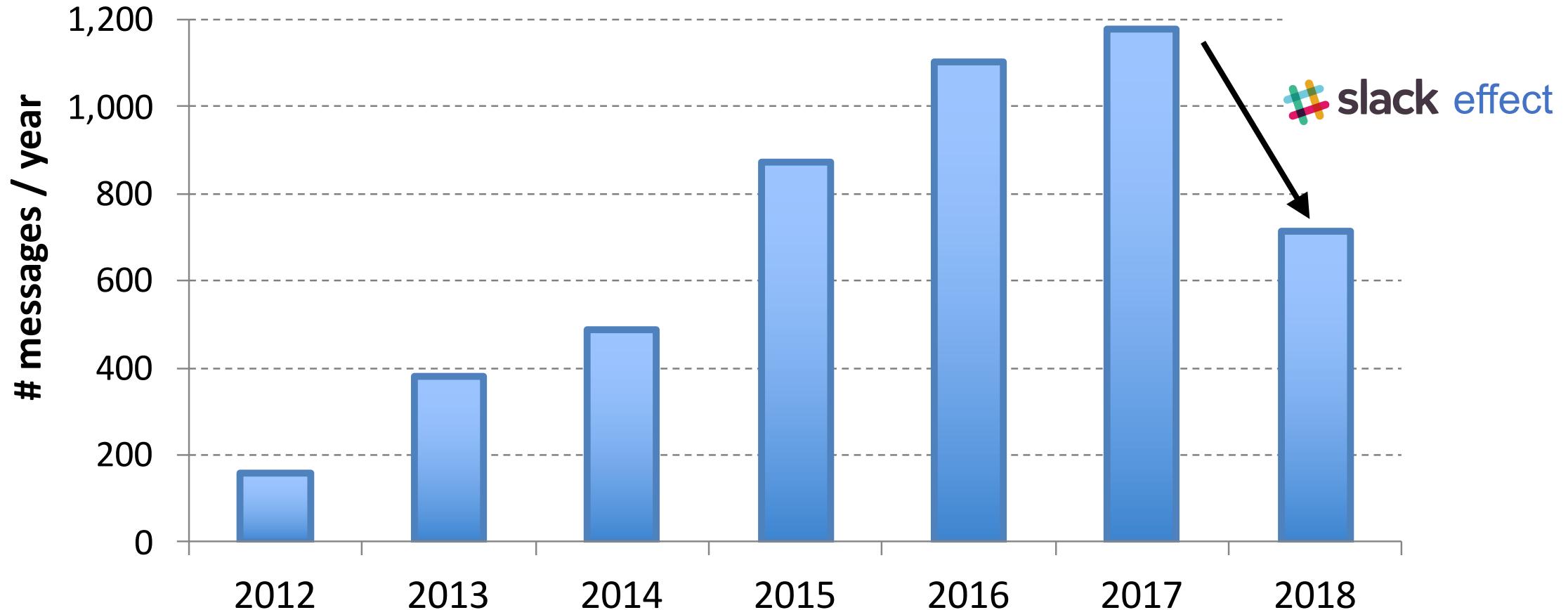
Over 250 people subscribed to EasyBuild mailing list, still growing...





easybuild mailing list (traffic)

Traffic on EasyBuild mailing list peaked in 2017,
most likely due to EasyBuild Slack channel.



easybuild slack (+ IRC) channel

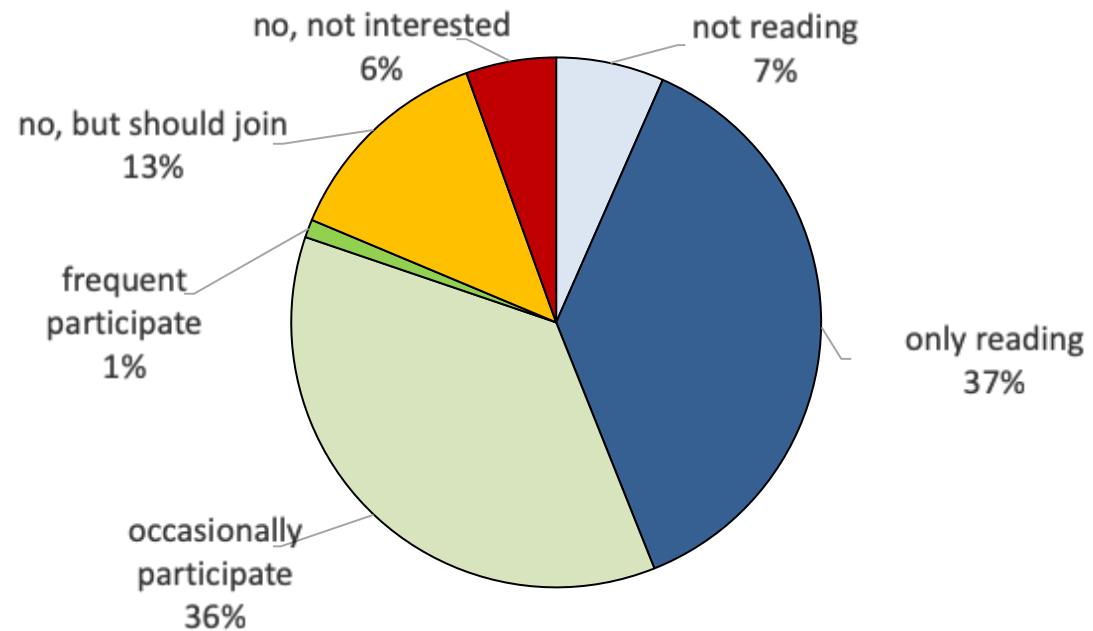
- #general on easybuild.slack.com is linked with #easybuild on IRC (freenode)
- ~130 members on Slack channel (+ about 15 (also) on IRC)



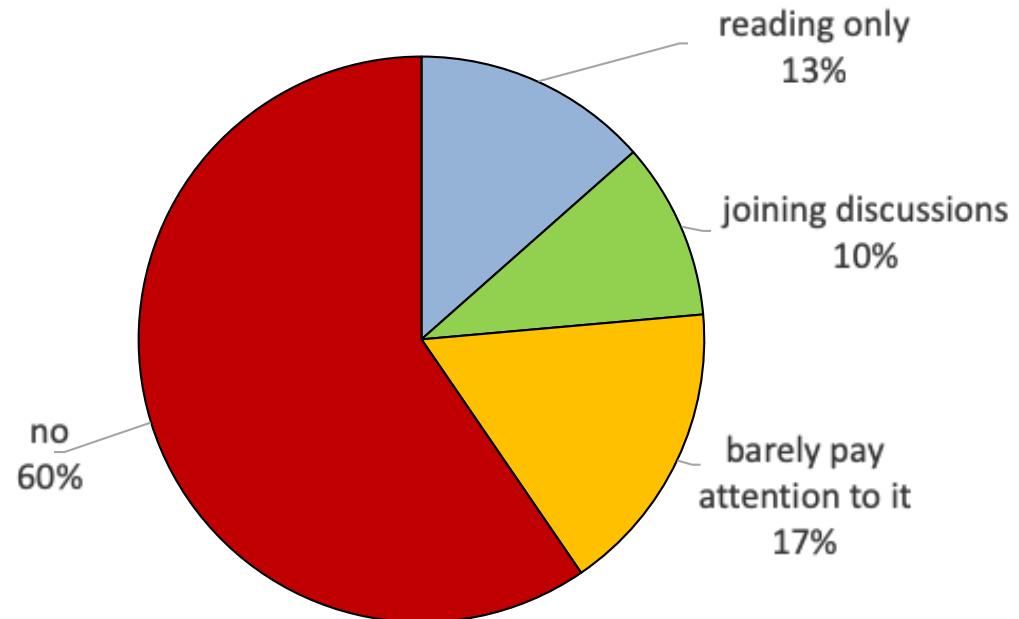
2nd easybuild User Survey



Are you subscribed to the EasyBuild mailing list?



Do you use the EasyBuild IRC/Slack channel?



(very similar to previous survey)

A significant part of the EasyBuild community remains silent.

IRC/Slack is not the right place to make "big" decisions,
since less than 25% actually pays attention to it.



easybuild documentation stats

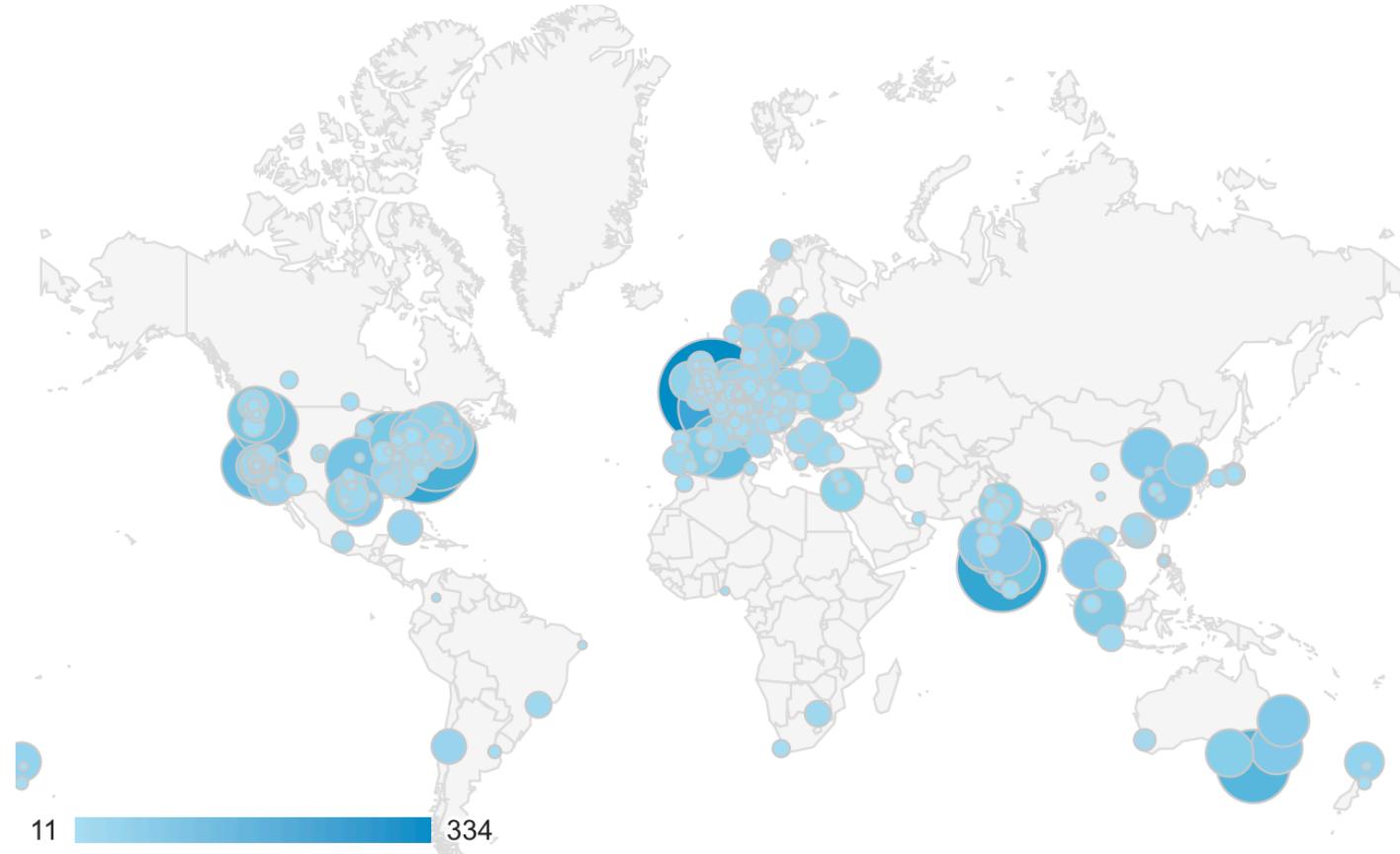


EasyBuild documentation at <https://easybuild.readthedocs.io>

400-500 weekly visitors

(source: Google Analytics)

The sun never sets on easybuild ...

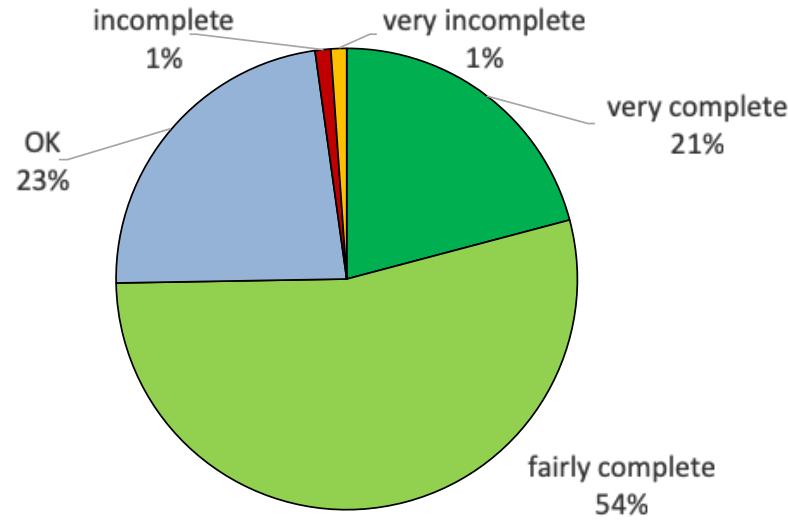


cities from where <https://easybuild.readthedocs.io> was visited
at least 10 times during the last year
(source: Google Analytics)

2nd easybuild User Survey

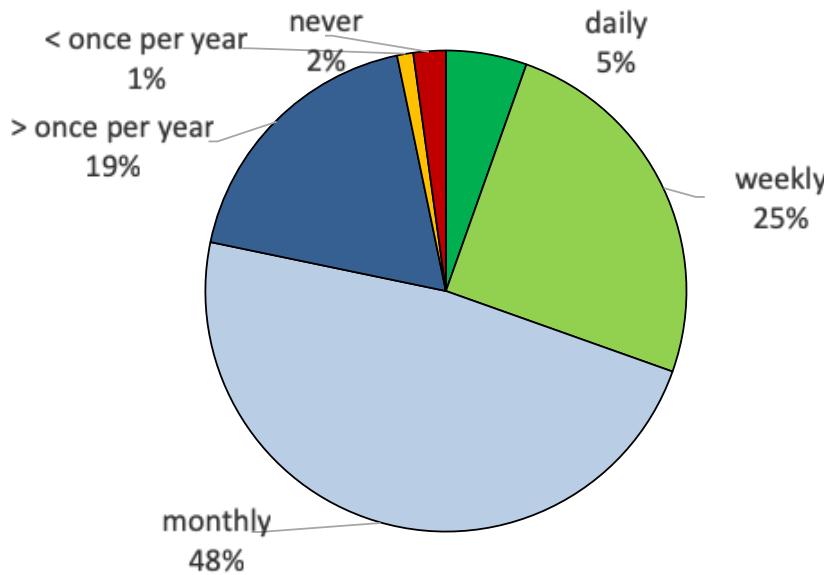


How complete is the EasyBuild documentation?



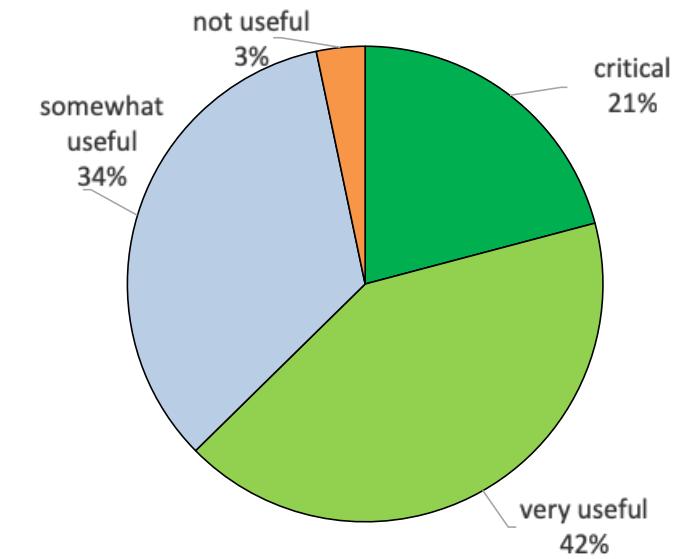
fairly to very complete
according to ~75%
(small decline compared to
previous survey, was 81%)

How often do you consult the EasyBuild documentation?



> 75% consults documentation
monthly or more frequent

How useful is the EasyBuild documentation to you?

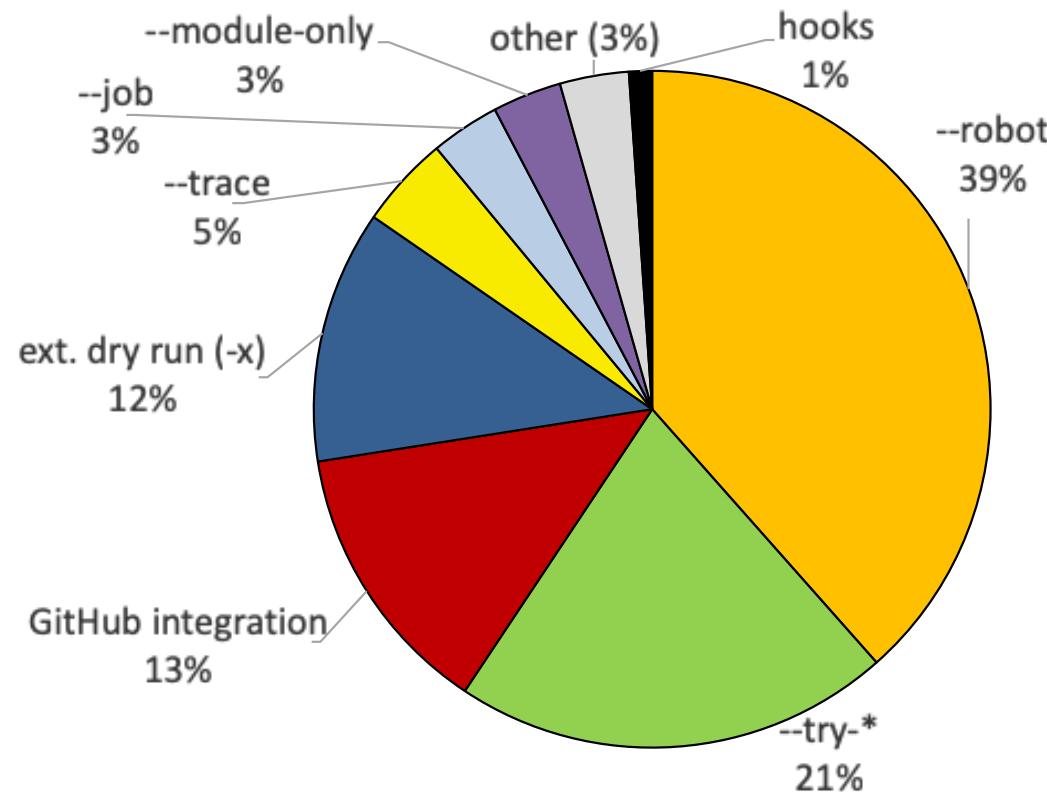


97% finds it at least
somewhat useful
(nobody answered
"totally useless")

2nd easybuild User Survey



What is your favourite EasyBuild feature?

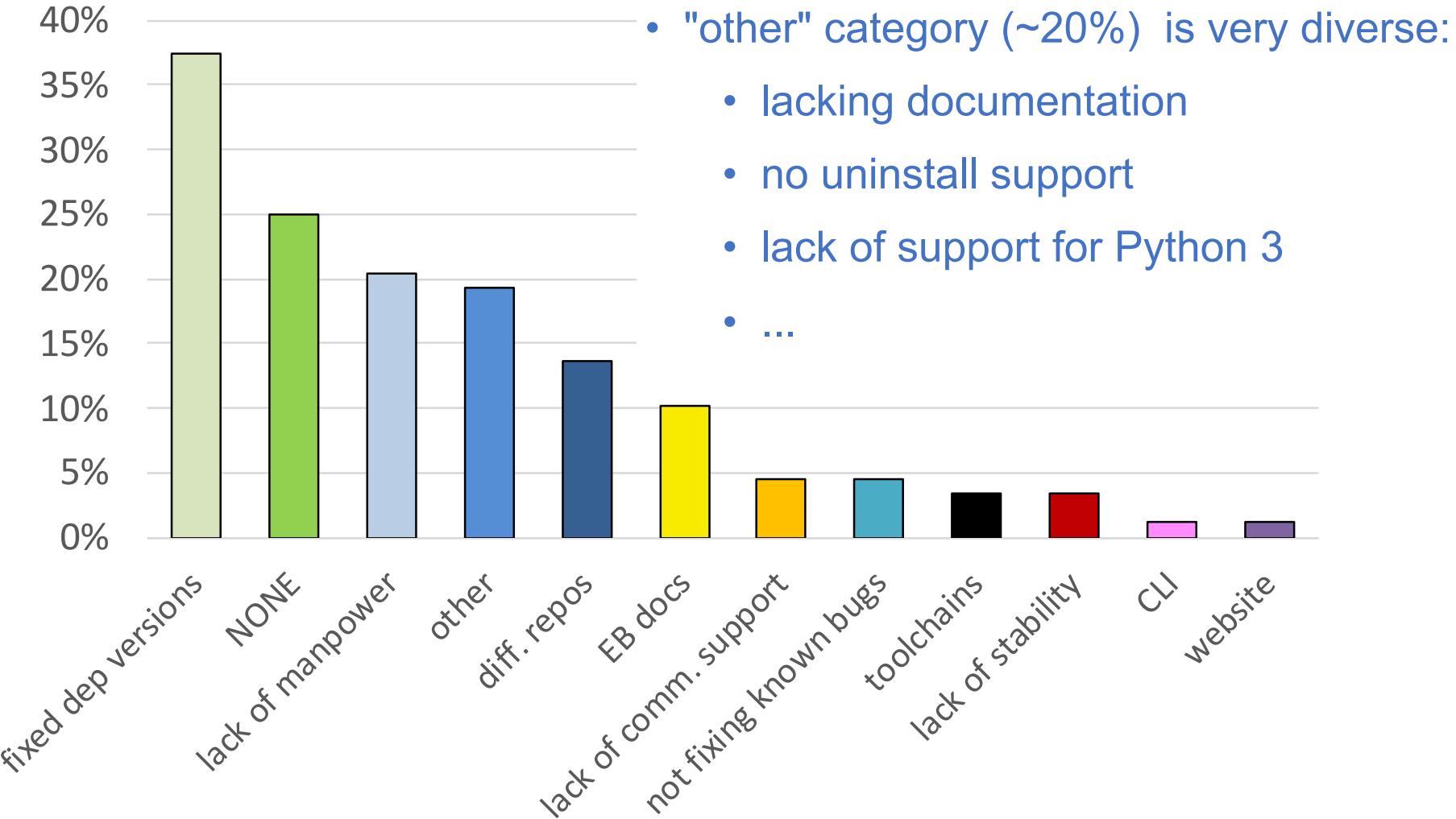


quite different compared to previous survey
(but list of suggested answers is also different)

2nd easybuild User Survey



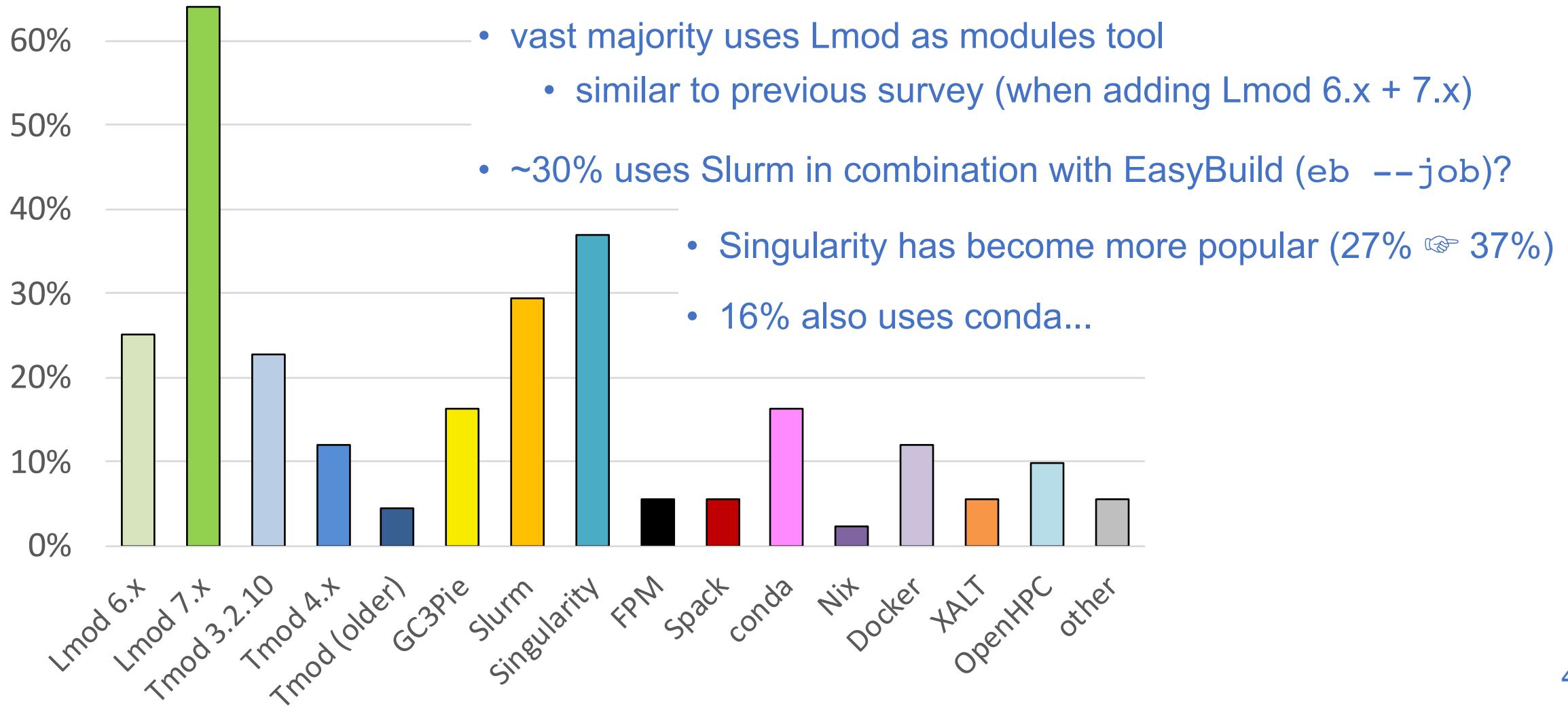
Which parts of EasyBuild do you not like? (check all that apply)



2nd easybuild User Survey



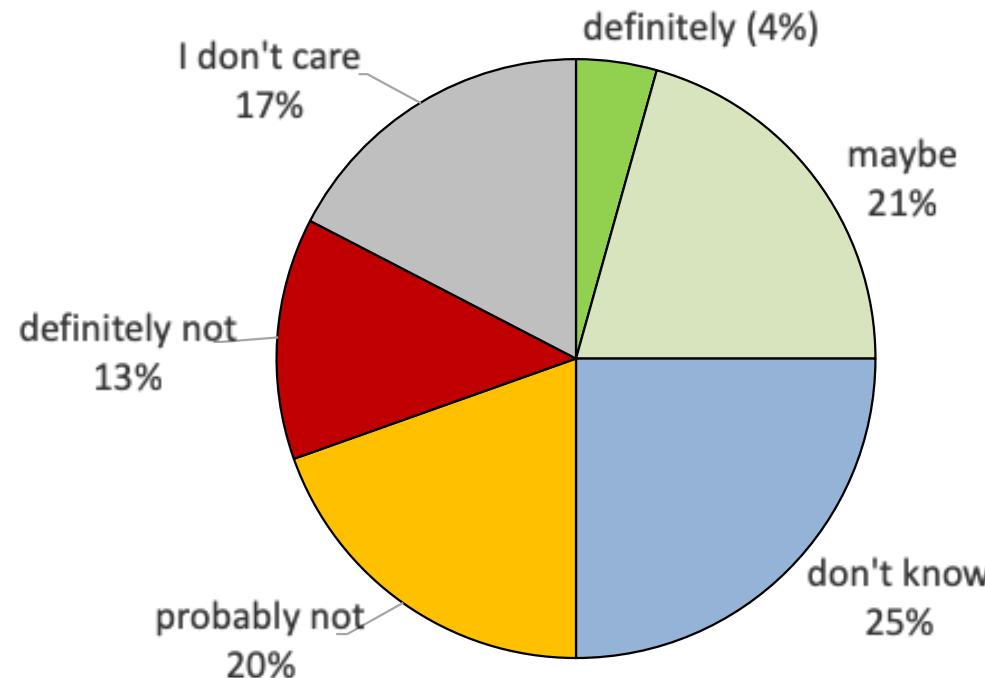
Do you use any other tools/projects in combination with EasyBuild?



2nd easybuild User Survey



If a commercial support option would be available,
would you consider purchasing it?



(very similar to previous survey)

Using a custom module naming scheme

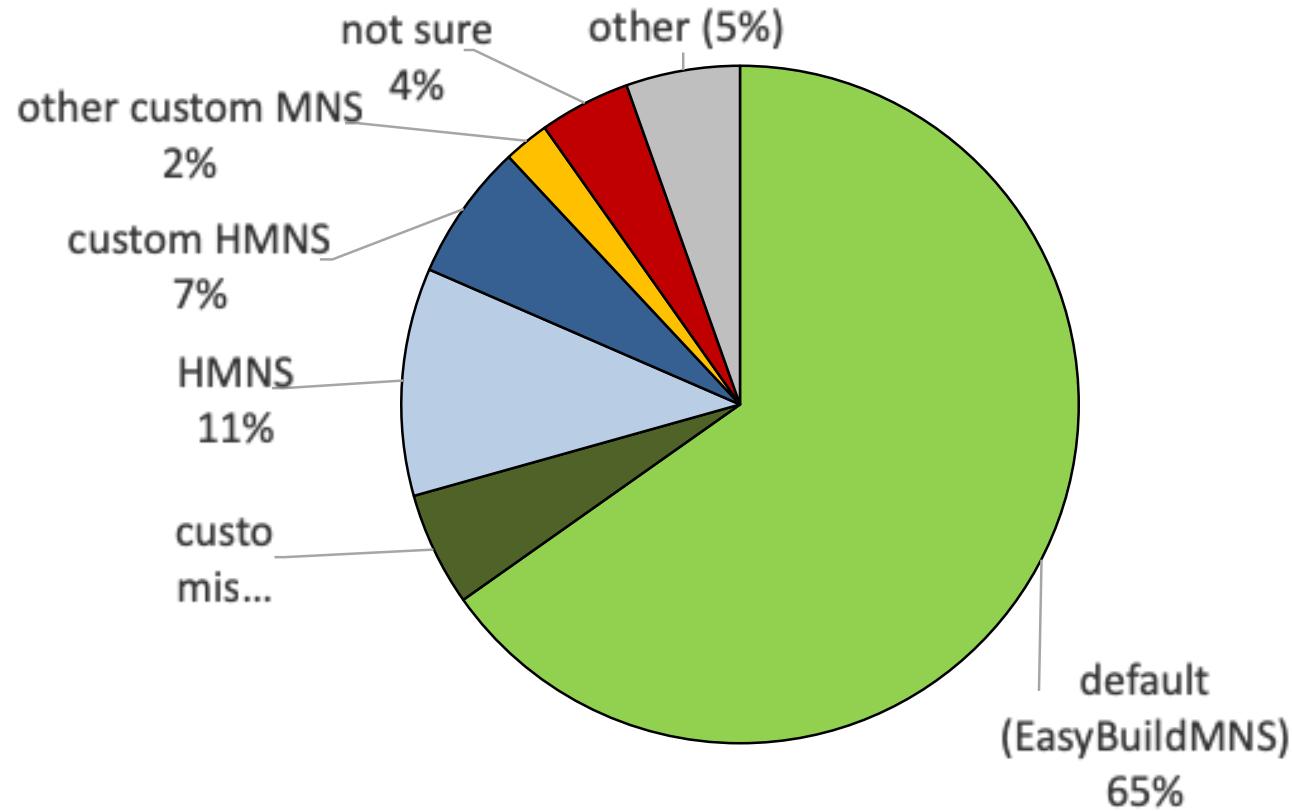


- a couple of different module naming schemes are included in EasyBuild
 - see `--avail-module-naming-schemes`
 - specify active module naming scheme via `--module-naming-scheme`
 - default: `EasyBuildMNS (<name>/<version>-<toolchain>-<versionsuffix>)`
- **you can implement your own module naming scheme relatively easily**
 - specify how to compose module name using provided metadata
 - via Python module that defines custom derivative class of `ModuleNamingScheme`
 - make EasyBuild aware of it via `--include-module-naming-schemes`
- decouple naming of install dirs vs modules via `--fixed-installdir-naming-scheme`

2nd easybuild User Survey



Which module naming scheme do you use?



(very similar to previous survey)

Flat module naming scheme



- all modules are always available for loading
- long(er) module names
- 'module avail' may be overwhelming for users
- too easy to load incompatible modules together

legend

(not available)
(available)
(loaded)

GCC/5.3.0

GCC/6.1.0

OpenMPI/1.10.2-GCC-5.3.0

OpenMPI/2.1.0-GCC-5.3.0

OpenMPI/1.10.3-GCC-6.1.0

OpenMPI/2.1.0-GCC-6.1.0

FFTW/3.3.4-gompi-2016.04

FFTW/3.3.6-gompi-2016.04

FFTW/3.3.4-gompi-2016.07

FFTW/3.3.6-gompi-2016.07

Hierarchical module naming scheme (1)



- modules are organised in a tree-like fashion
- initially, only 'core' modules are available for loading
- typically 3 hierarchy levels: core, compiler-dependent, MPI-dependent
- other modules are only visible via 'module spider'

legend

(not available)
(available)
(loaded)

Core level

GCC/5.3.0

GCC/6.1.0

OpenMPI/1.10.6

OpenMPI/2.1.0

*Compiler level
(compiler-dependent)*

FFTW/3.3.4

FFTW/3.3.5

FFTW/3.3.6

FFTW/3.3.4

FFTW/3.3.5

FFTW/3.3.6

*MPI level
(MPI-dependent)*

Hierarchical module naming scheme (2)



- Core modules may extend \$MODULEPATH with an additional location
- loading a Core module may make more modules available
- in this example, loading a GCC module makes OpenMPI modules available

legend

(not available)

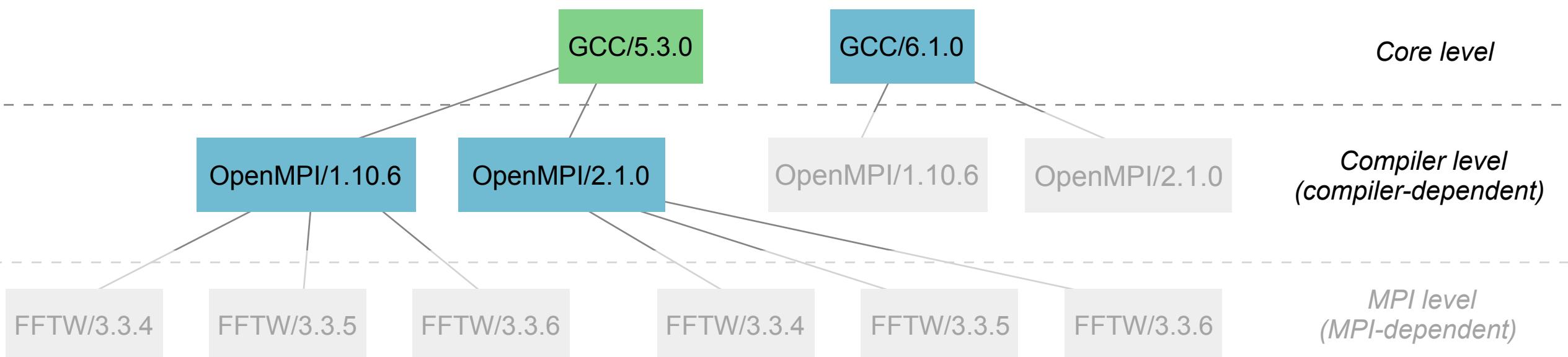
(available)

(loaded)

Core level

Compiler level
(compiler-dependent)

MPI level
(MPI-dependent)



Hierarchical module naming scheme (3)

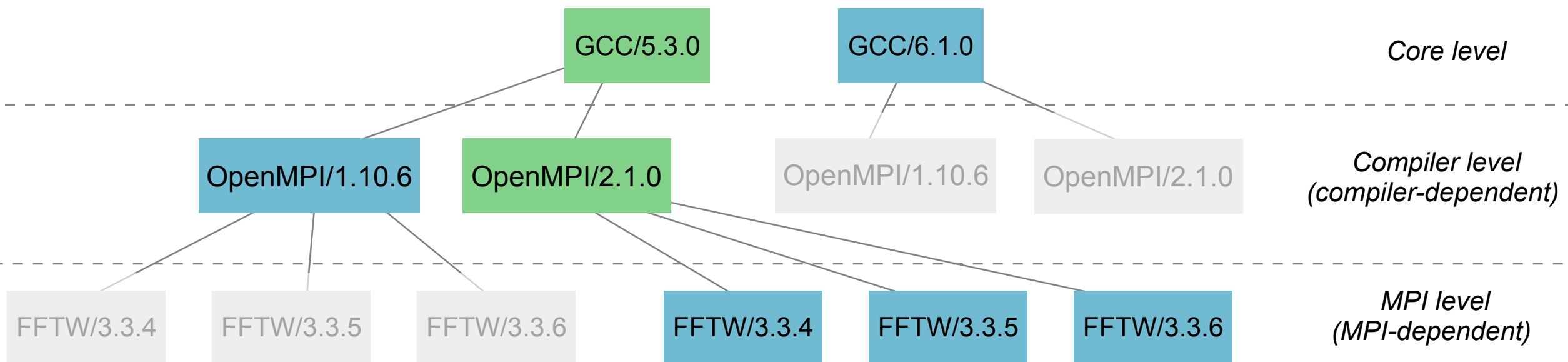


- even more modules may be made available by loading other modules
- for example, loading an OpenMPI modules reveals MPI-dependent modules
- EasyBuild can organise modules in hierarchy for you!

legend

(not available)
(available)
(loaded)

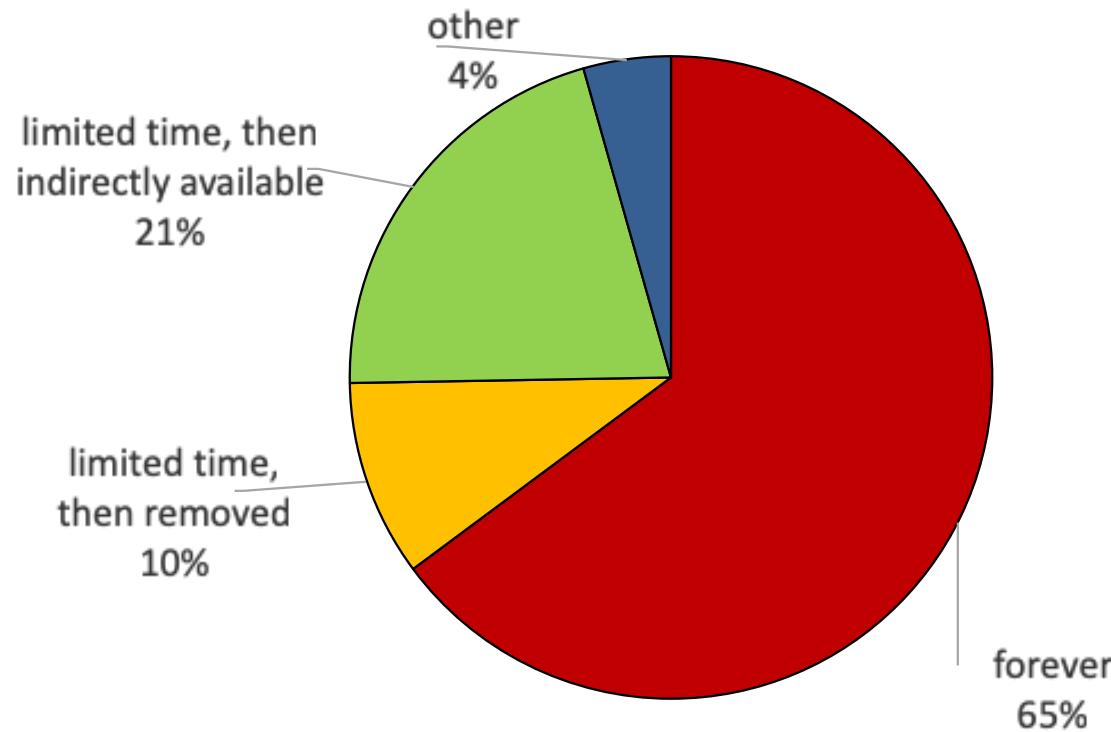
Core level



2nd easybuild User Survey



How long are modules installed with EasyBuild available to users?



increase for "limited time, then indirectly available": 14% ↗ 21%

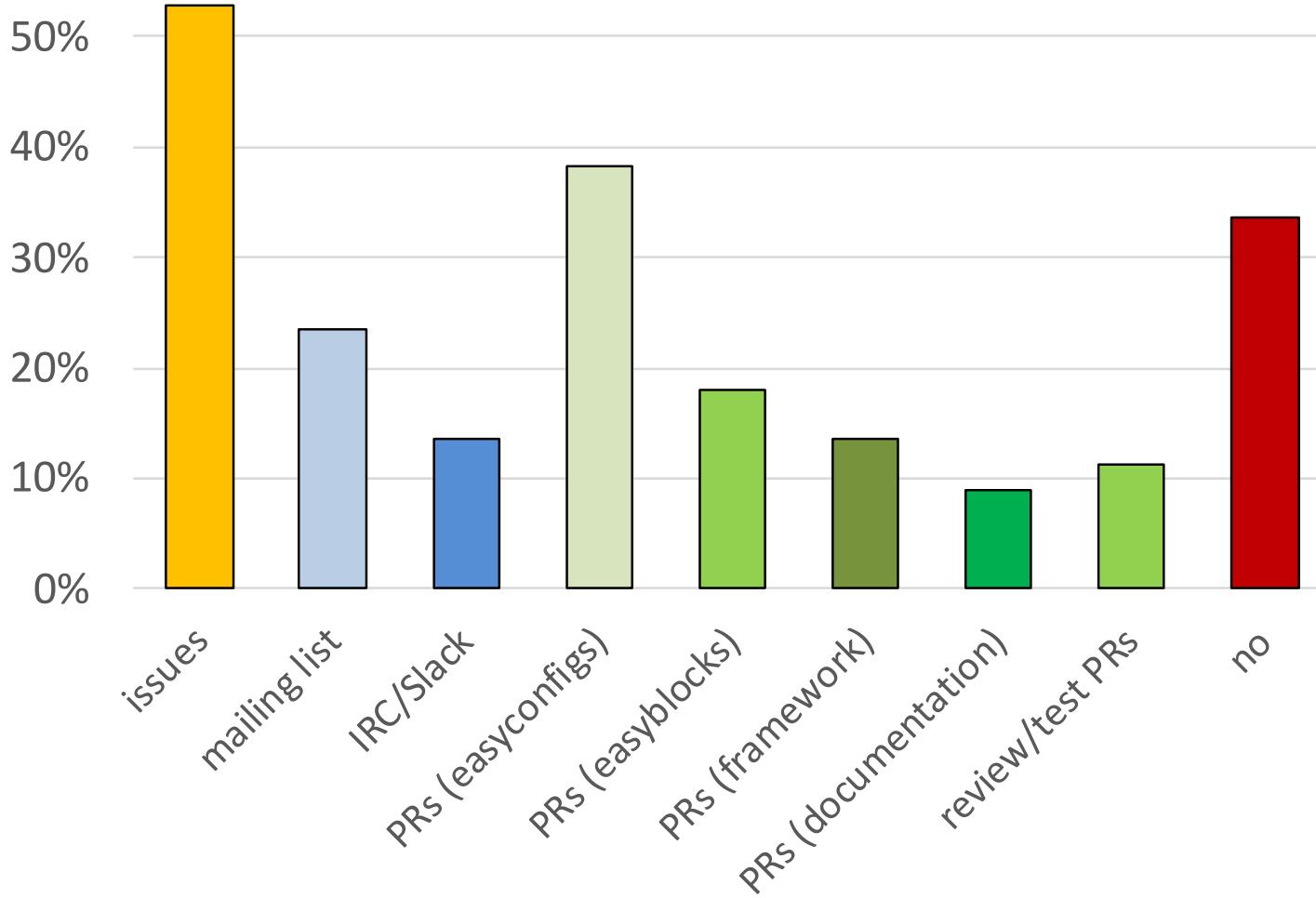
Contributors, contributors, contributors...



2nd easybuild User Survey

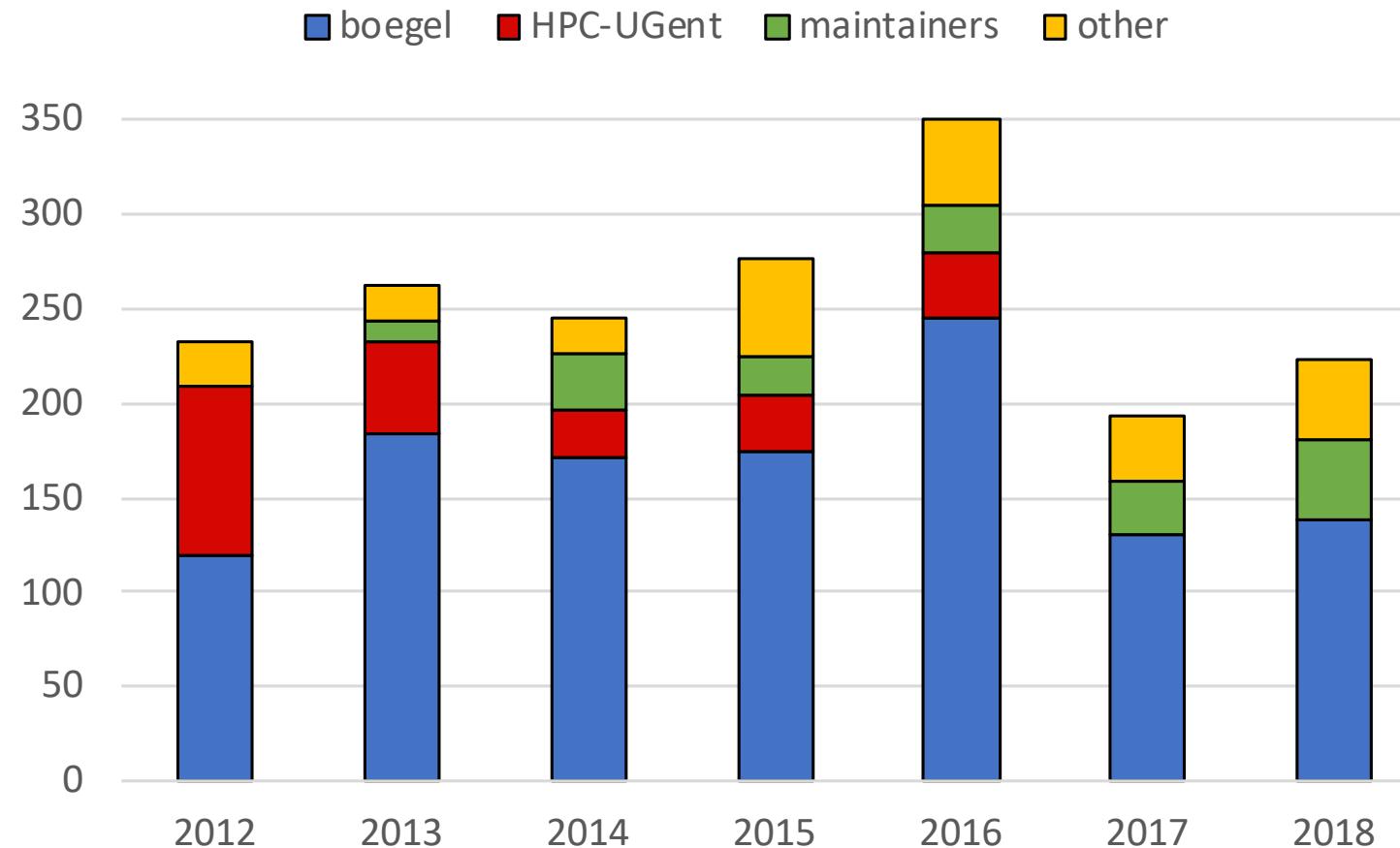


Do you actively contribute to EasyBuild? (check all that apply)



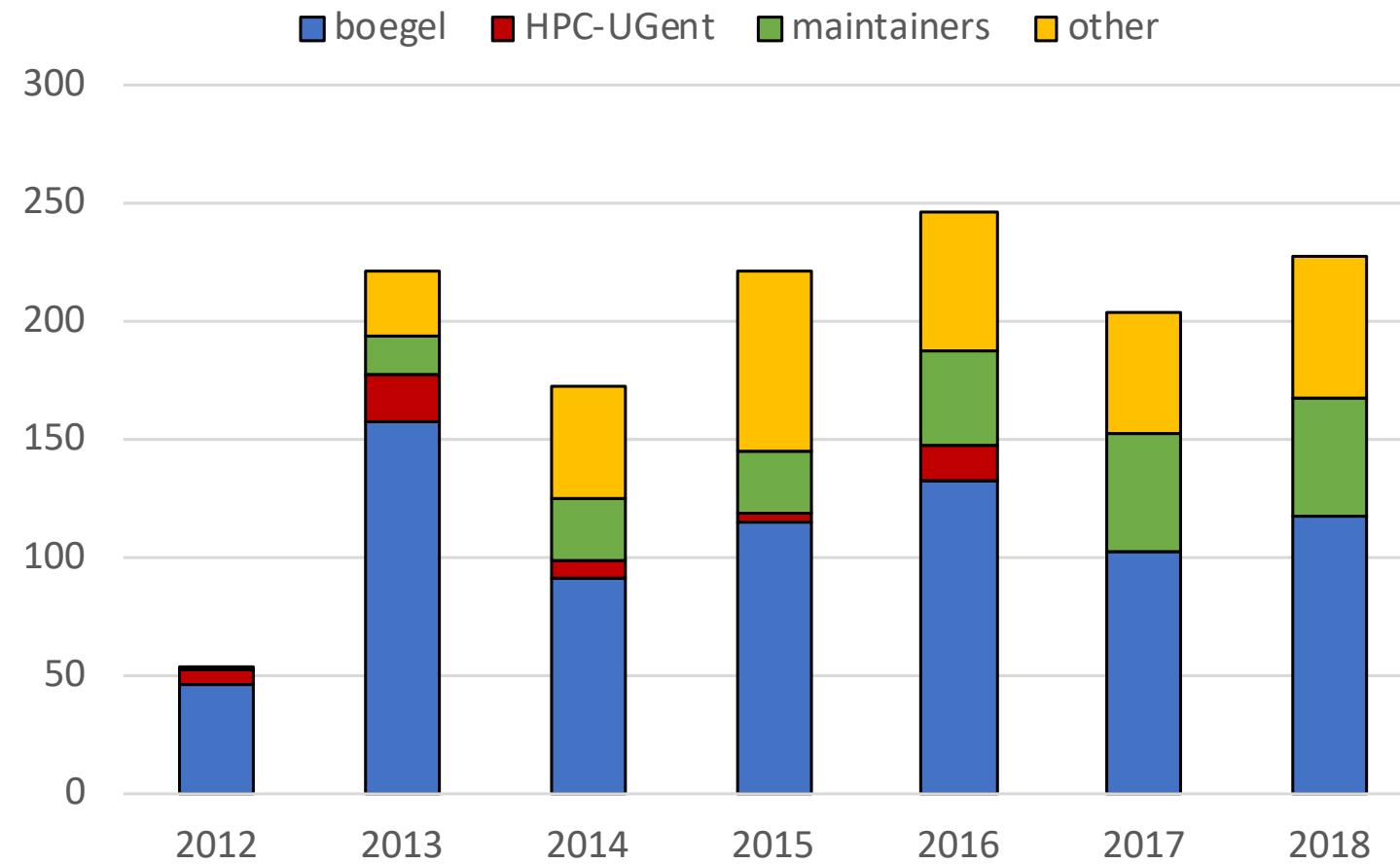
- Overall minor decline:
 - mailing list (35% ↘ 23%)
 - IRC/Slack (18% ↘ 13%)
 - easyconfig PRs (49% ↘ 38%)
 - easyblock PRs (26% ↘ 18%)
 - framework PRs (18% ↘ 13%)
 - review/testing of PRs (18% ↘ 11%)
- Increase for issues (47% ↗ 53%)
- Similar for documentation + no contr.

Overview of contributions (framework)



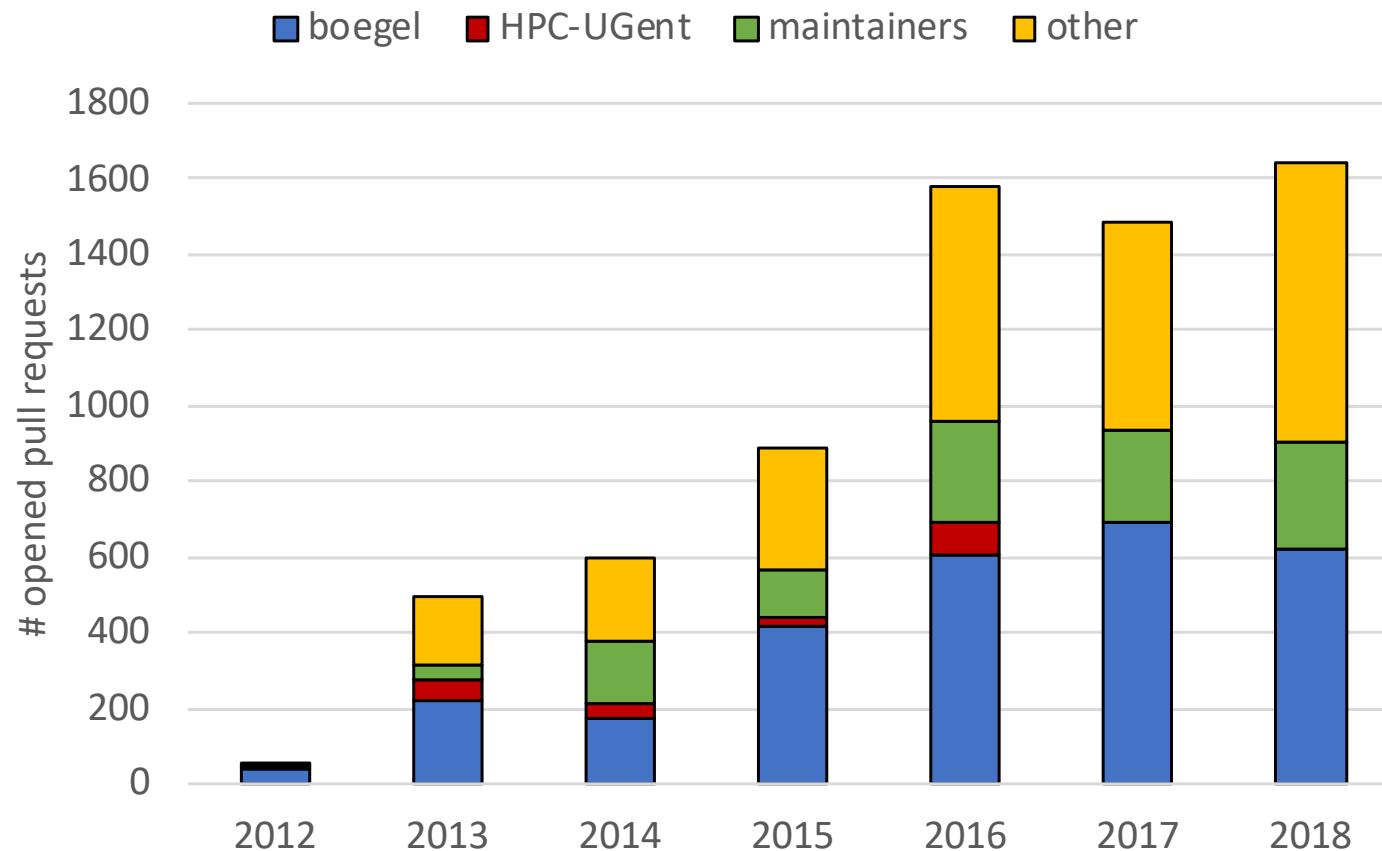
- less PRs than back in 2016 (mostly due to lack of time by boegel)
- ~20% of PRs by maintainers other than boegel + additional 20% by others

Overview of contributions (easyblocks)



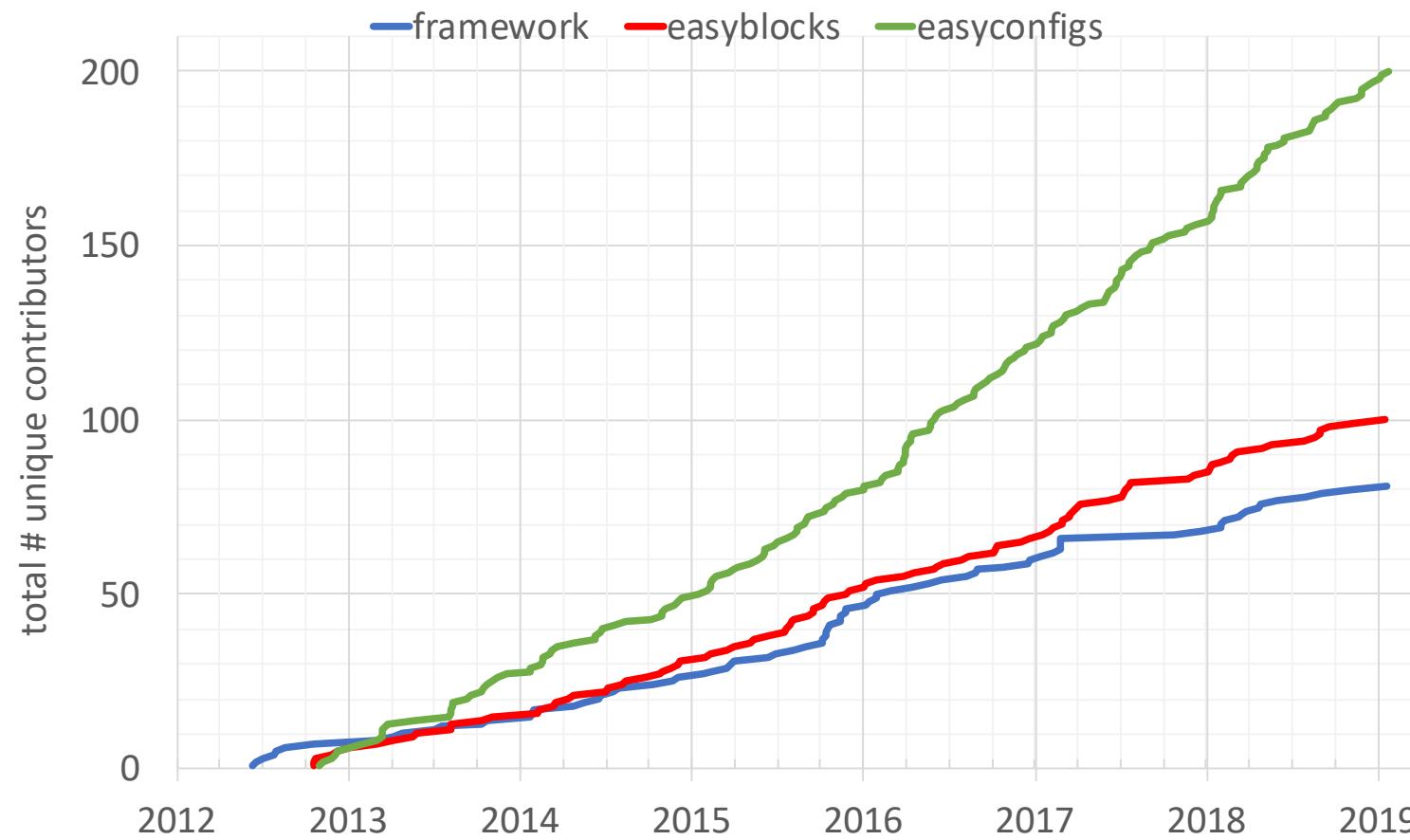
- 25% of PRs by maintainers other than boegel + additional 25% by others
- mostly maintenance of existing easyblocks

Overview of contributions (easyconfigs)



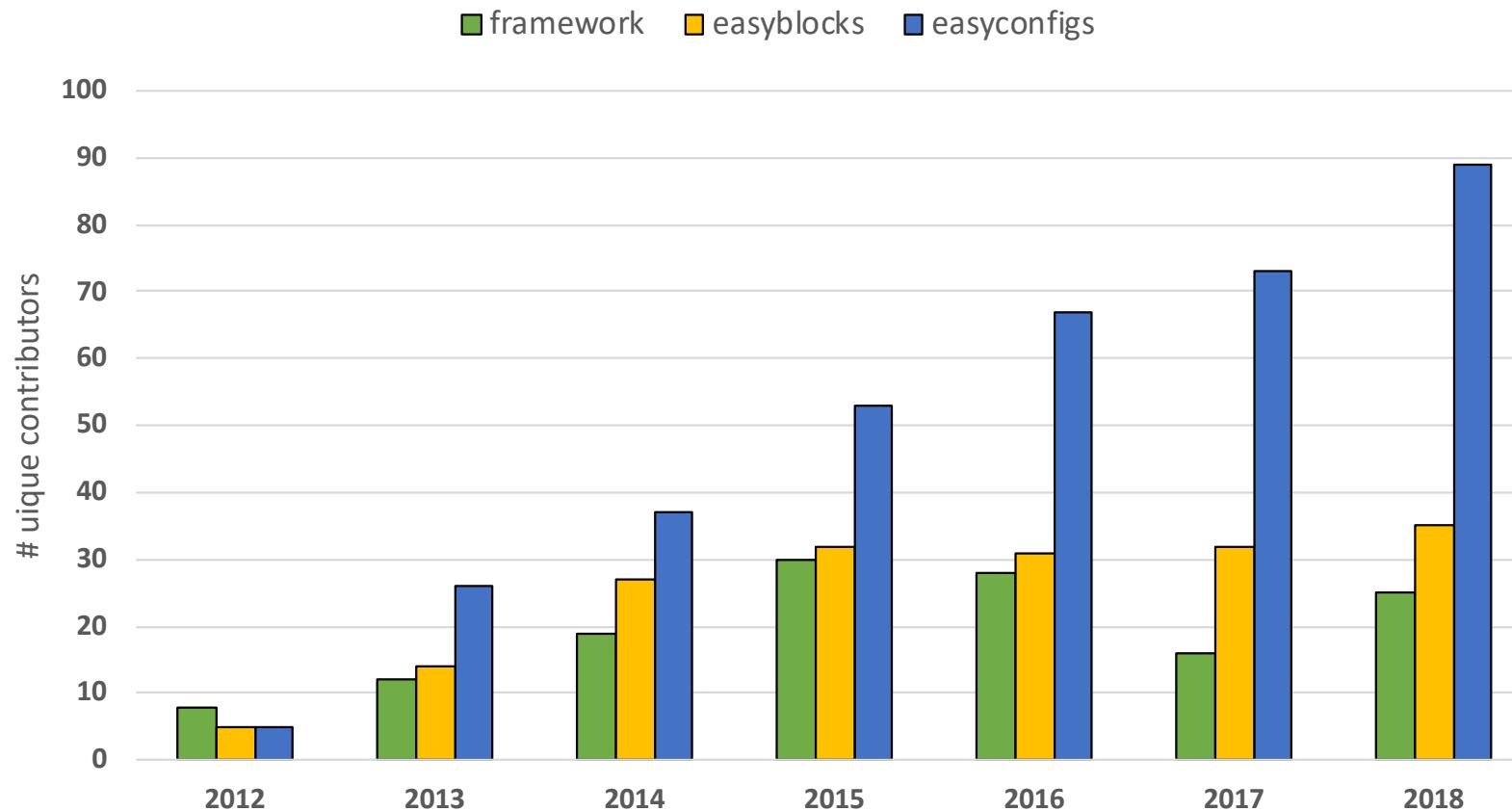
- explosive growth in easyconfig PRs since 2016 (partially thanks to `--new-pr`)
- stabilised in last couple of years; ~1,600 PRs in 2018
- growing ratio of PRs from contributors that are not maintainers (~45% in 2018)

Overview of unique contributors (total)



- over 80 unique contributors to EasyBuild framework so far
- recently reached **100 total unique contributors for easyblocks, 200 for easyconfigs!**
- steadily growing number of unique contributors to easyconfigs

Overview of unique contributors (per year)



- ~25 unique contributors per year to EasyBuild framework
- 30-35 unique contributors per year to easyblocks, fairly stable in recent years)
- steadily growing number of unique contributors per year for easyconfigs (100 in 2019?)

Maintainers, maintainers, maintainers, ...





easybuild maintainers

Since summer 2017, there are **10 EasyBuild maintainers**.

Since May 2018, there is a rotating "maintainer-of-the-week" (MotW) role.



Damian Alvarez - @damianam
(Jülich Supercomputing Centre)



Miguel Dias Costa - @migueldiascosta
(National University of Singapore)



Pablo Escobar - @pescobar
(sciCORE, University of Basel)



Kenneth Hoste - @boegel
(HPC-UGent)



Adam Huffman - @verdurin
(Big Data Institute, University of Oxford)



Alan O'Cais - @ocaisa
(Jülich Supercomputing Centre)



Bart Oldeman - @bartoldeman
(ComputeCanada)



Ward Poelmans - @wpoely86
(Free University of Brussels)



Åke Sandgren - @akesandgren
(Umeå University, Sweden)

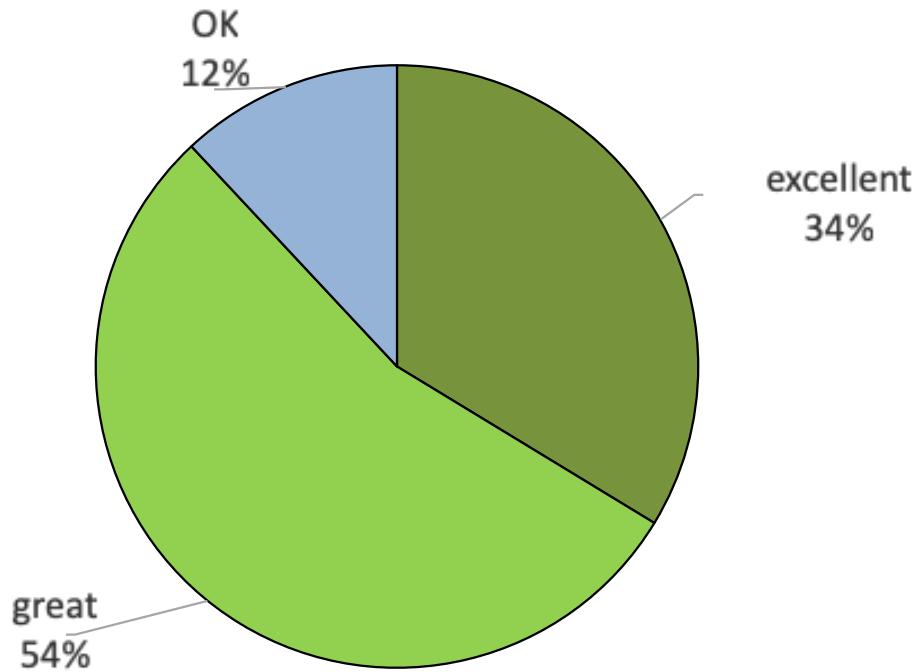


Davide Vanzo - @vanzod
(Vanderbilt University)

2nd easybuild User Survey



How would you rate the overall quality of EasyBuild?



100% positive!

(nobody answered "could be better" or "pretty bad")

(very similar to previous survey)

2nd easybuild User Survey



Do you have any suggestions for additional features? (1/2)

- support for uninstalling packages (with check on dependencies)
- better support for filtering/hiding dependencies
- "containers, containers, containers"
- support for skipping sanity check (supported since EasyBuild 3.7.0 via `skipsteps`)
- "steep learning curve"
- missing documentation on best practices / common workflows

2nd easybuild User Survey



Do you have any suggestions for additional features? (2/2)

- list of dependency versions per toolchain (esp. for common toolchains)
- loosen up the restriction on dependency versions
- semi-auto creation of easyconfigs based on existing ones (WIP: eb --new)
- installing multiple packages in a single-directory more easily
- better error reporting / readability of build logs
- support for using multiple module naming schemes at once

Future work: changes w.r.t. Python (packages)



We're actively looking into some changes w.r.t. handling of:

- **installation of Python interpreter + standard library** (with batteries incl.)
 - now done using "full" toolchain (foss/intel/...)
 - looking into installing Python with subtoolchain (GCCcore)
 - + separate bundle with Python packages that require MPI/BLAS/LAPACK
- **installation of Python packages for different Python versions**
 - now done with separate easyconfigs/installation directories/modules
 - considering switch to ComputeCanada approach (*see Bart's remote talk on Thu*)
 - single easyconfig/installdir/module combined some Python sys.path magic to automatically pick up the right installation based on active Python version

Future work: testing contributions in isolation



Testing contributions in isolation:

- mainly relevant for easyconfig PRs
- avoid that OS packages on test system affect test result
- test contributions both in:
 - minimal environment
 - fully loaded system (Boost, zlib, etc. installed in OS)
 - different (versions of) Linux distros
- via (standardised?) Singularity images

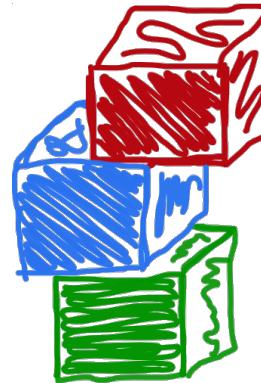


5th easybuild User Meeting

Party edition?

Sometime late 2019 or early 2020...

Let me know if you're interested in hosting it!



easybuild

Questions?

kenneth.hoste@ugent.be
easybuild@lists.ugent.be

<http://easybuilders.github.io/easybuild>
<http://easybuild.readthedocs.io>



State of the Union

4th EasyBuild User Meeting

January 30th 2019 - Louvain-la-Neuve (Belgium)

https://users.ugent.be/~kehoste/EasyBuild_20190130_state-of-the-union.pdf

kenneth.hoste@ugent.be

easybuild@lists.ugent.be

<https://easybuilders.github.io/easybuild>

<https://easybuild.readthedocs.io>