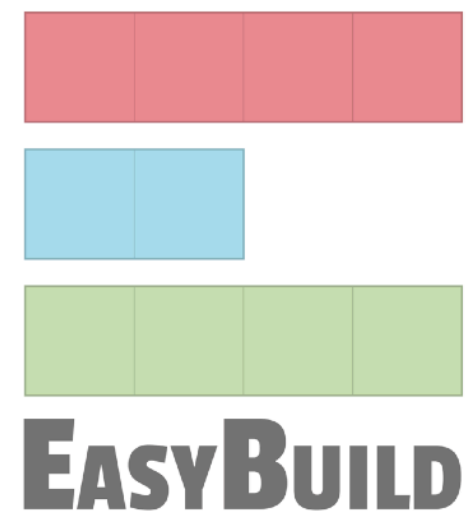# State of the union

Kenneth Hoste (HPC-UGent)

kenneth.hoste@ugent.be

9th EasyBuild User Meeting

23 April 2024 - Umeå (Sweden)

https://easybuild.io/eum24

EasyBuild in a nutshell

- **Framework to get scientific software installed** (preferably from source)

- Focus on Linux & HPC systems, specific attention to *performance* of installed software

- Implemented in Python, (currently still) supports both Python 2.7 and 3.5+ *(soon only Python 3.6+)*

- Integrates with environment modules tool (Lmod, or Tcl-based Environment Modules)

- **Open source since 2012** (GPLv2), available via GitHub and Python Package Index (PyPI)

- Supports different compiler toolchains + close to 3,500 software packages (excl. extensions)

- Originally created by HPC-UGent, now supported & developed by a **worldwide community...**

EESSI in a nutshell

- **Community project** that evolved from the EasyBuild community, started in 2020

- Main goal: build a **publicly accessible, shared, optimized software stack** for scientific computing

- Focus on easy-of-use, performance, testing, automating, collaboration, …

- Powered by open source software, incl. EasyBuild, Lmod, CernVM-FS, Gentoo Prefix, ReFrame, …

- Effort is currently funded through MultiXscale EuroHPC Centre-of-Excellence

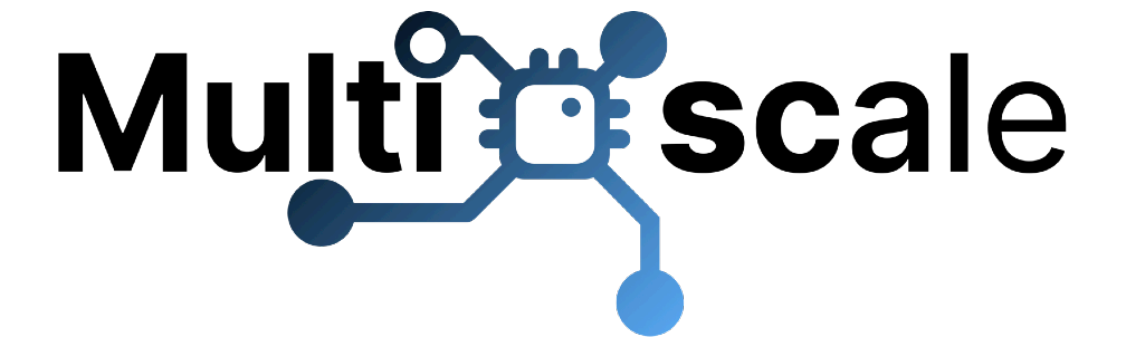- **We will focus extensively on EESSI during 3rd day of EUM'24**

eessi.io | github.com/EESSI | eessi.io/docs

MultiXscale in a nutshell

multixscale.eu

- **EuroHPC Centre of Excellence**

  Exascale-oriented application co-design and delivery for multiscale simulations

- Total funding: ~€ 6,000,000 (2023-2026), 16 partners in total, including HPC-UGent

- Two aspects:

  - Services to facilitate development and deployment of scientific software => (technical partners)

  - Development of software for efficient large-scale multiscale simulations (scientific partners)
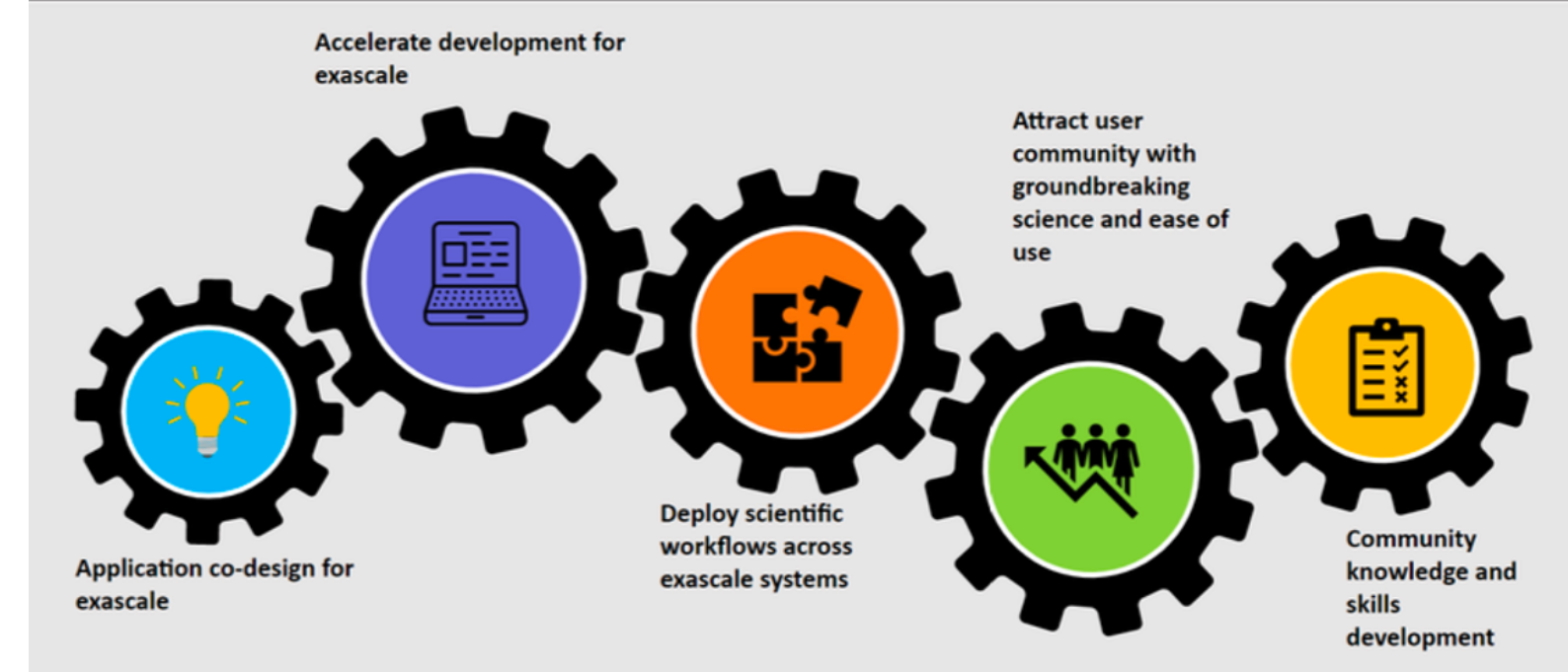
- **Source of funding for EESSI (and indirectly EasyBuild)**

  => More people can spend time working on these projects
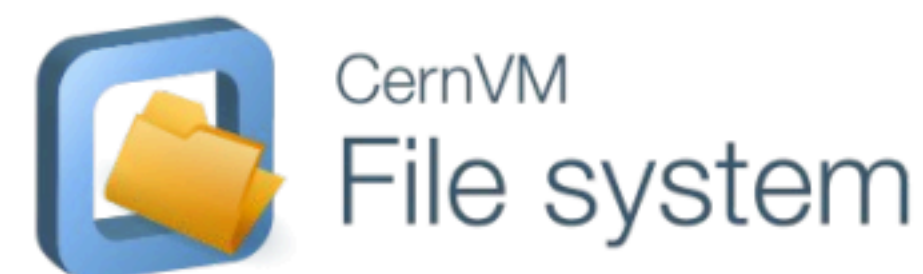
- **Gets us a seat at the table in EuroHPC community…**

MultiXscale and EESSI were highlighted as
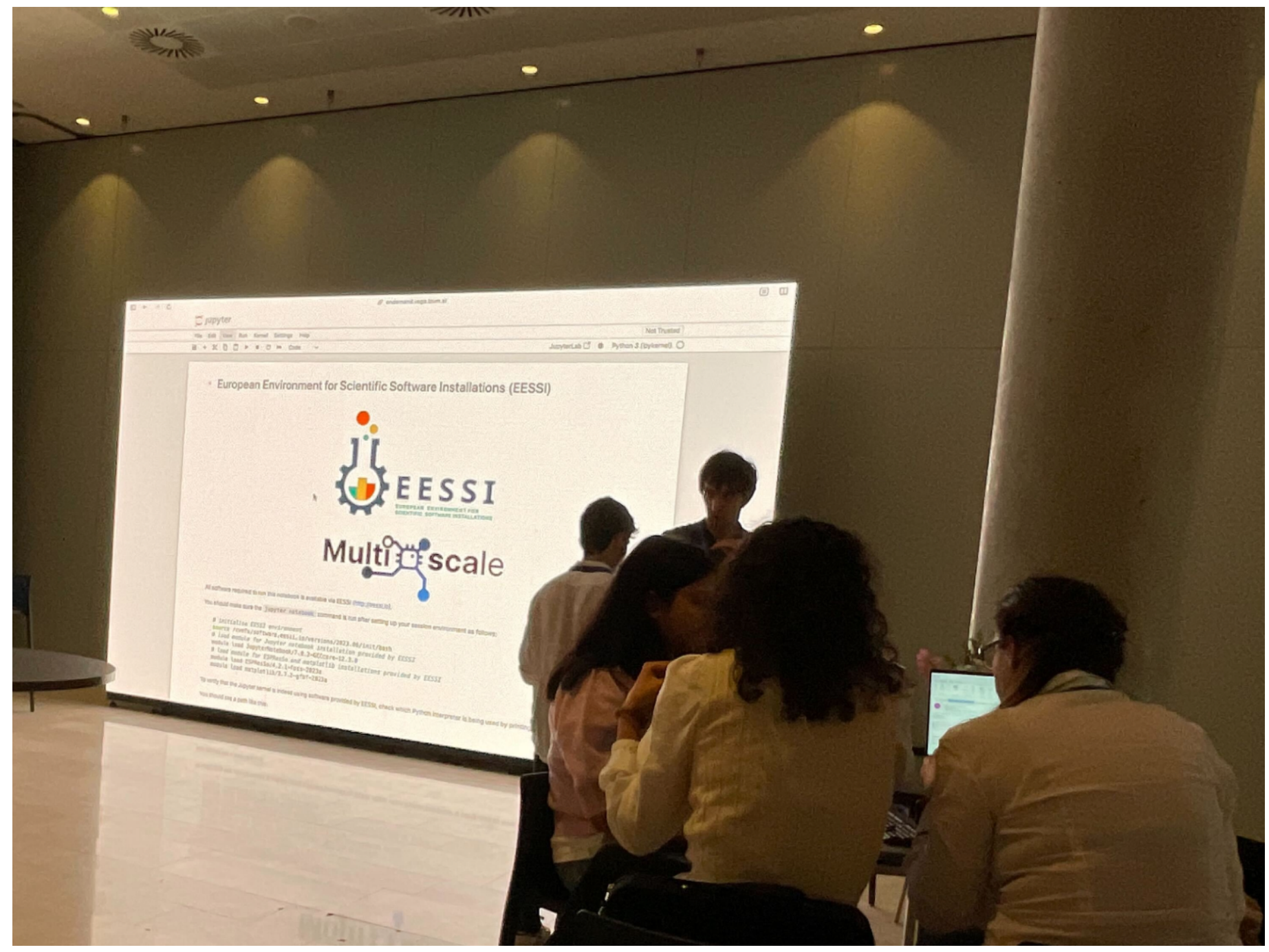*"example of a CoE that aligns well with goals of EuroHPC JU"*
at EuroHPC User Day (Dec'23, Brussels)

Half-day online tutorial on CernVM-FS and EESSI,
in collaboration with CernVM-FS developers

205 registrations, ~130 live attendees

Slides + recording available via

multixscale.github.io/cvmfs-tutorial-hpc-best-practices
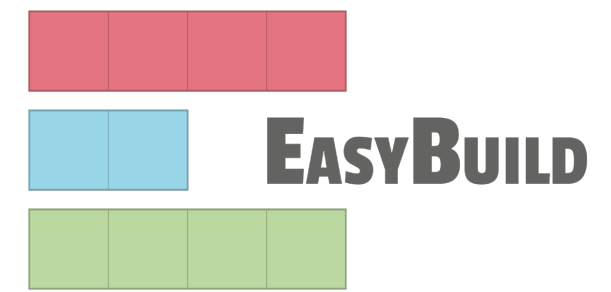
5

MultiXscale + EESSI promotion at EuroHPC Summit
(Antwerp, March'2024)

Poster session, presentations, panel sessions, …

EasyBuild User Survey

- Anonymous survey, **7th edition** (was held in March-mid April 2024)

- Get feedback from / gain insight into EasyBuild community

- 49 questions, covers wide variety of aspects

- 108 people participated

- Full survey results (raw) ~~are~~ will be available at https://docs.easybuild.io/user-survey

- Highlights are covered in this talk - questions with "boring" results are skipped

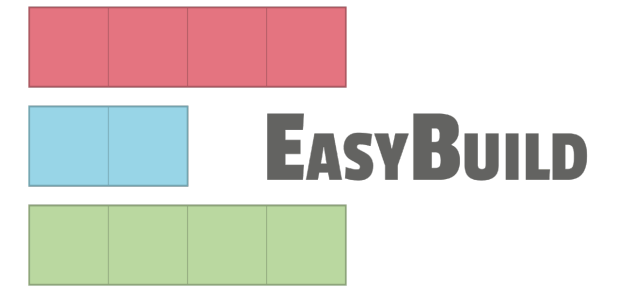- 7 years of survey results allows for observing evolution and trends…
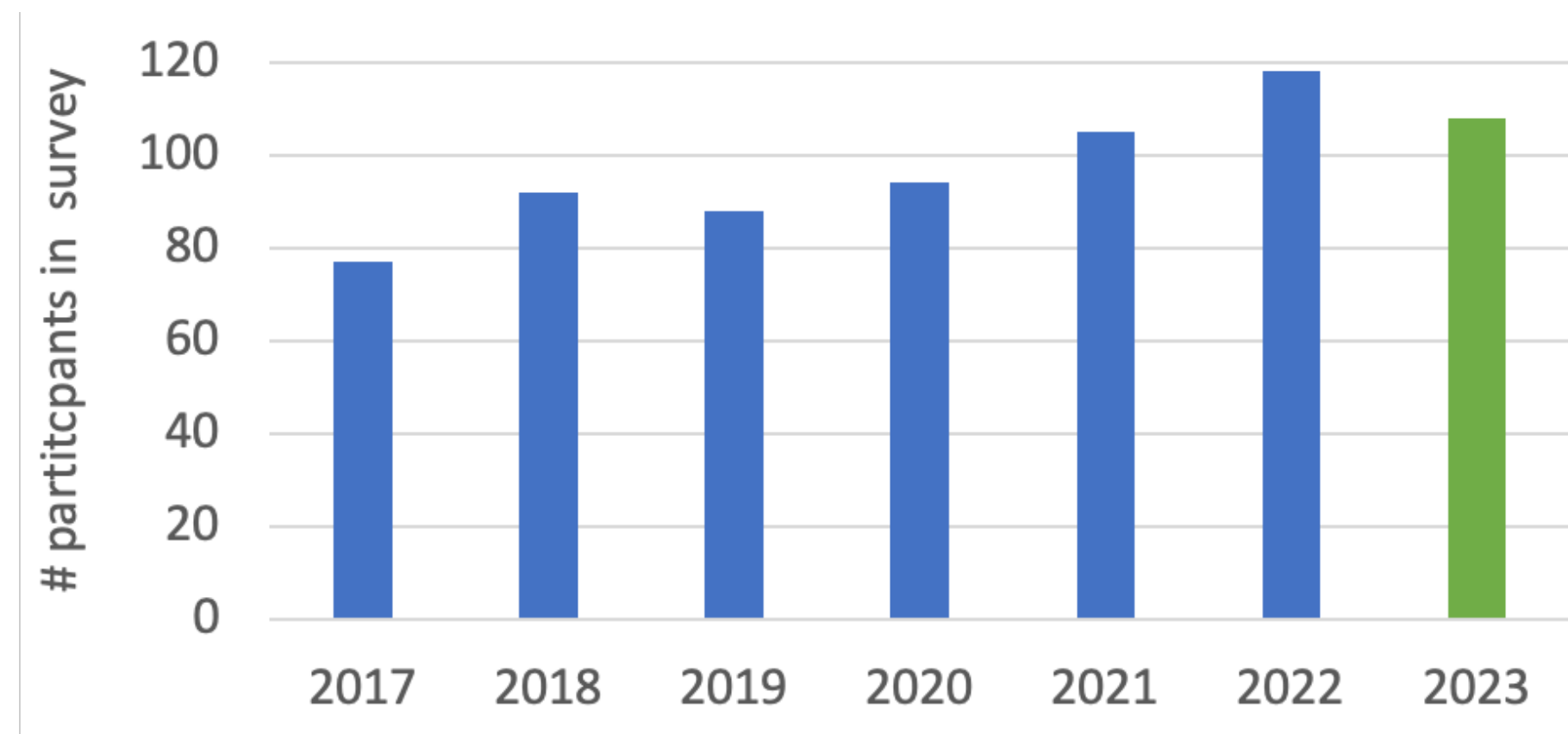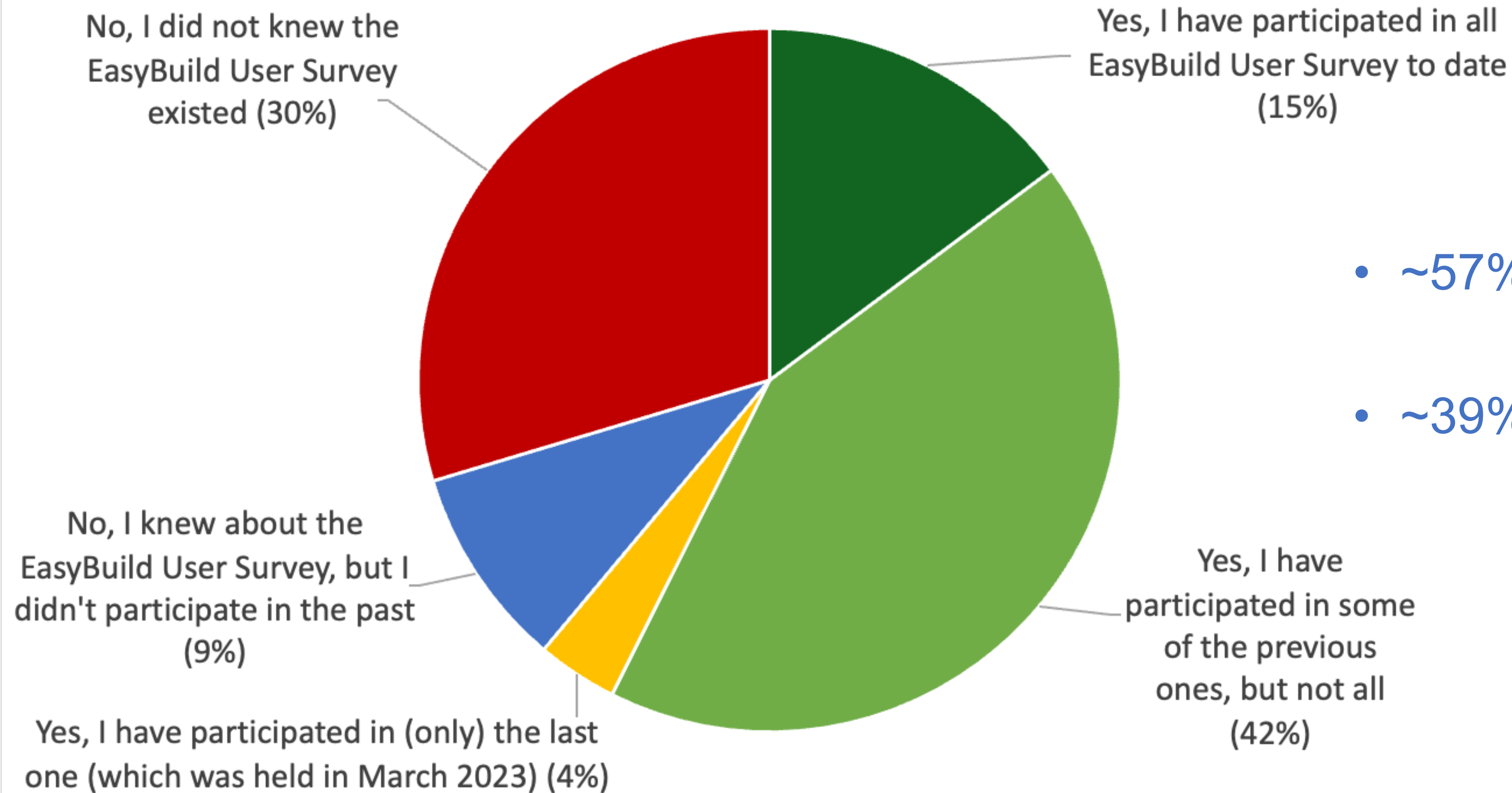
EasyBuild User Survey
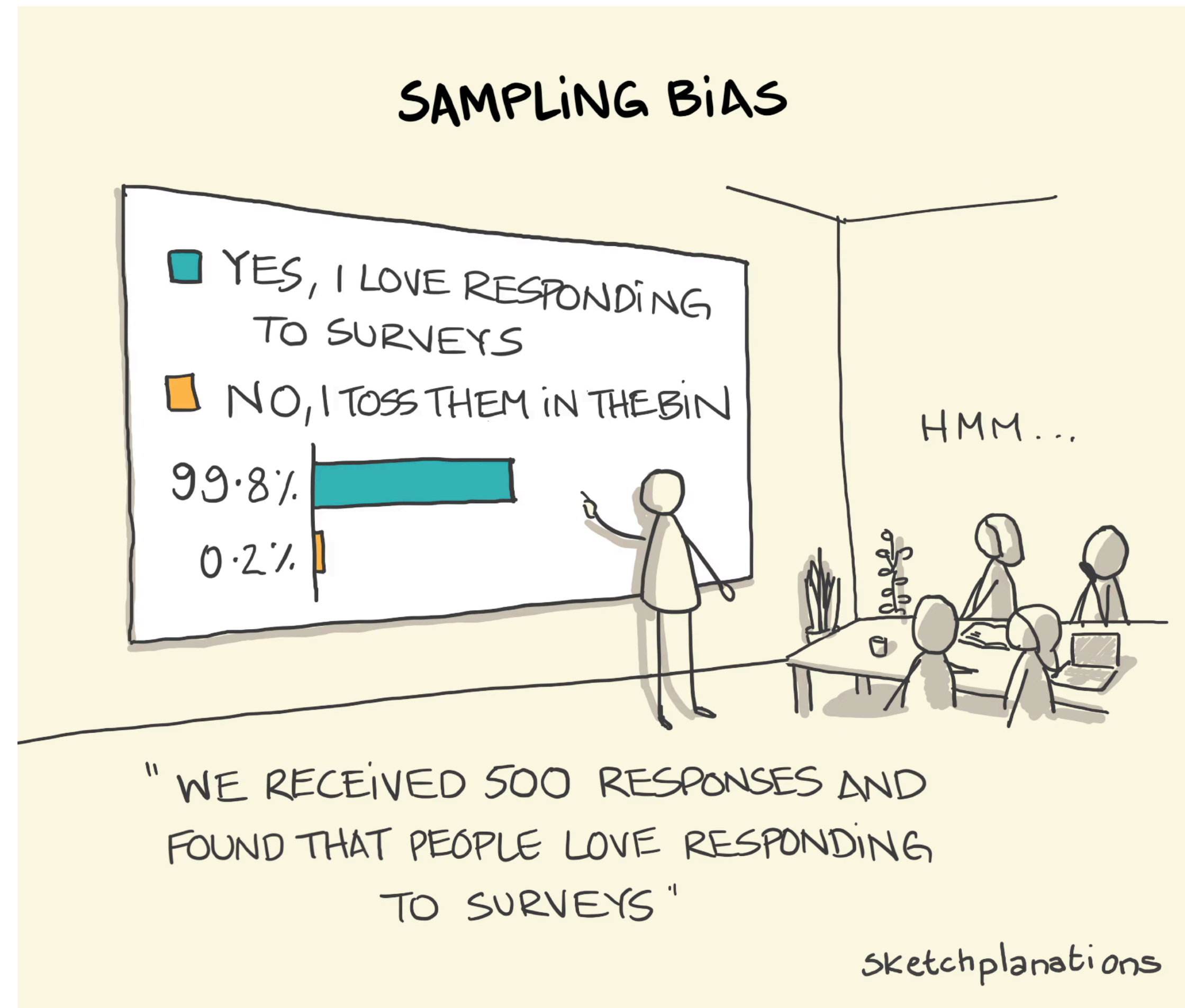
Survey results are *assumed* to give a representative view on EasyBuild community,

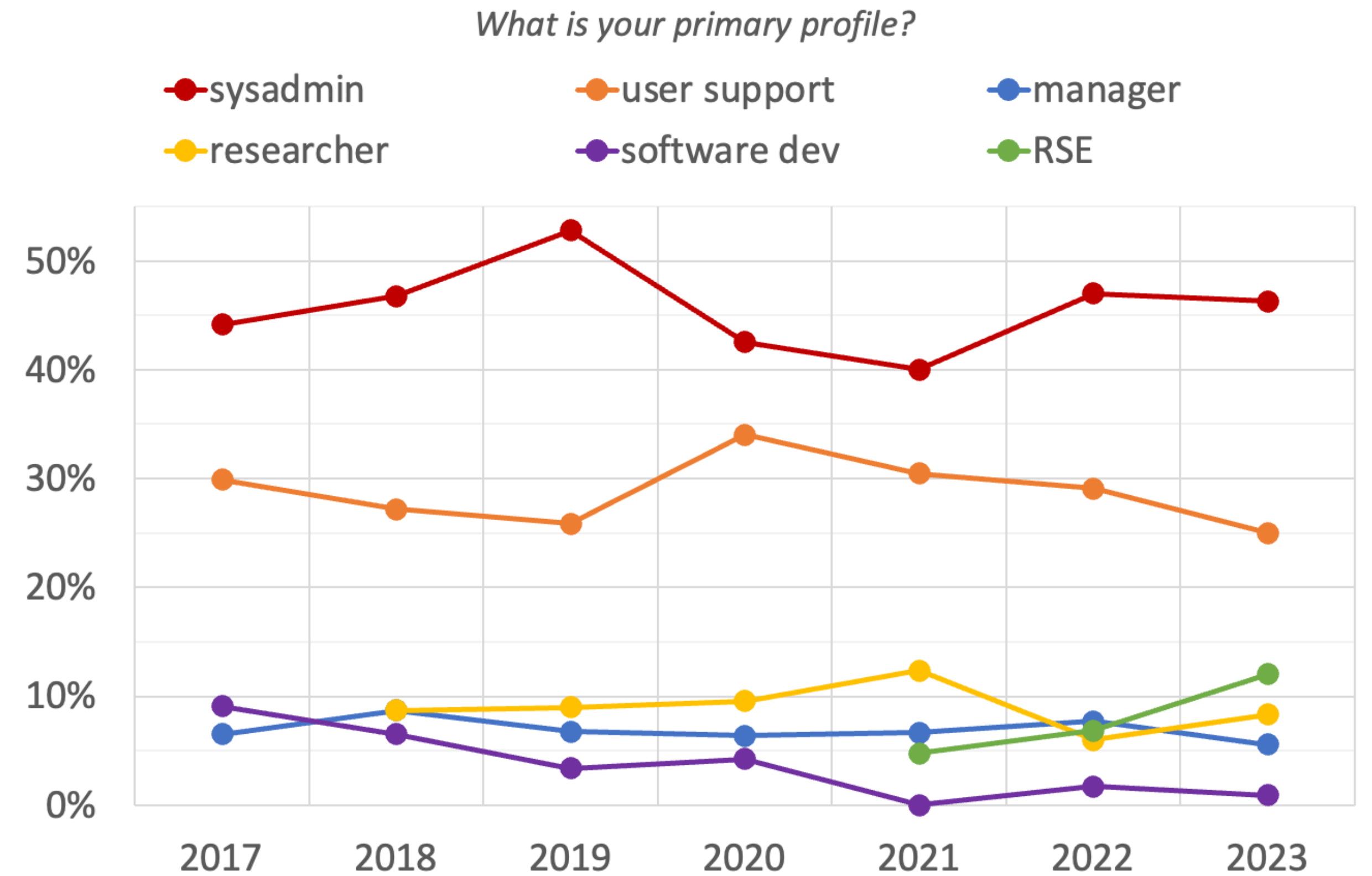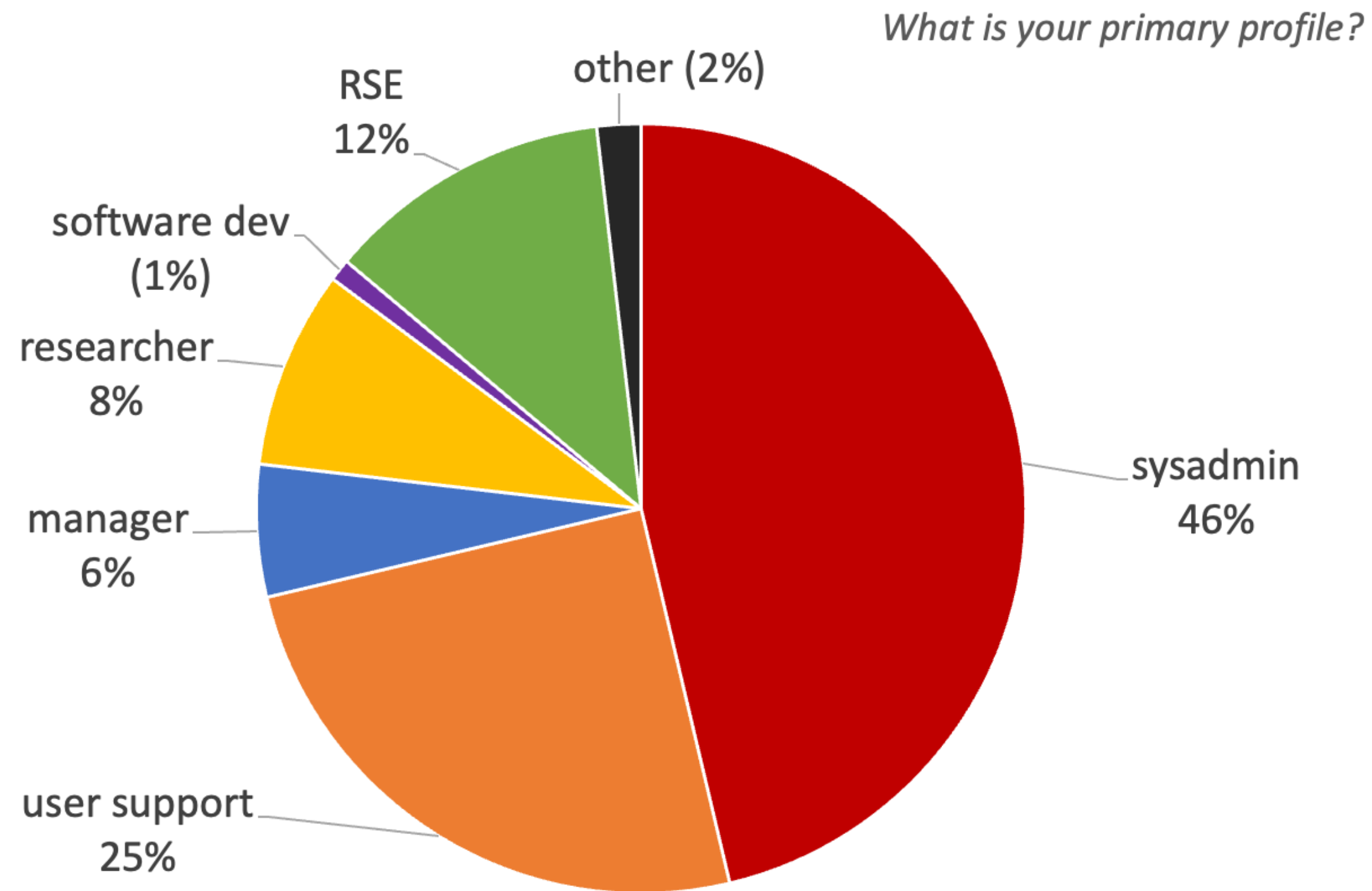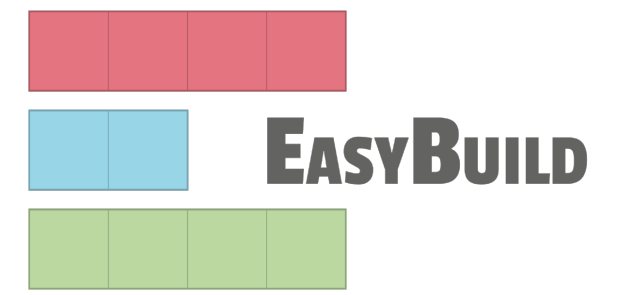but results should be interpreted with a grain of salt…

EasyBuild User Survey

Have you participated in the EasyBuild User Survey before?

No, I did not knew the EasyBuild User Survey existed (30%)

Yes, I have participated in all EasyBuild User Survey to date (15%)

No, I knew about the EasyBuild User Survey, but I didn't participate in the past (9%)

Yes, I have participated in (only) the last one (which was held in March 2023) (4%)

Yes, I have participated in some of the previous ones, but not all (42%)

- ~57% are experienced survey participants

- ~39% participate for the first time

EasyBuild demographics - primary profile

What is your primary profile?

- sysadmin 46%
- user support 25%
- manager 6%
- researcher 8%
- software dev (1%)
- RSE 12%
- other (2%)

What is your primary profile?

- sysadmin
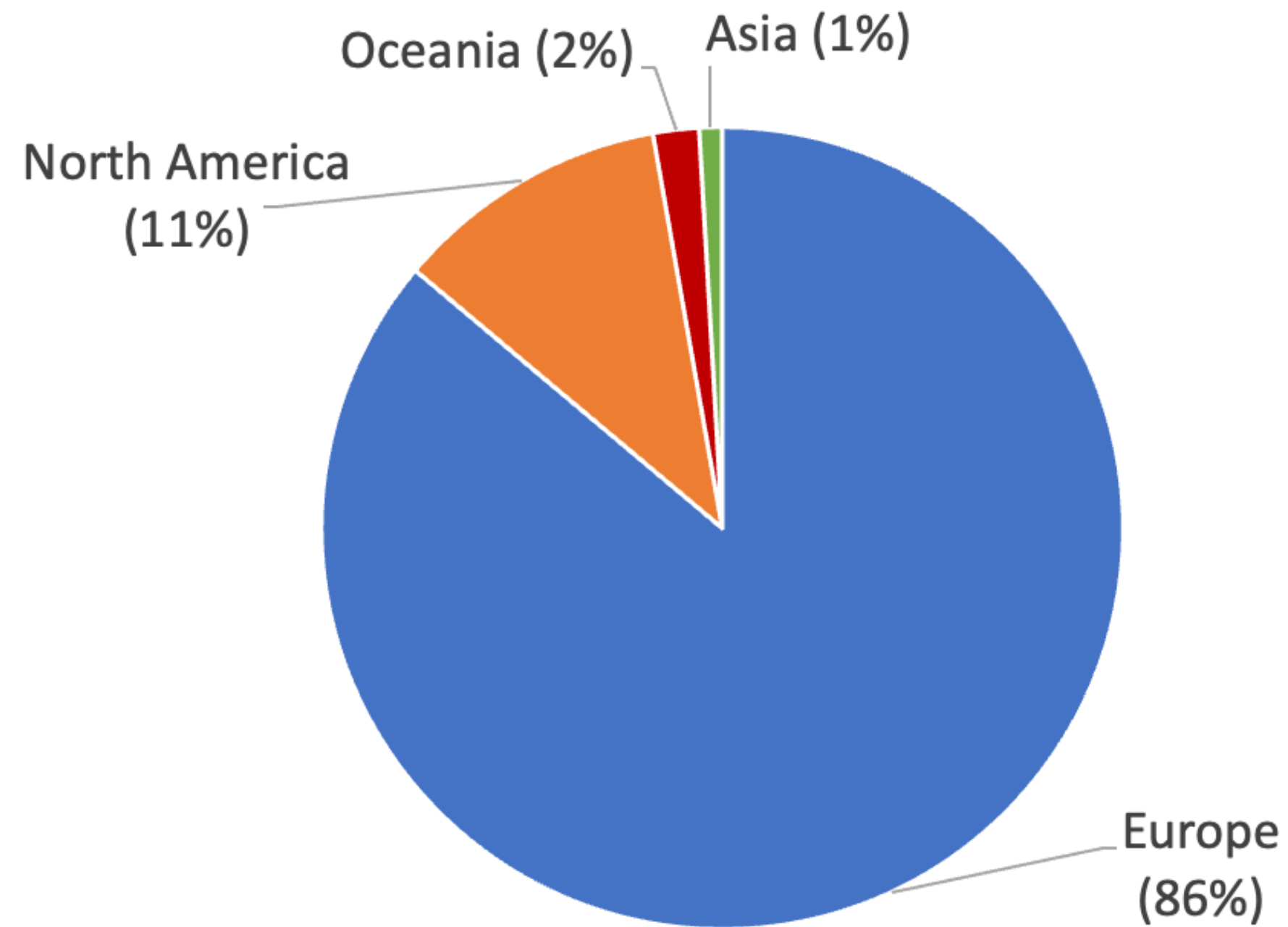- user support
- manager
- researcher
- software dev
- RSE

- Increase for "RSE" profile, at cost of "user support"
- Sysadmin profile still dominates, only 1% software dev.

EasyBuild demographics - geographical
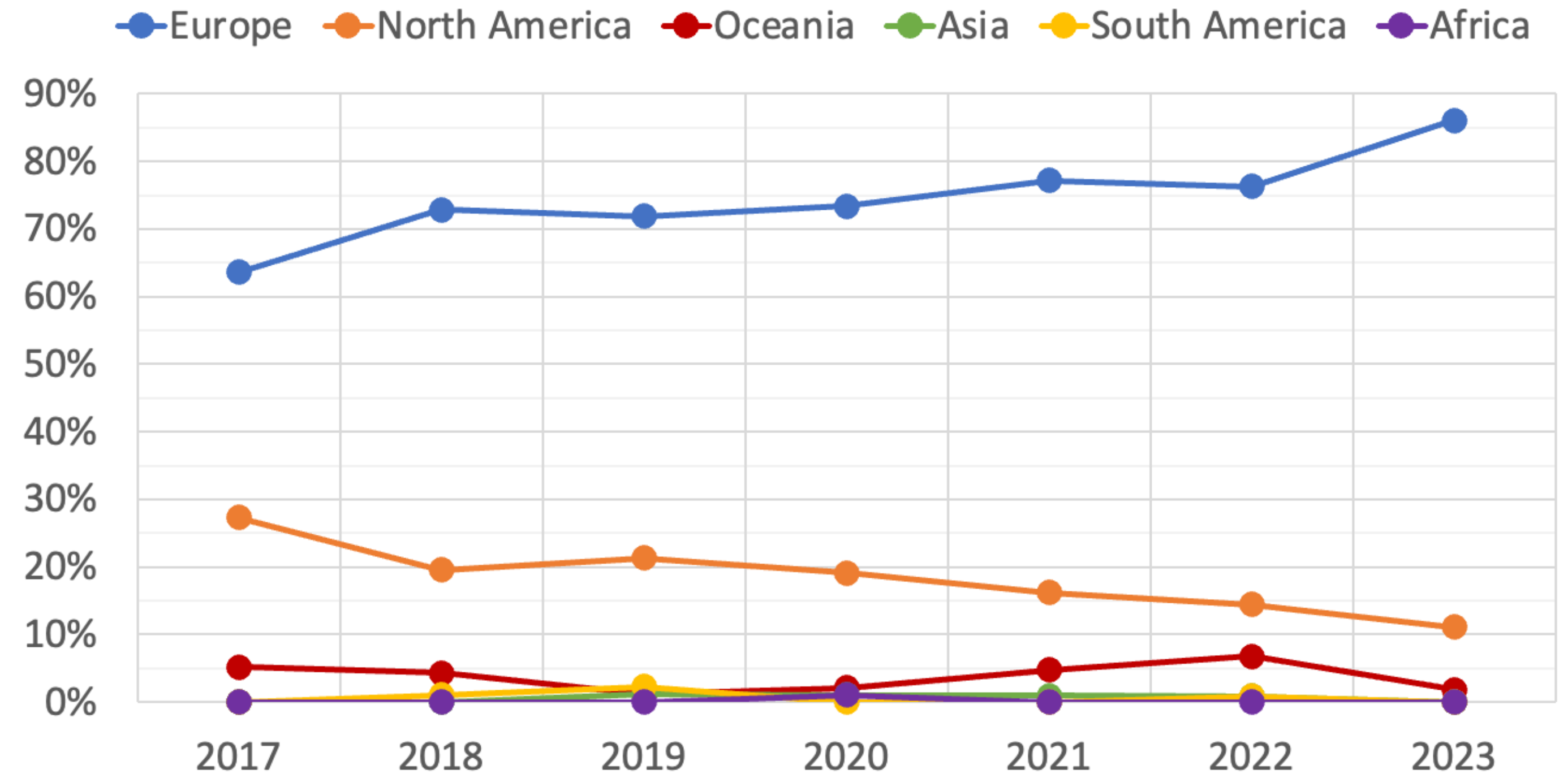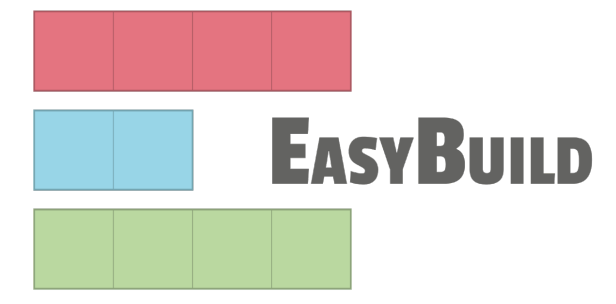

In which part of the world are you located?

Asia (1%)
Oceania (2%)
North America (11%)
Europe (86%)
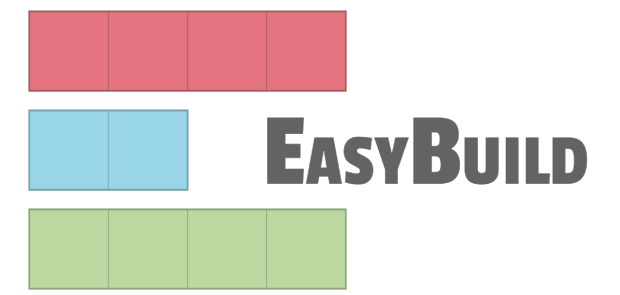
In which part of the world are you located?

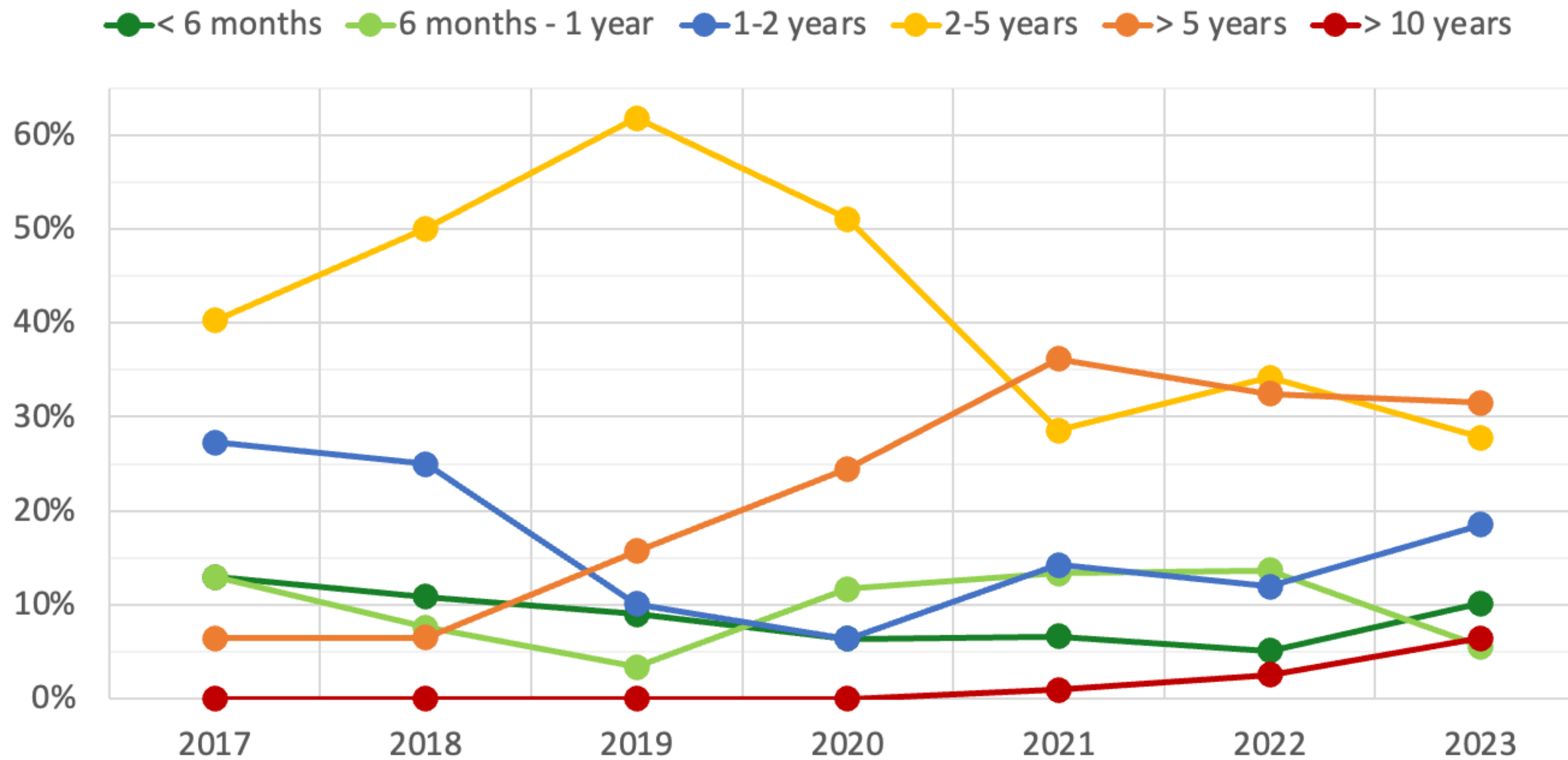Europe — North America — Oceania — Asia — South America — Africa

- EasyBuild community becoming (even more) European?

- No survey participants from South America / Africa

EasyBuild demographics

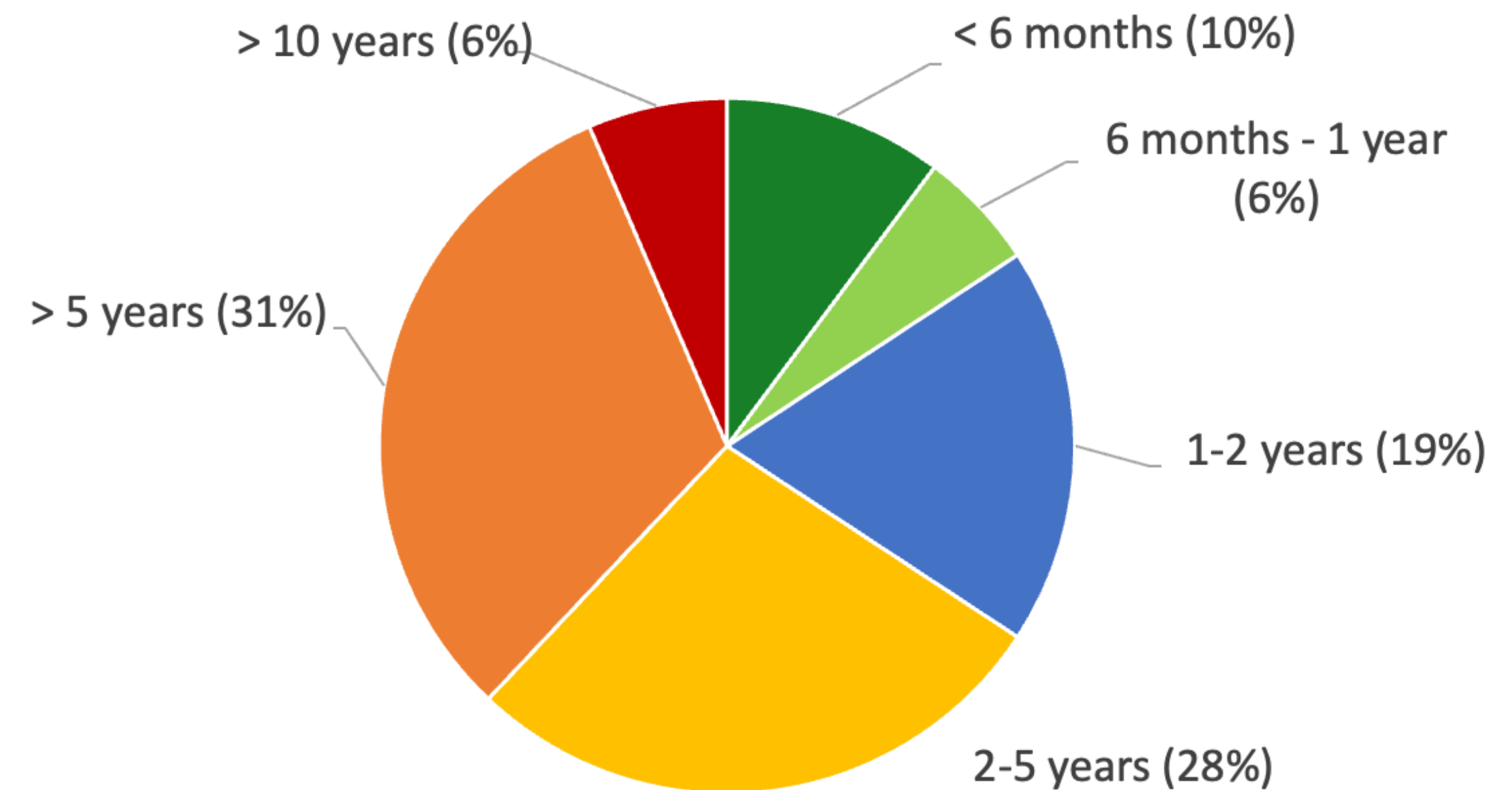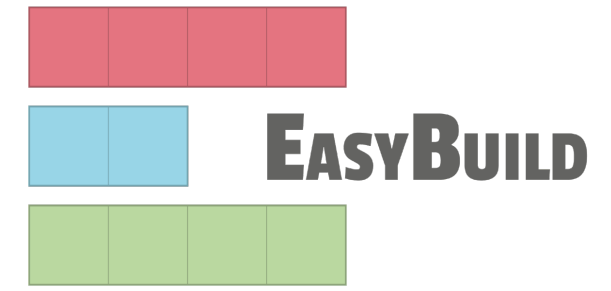*How long have you been using EasyBuild?*

Steady and significant influx of new people in past year: ~15% of current community

Minority group (~5%) with *over a decade* of experience with EasyBuild is starting to form…

Mix of new & experienced EasyBuilders

- We should be aware of people new to EasyBuild (or even installing software on HPC systems)

- They can learn a lot from experienced EasyBuild users, **but also the other way around!**

- We should try and look beyond the questions they ask: **what can we do to improve EasyBuild?**

- **Feedback by newcomers is very valuable**, experienced EasyBuilders "go blind" to certain things

- Influx of new people is partially fueled by EESSI & MultiXscale

    - They are funded to work with/on EasyBuild (but in a specific context)
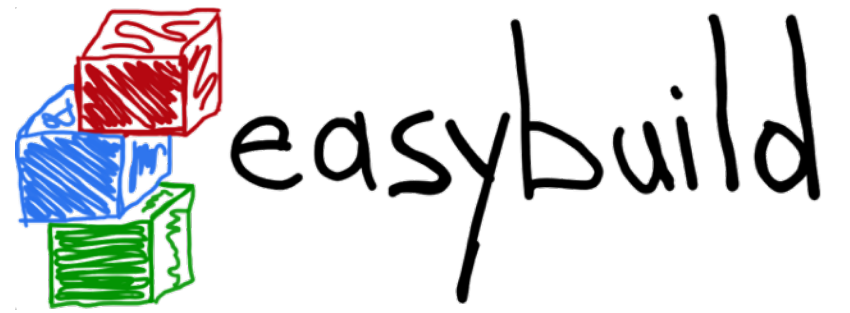
EasyBuild demographics
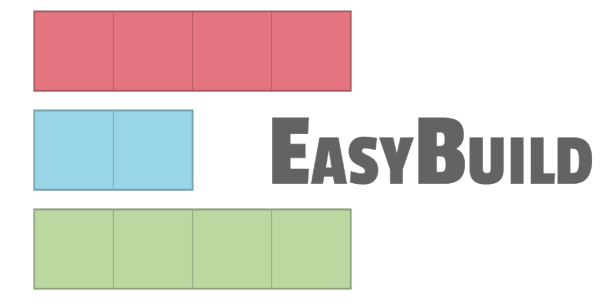
How did you first learn about EasyBuild?
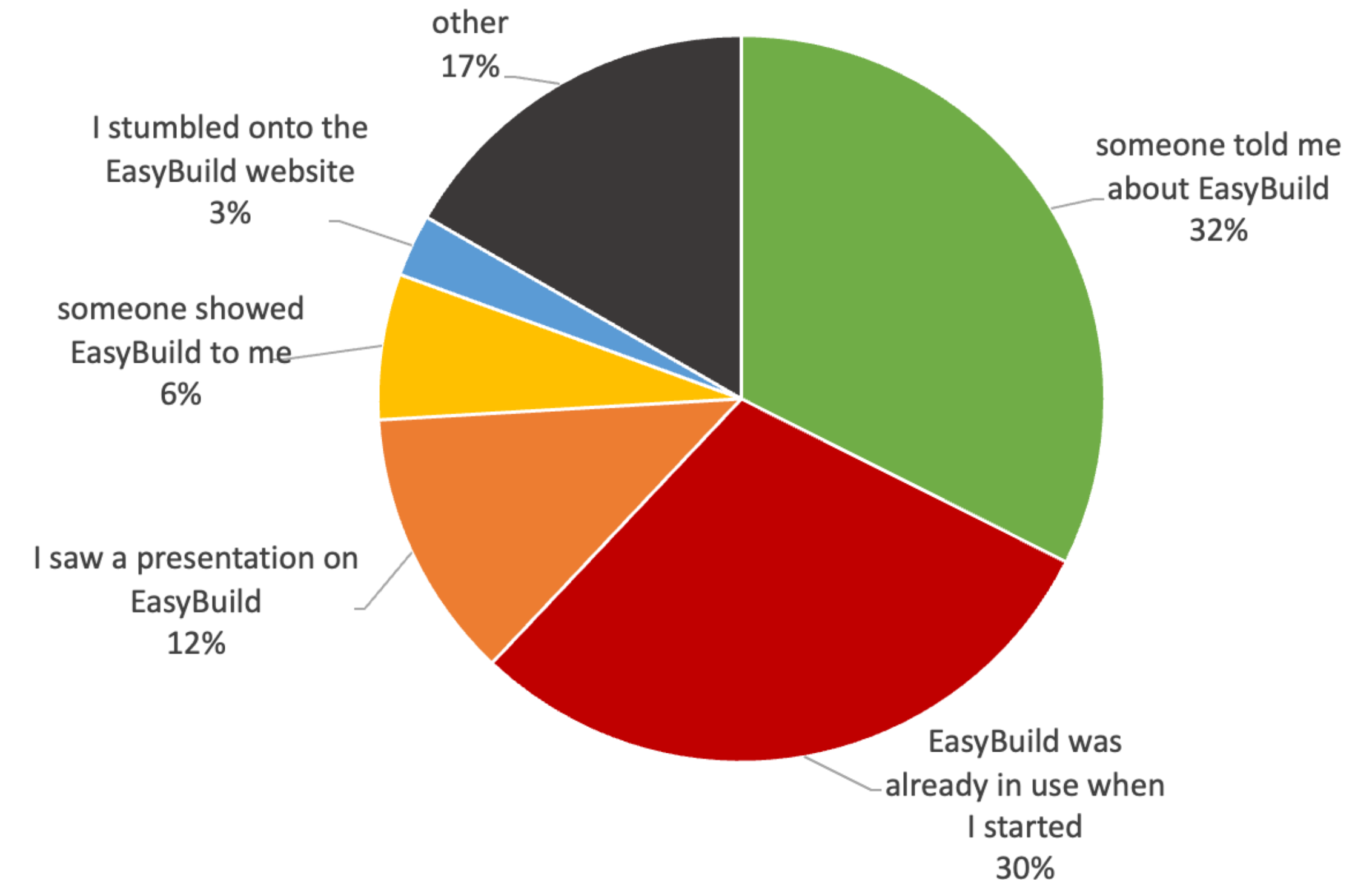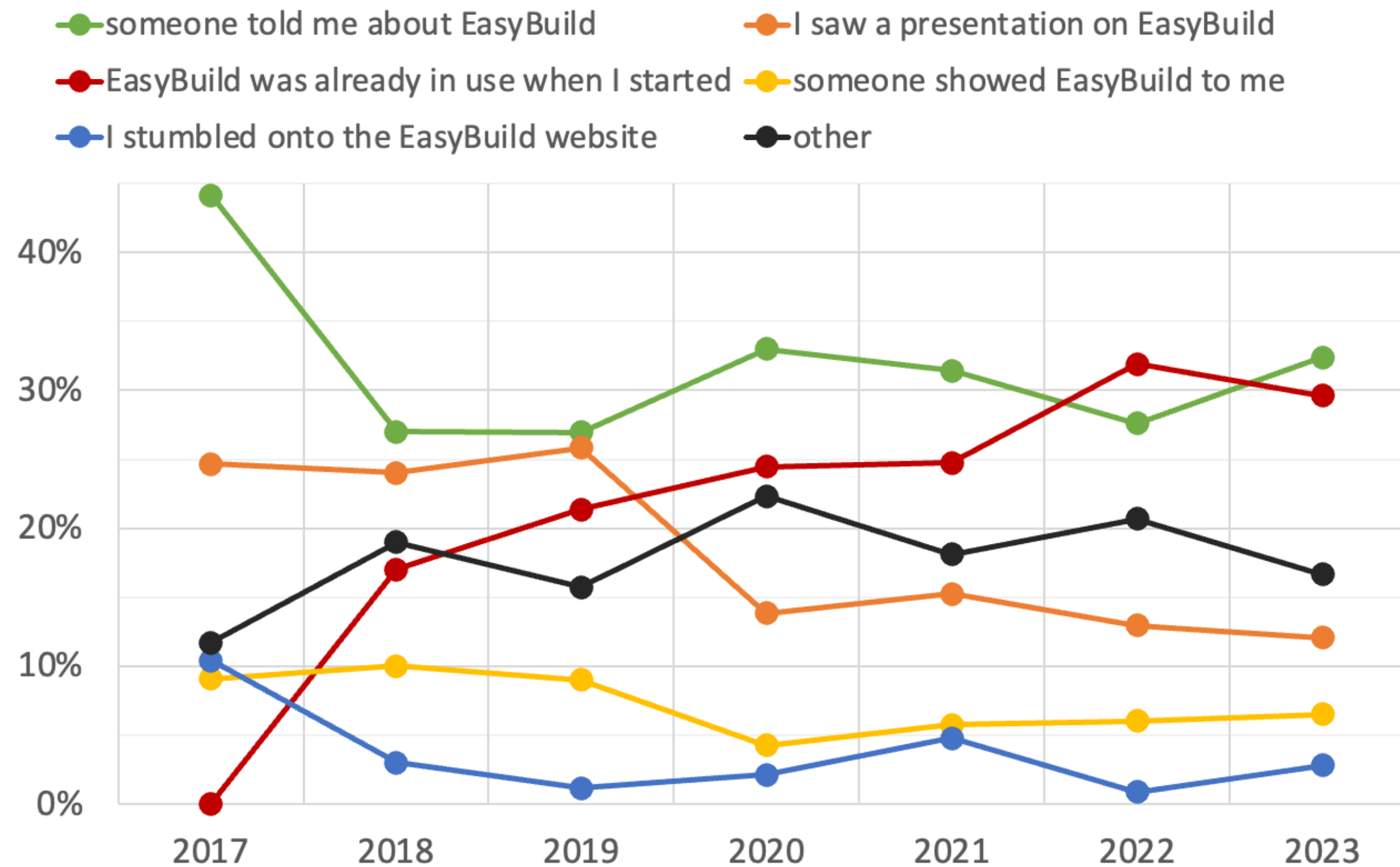
Word-of-mouth (peers, talks, demos, social media) remains dominant (>50%)

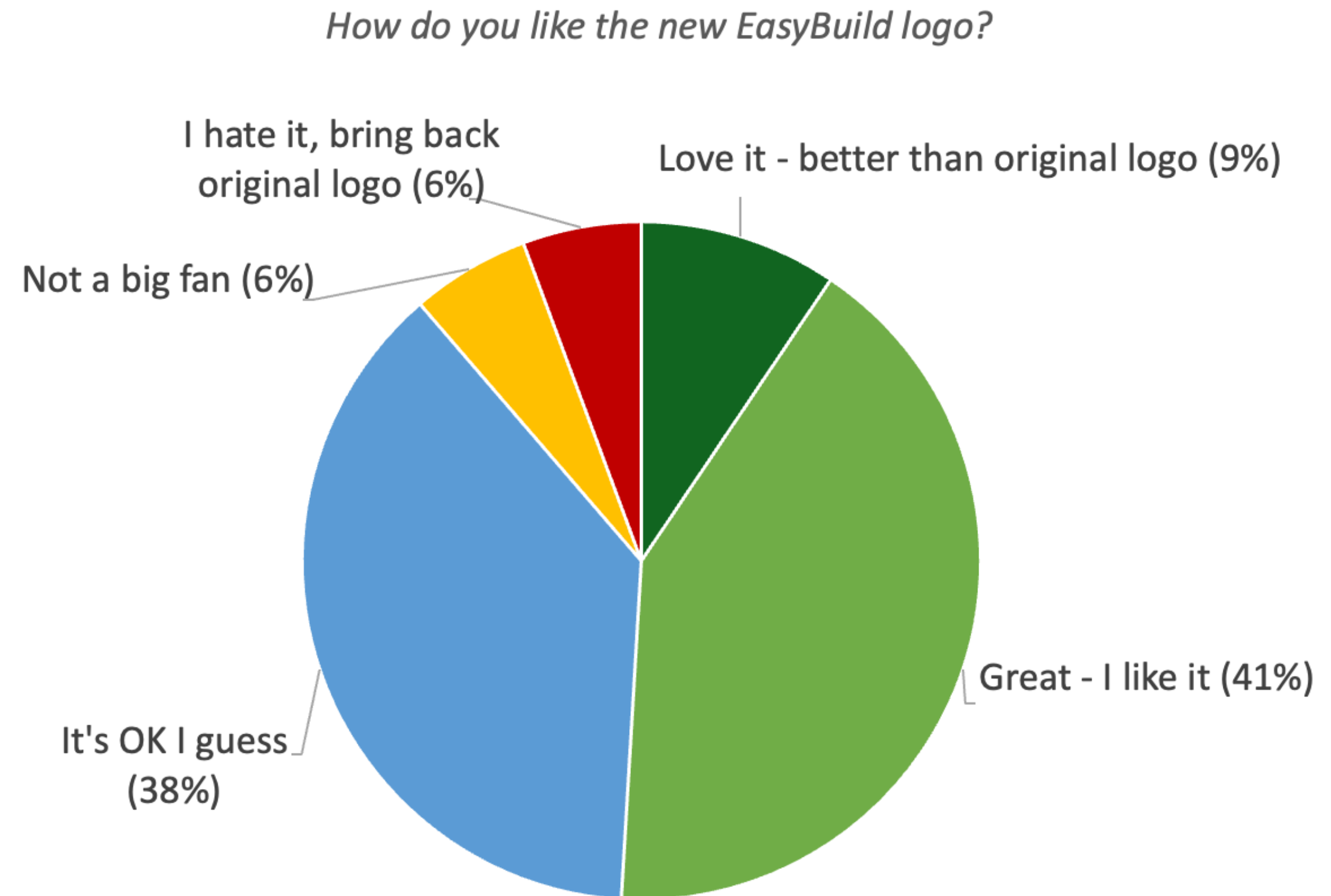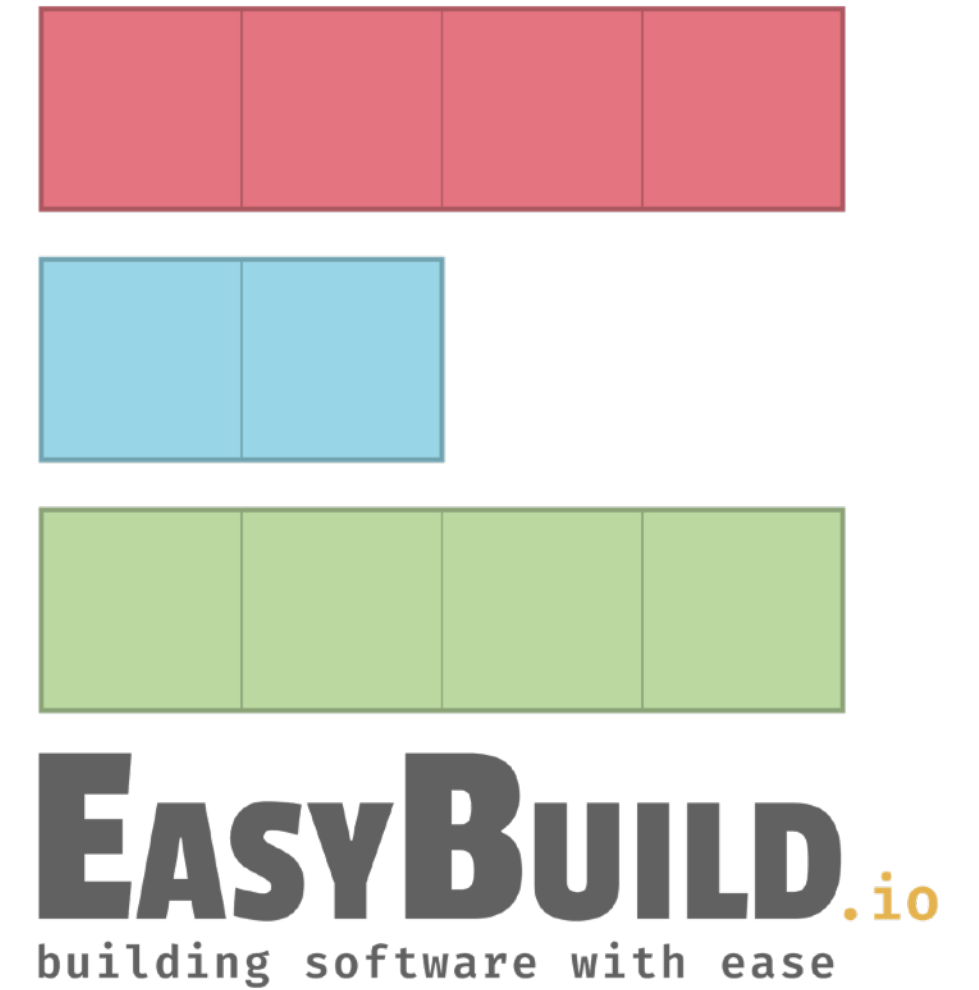People who learned about EasyBuild on the job is 2nd largest group

New(ish) EasyBuild logo

A new logo for EasyBuild was introduced on 14 Nov 2022,
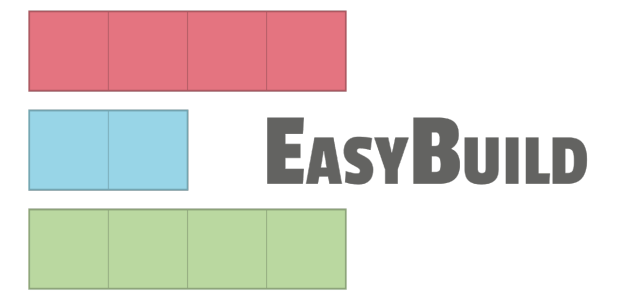
exactly 10 years after the release of EasyBuild v1.0 .
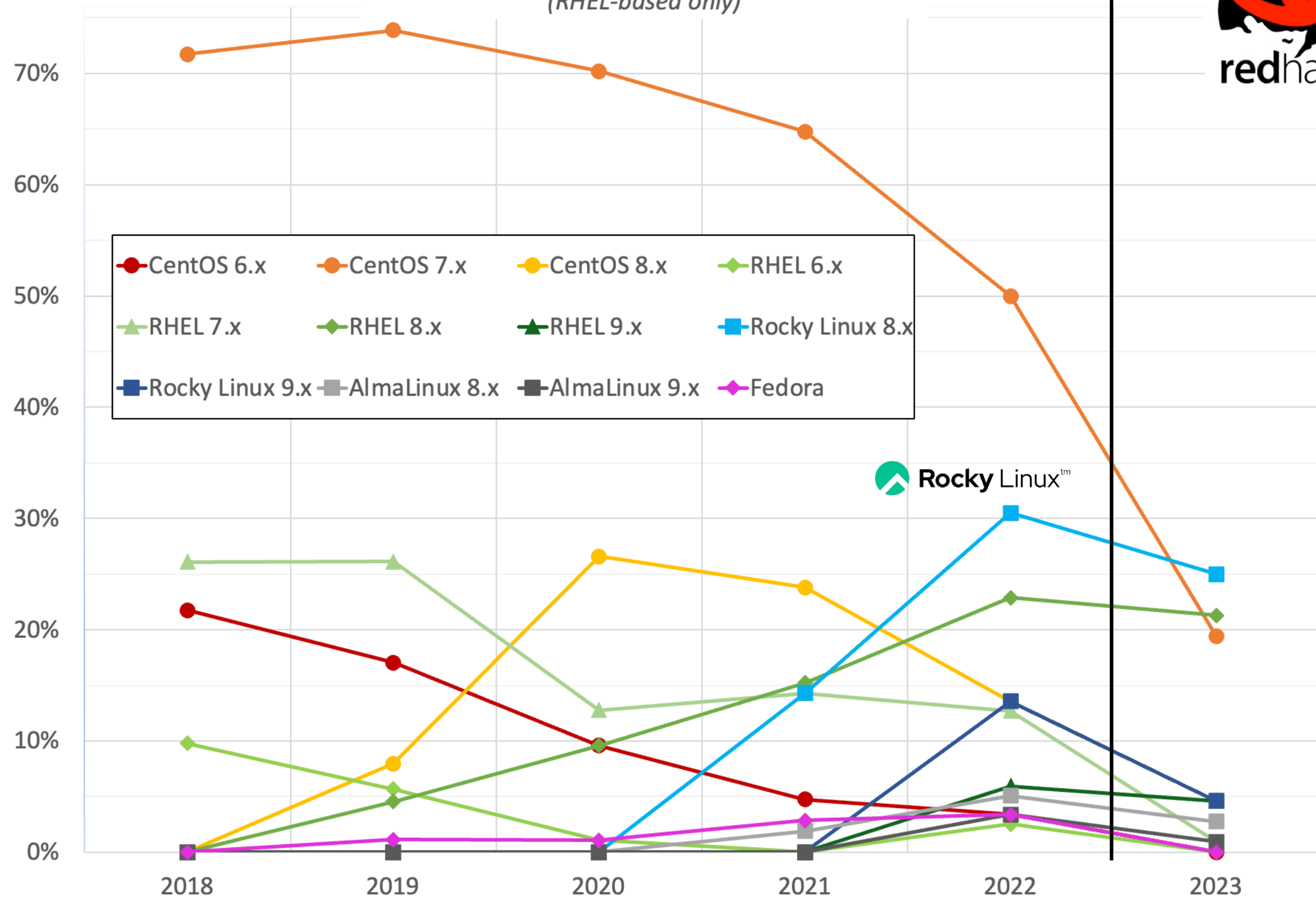
More information on (the motivation for) new logo:

https://easybuild.io/new-logo-2022.html

New logo "approved" by majority,

small group (~12%) doesn't like it.

How do you like the new EasyBuild logo?

I hate it, bring back original logo (6%)

Not a big fan (6%)

Love it - better than original logo (9%)

Great - I like it (41%)

It's OK I guess (38%)

Operating systems in EasyBuild community

On which operating system(s) do you use EasyBuild?
(RHEL-based only)

Legend: CentOS 6.x, CentOS 7.x, CentOS 8.x, RHEL 6.x, RHEL 7.x, RHEL 8.x, RHEL 9.x, Rocky Linux 8.x, Rocky Linux 9.x, AlmaLinux 8.x, AlmaLinux 9.x, Fedora
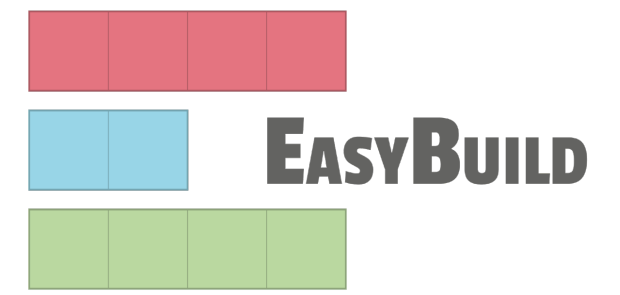
- Vast majority of EasyBuild community is using a RHEL-based OS

- **CentOS 7 is losing ground quickly** (EOL June 2024), but still at ~19%

- Rocky Linux 8 (~25%) is replacing CentOS 7, along with RHEL 8 (~21%)

- CentOS 6 + RHEL 6 are (finally) gone…

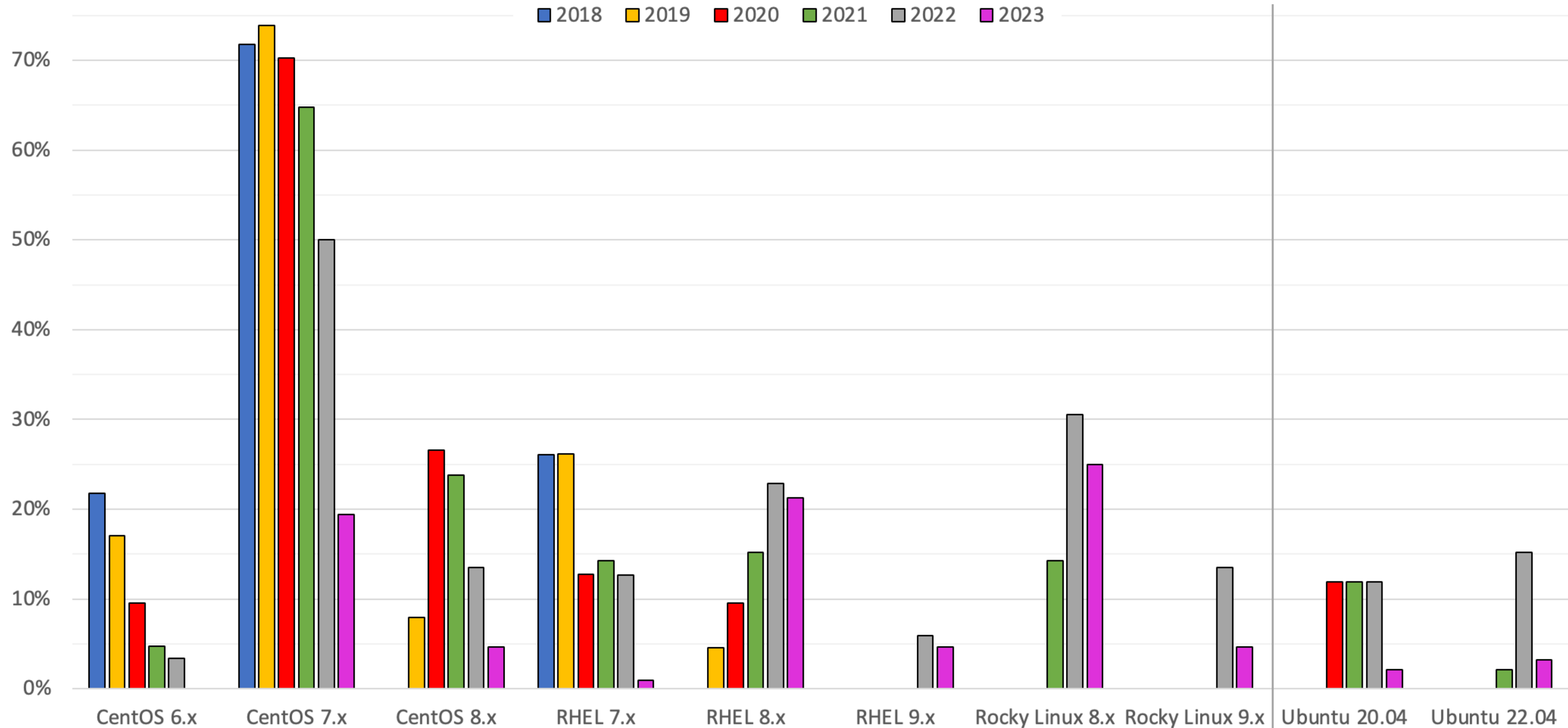- Adooption of RHEL9-based OS still (very) low (~5% RHEL 9, ~5% Rocky Linux 9)

Note: in this survey (2023 results),
*you could only pick *one* option*

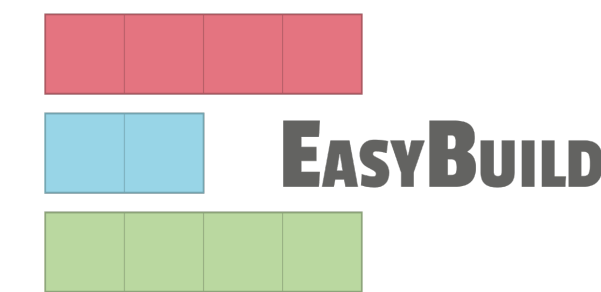16

Operating systems in EasyBuild community

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

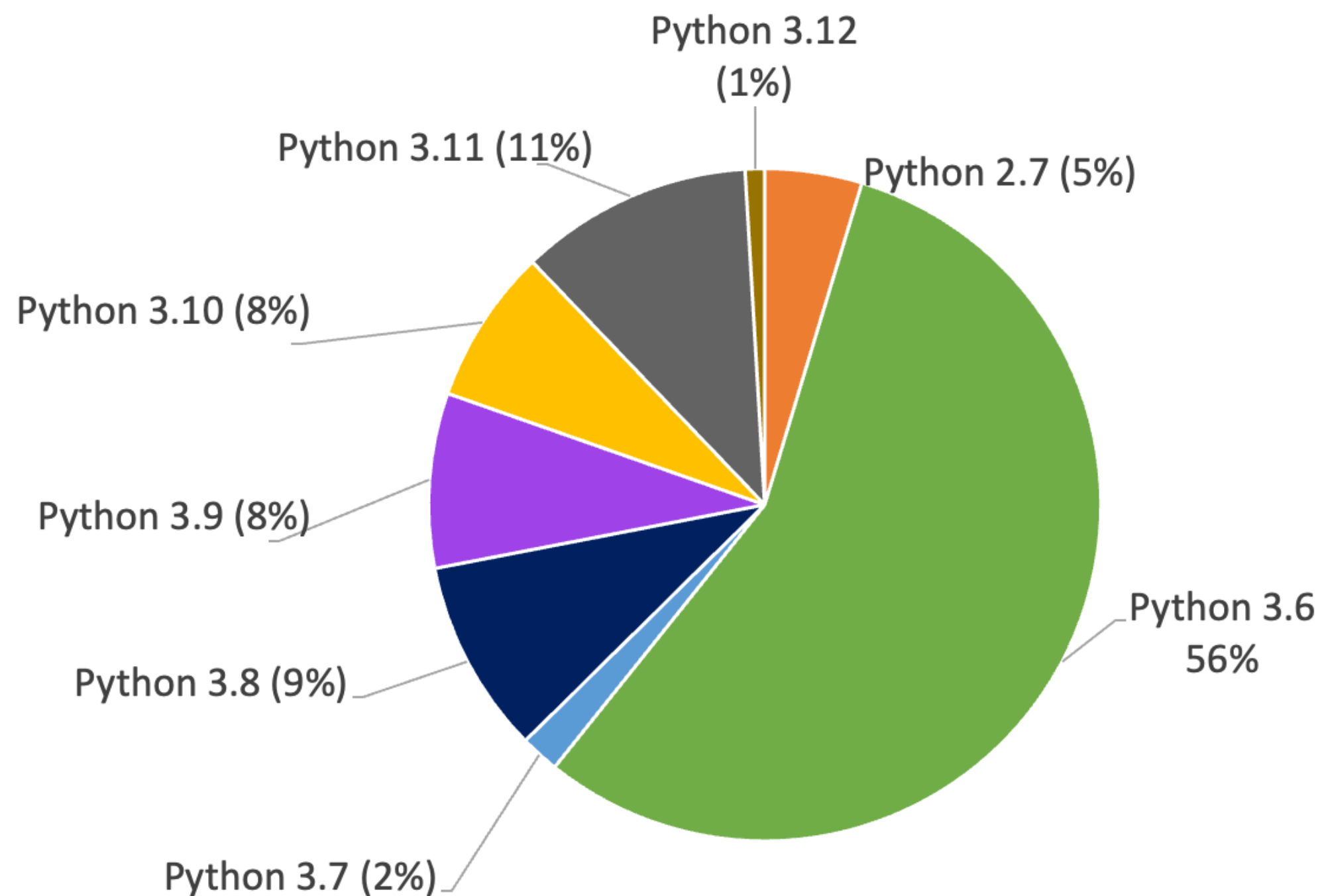■ 2018 ■ 2019 ■ 2020 ■ 2021 ■ 2022 ■ 2023

Non-RHEL based OSs are a minority in the EasyBuild community, only Ubuntu (was) > 10%

Python versions used to run EasyBuild

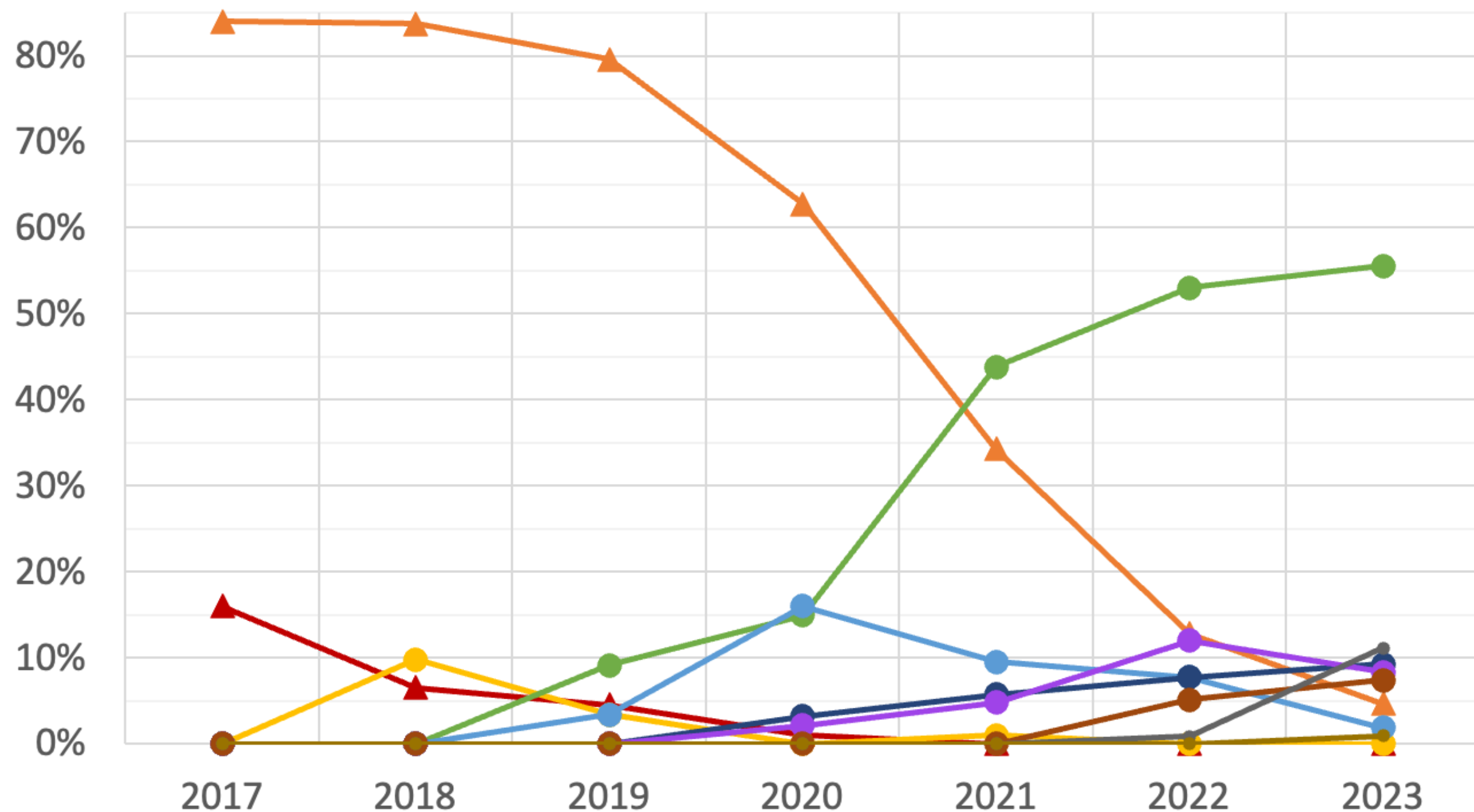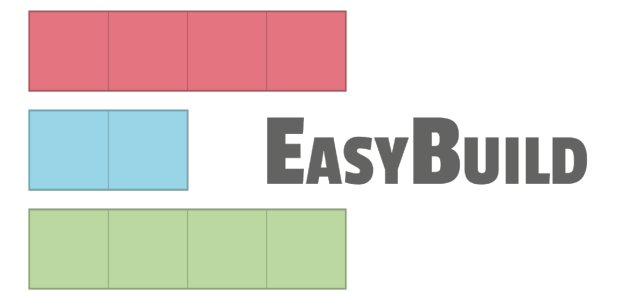Which Python version do you usually use to run EasyBuild?

The time is right to stop supporting running EasyBuild on top of Python 2.7 (EOL since 2020!)

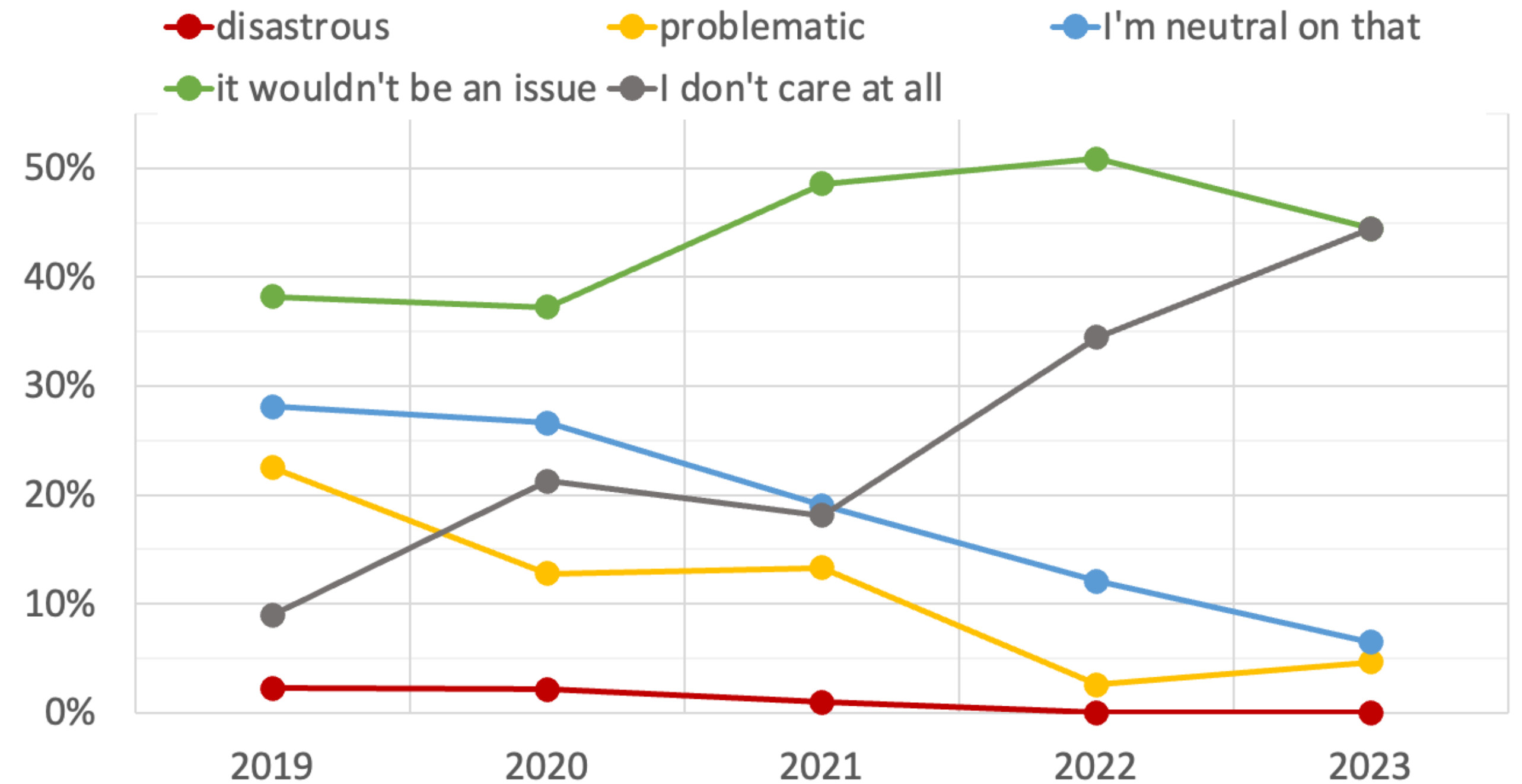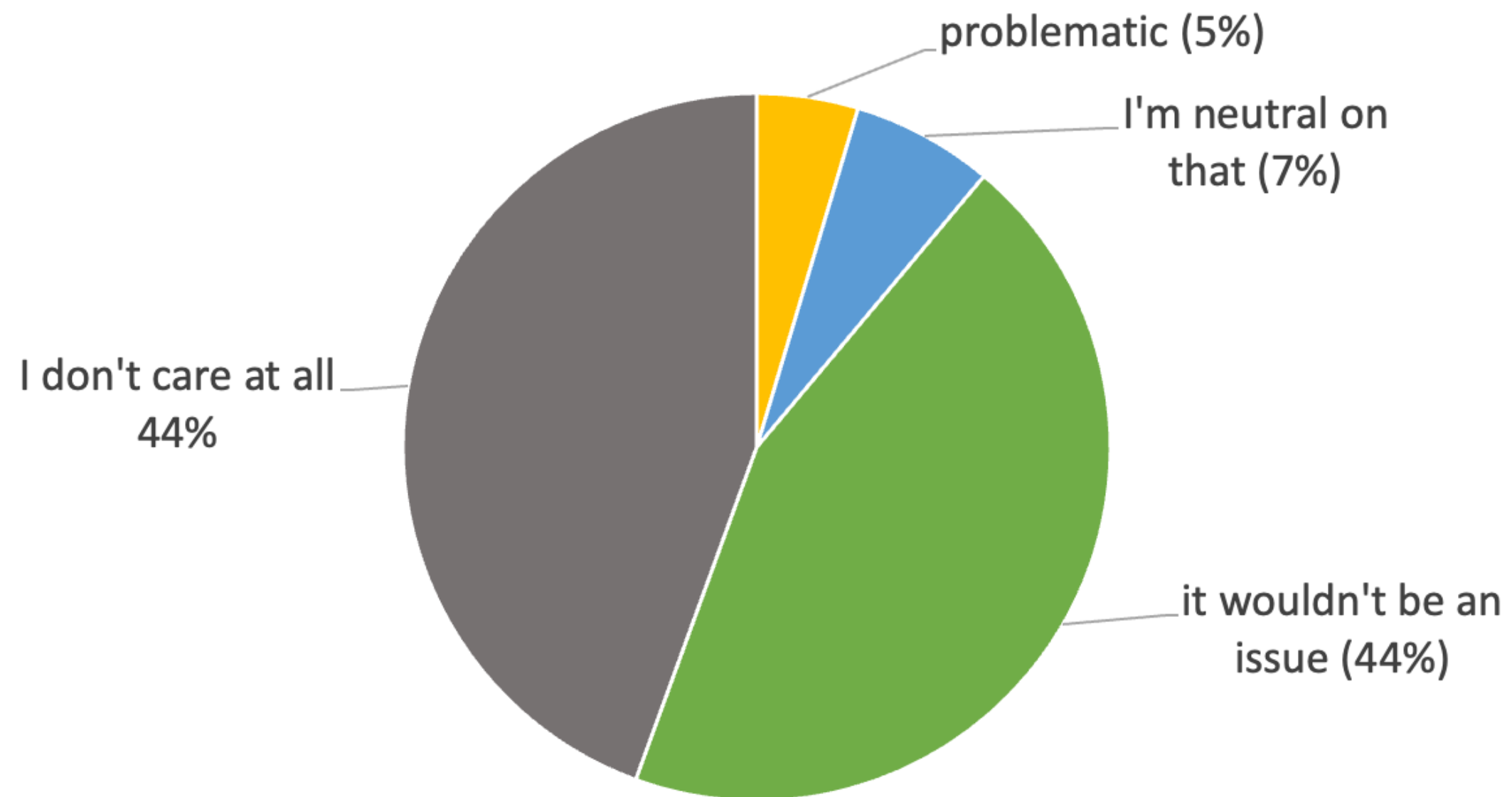Keeping support for using Python 3.6 (default Python in RHEL8-based OSs) is important!

Splintering across various Python 3 versions is a challenge going forward…

Community is fine with getting rid of Python 2.7

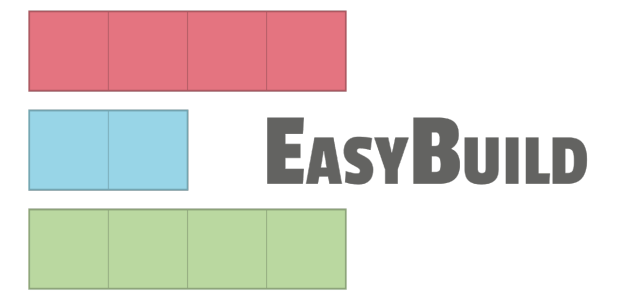*How troublesome would it be for you if EasyBuild becomes incompatible with Python 2.7?*

Support for running on top of Python 2.7 is deprecated since EasyBuild v4.7.0 (Jan'23),

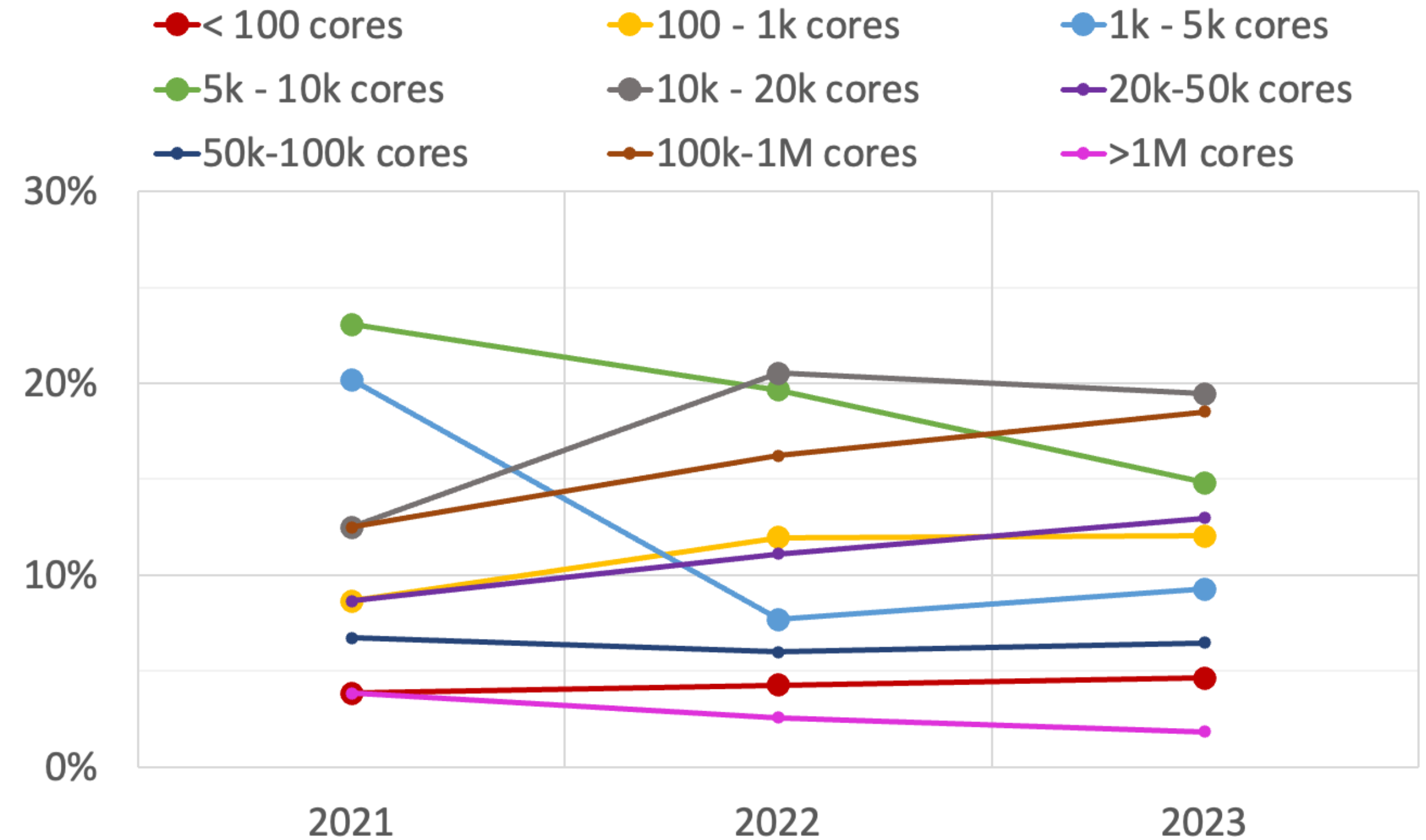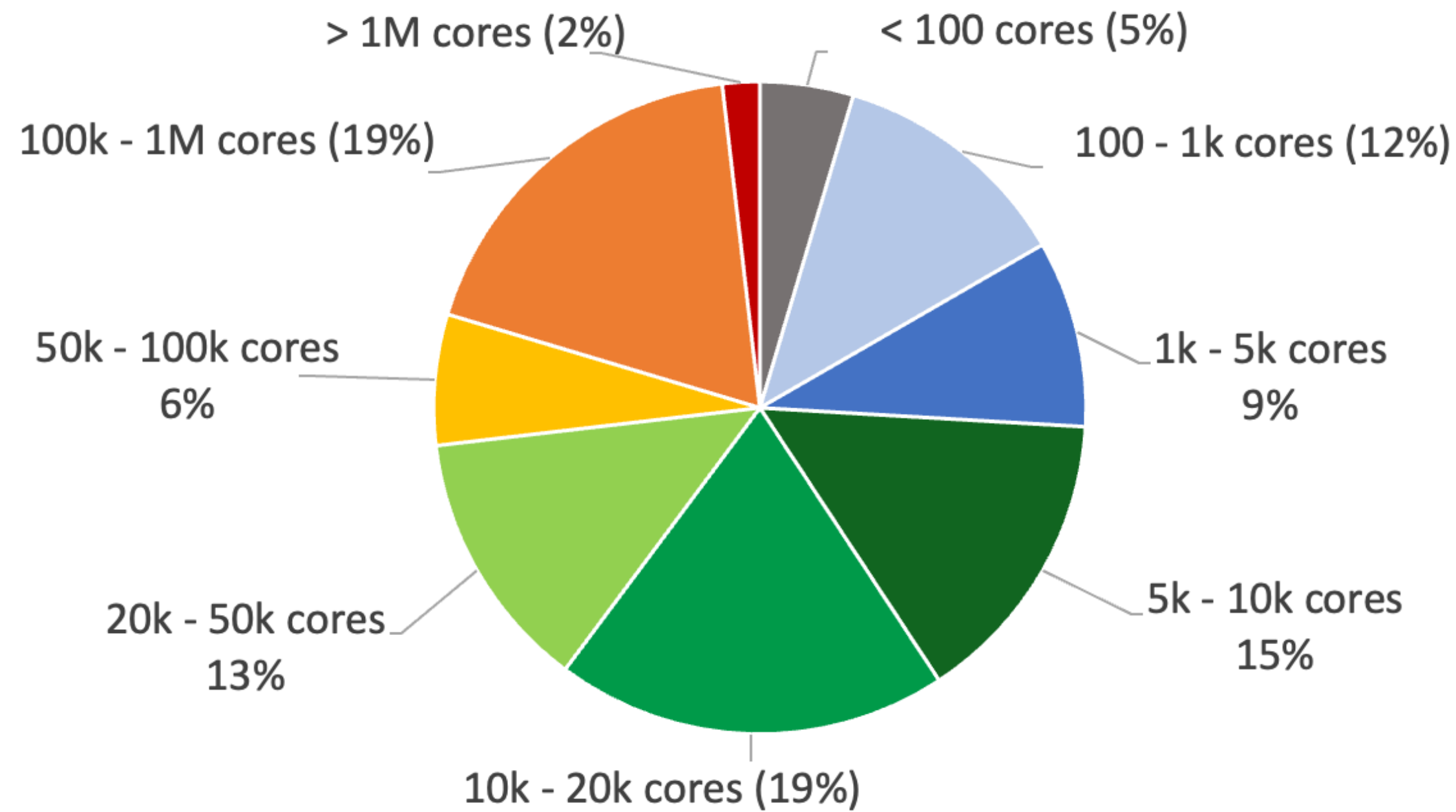and **will be *removed* in upcoming EasyBuild v5.0**.

95% of EasyBuild community is fine with that.

System sizes (CPU cores)

What's the aggregate core count for the HPC systems you manage, support, or use?
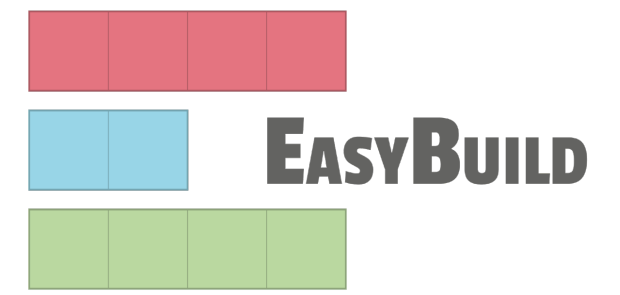

Huge mix of system sizes, minor shift to bigger systems
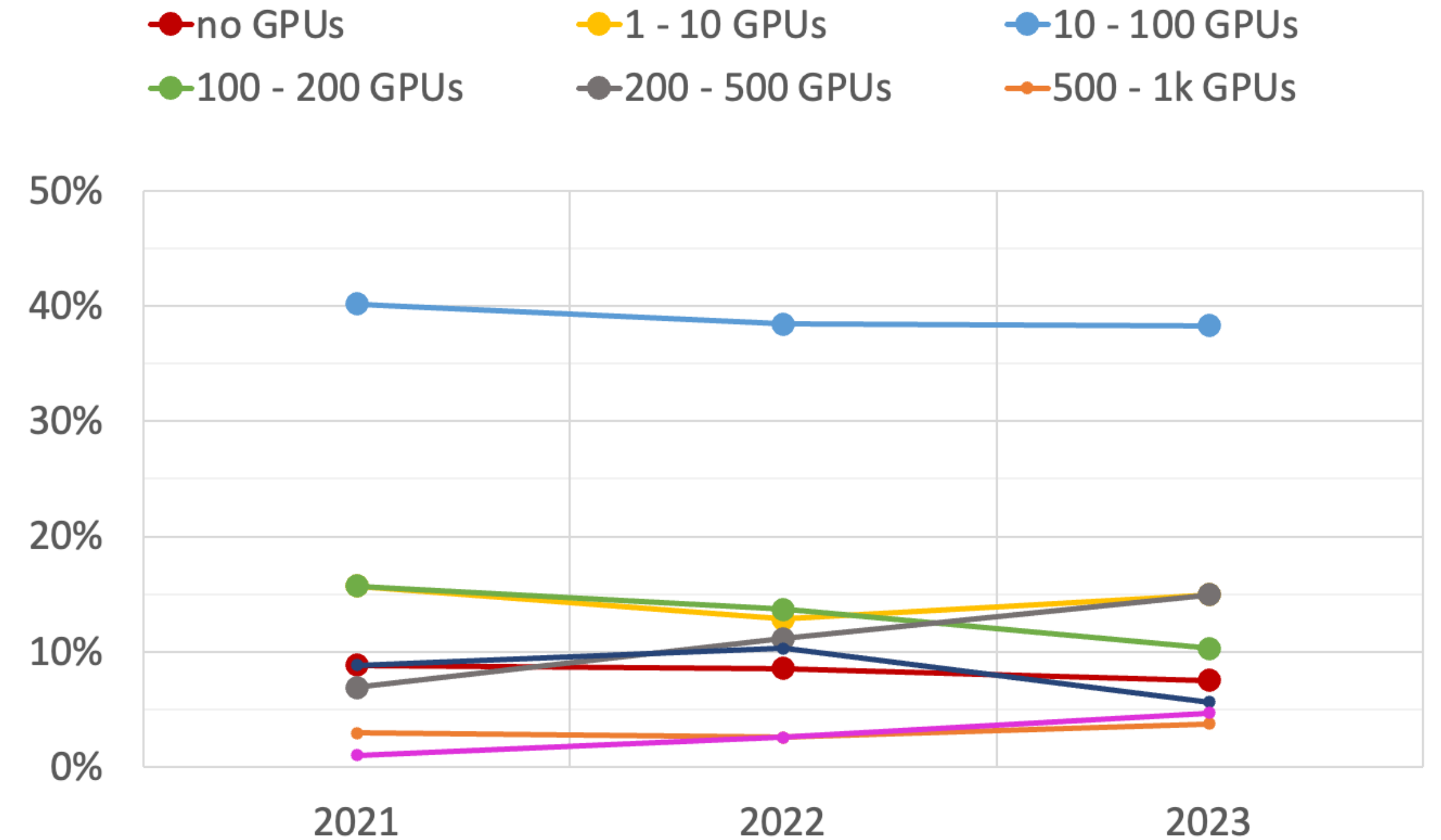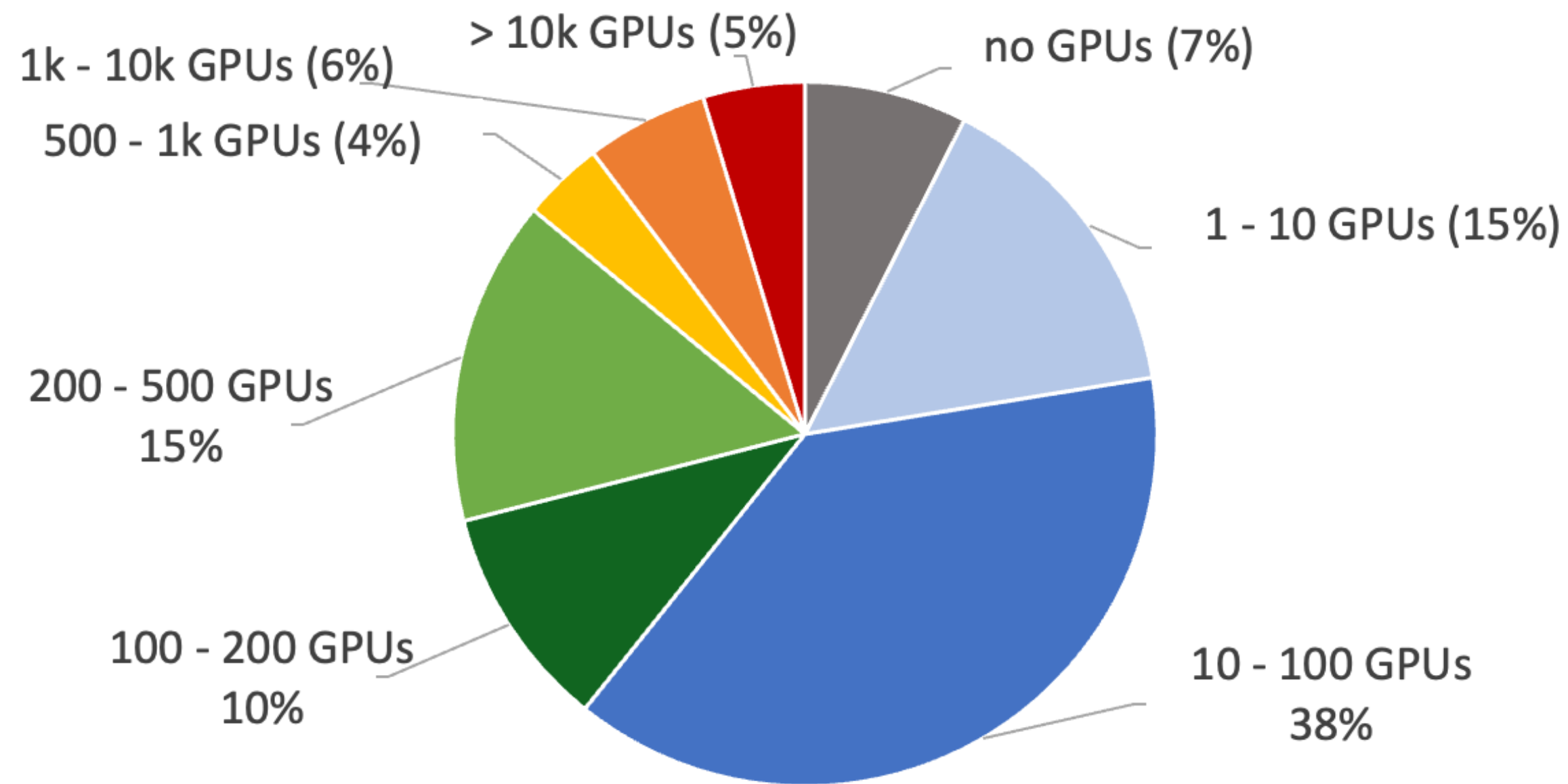
<10k cores: ~41%, 10k-100k cores: ~39%, >100k: ~20% cores

System sizes (GPUs)

What's the aggregate GPU count for the HPC systems you manage, support, or use?

Pie chart:
- no GPUs (7%)
- 1 - 10 GPUs (15%)
- 10 - 100 GPUs 38%
- 100 - 200 GPUs 10%
- 200 - 500 GPUs 15%
- 500 - 1k GPUs (4%)
- 1k - 10k GPUs (6%)
- > 10k GPUs (5%)

Line chart legend:
- no GPUs
- 1 - 10 GPUs
- 10 - 100 GPUs
- 100 - 200 GPUs
- 200 - 500 GPUs
- 500 - 1k GPUs

Majority of sites has "only" 1-100 GPUs (~53%), ~25% has 100-500 GPUs (both relatively steady trend)

Steady 7-8% has no GPUs

Processor families (Intel)

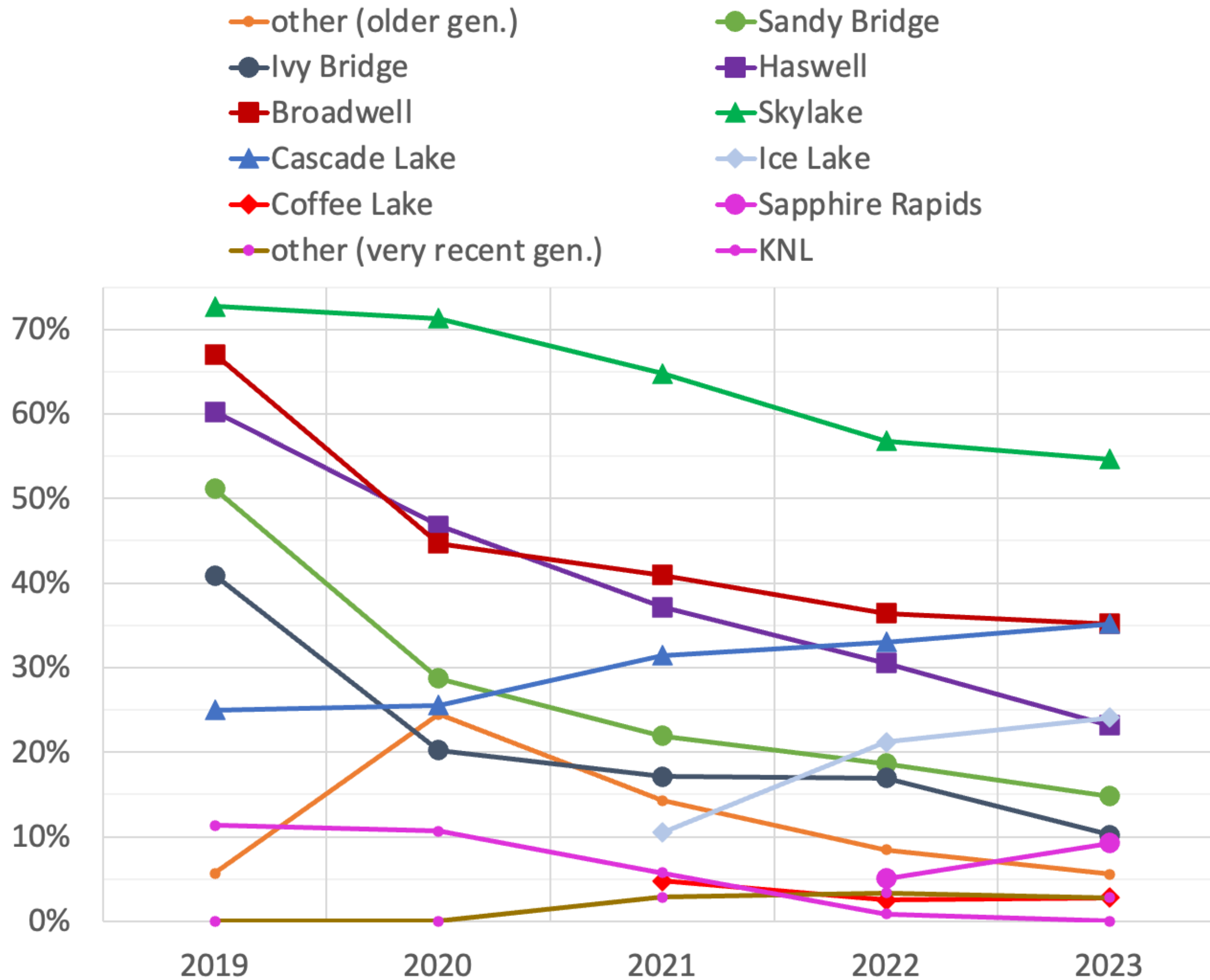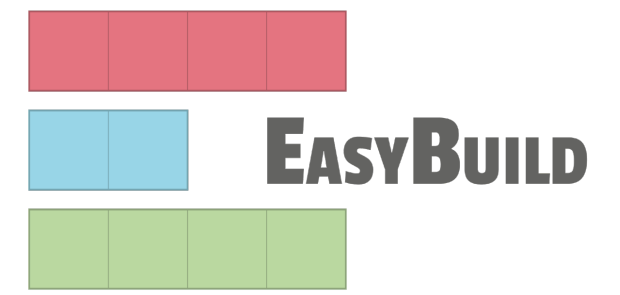*For which processor families are you installing software using EasyBuild?*

Legend:
- other (older gen.)
- Ivy Bridge
- Broadwell
- Cascade Lake
- Coffee Lake
- other (very recent gen.)
- Sandy Bridge
- Haswell
- Skylake
- Ice Lake
- Sapphire Rapids
- KNL

- Skylake is *still* most popular at 55%, slowly declining...

- Haswell still at ~23%, Broadwell ~35%

- Ivy Bridge, Sandy Bridge: **still** at 10-14%…

- On the rise (slowly): Cascade Lake (35%), Ice Lake (24%)

- Still a small part *older than Sandy Bridge* (~6%)…

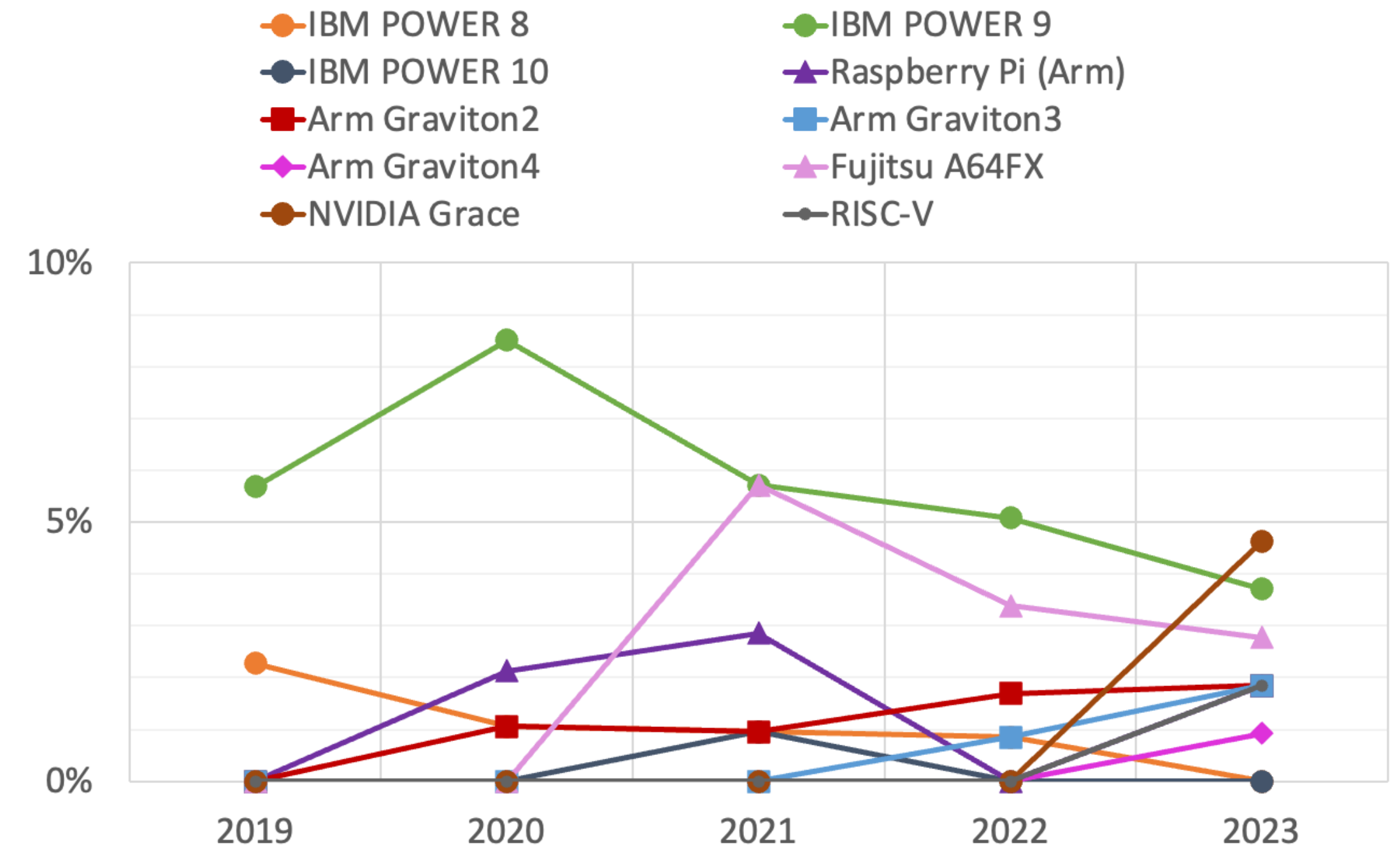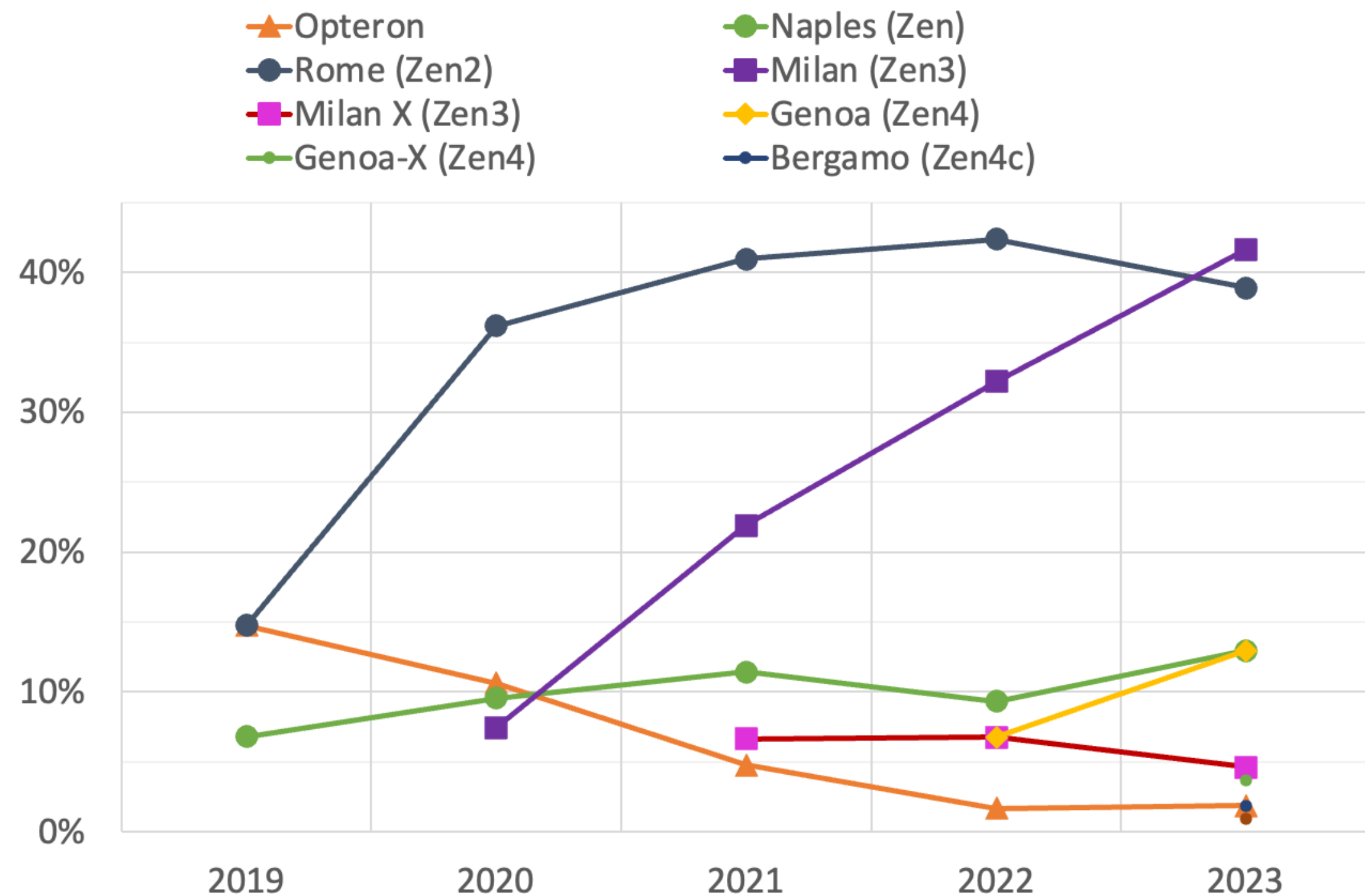- Sapphire Rapids is new kid on the block (now ~9%)
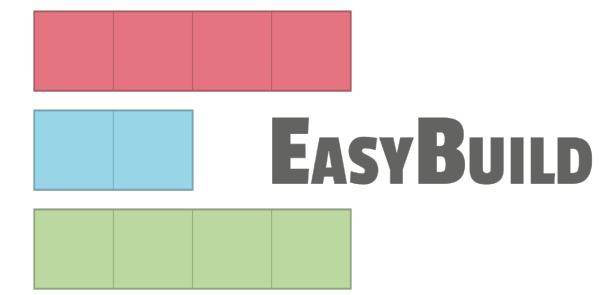
Processor families (non-Intel)

For which processor families are you installing software using EasyBuild?

- AMD Milan (Zen3, ~42%) surpassed Rome (Zen2, ~39%)

- Genoa (Zen4) rises to ~13%, Naples (Zen) still at ~13%

- IBM POWER 9 is still around (sort of, ~4%)

- Arm: mostly NVIDIA Grace, but still < 5%
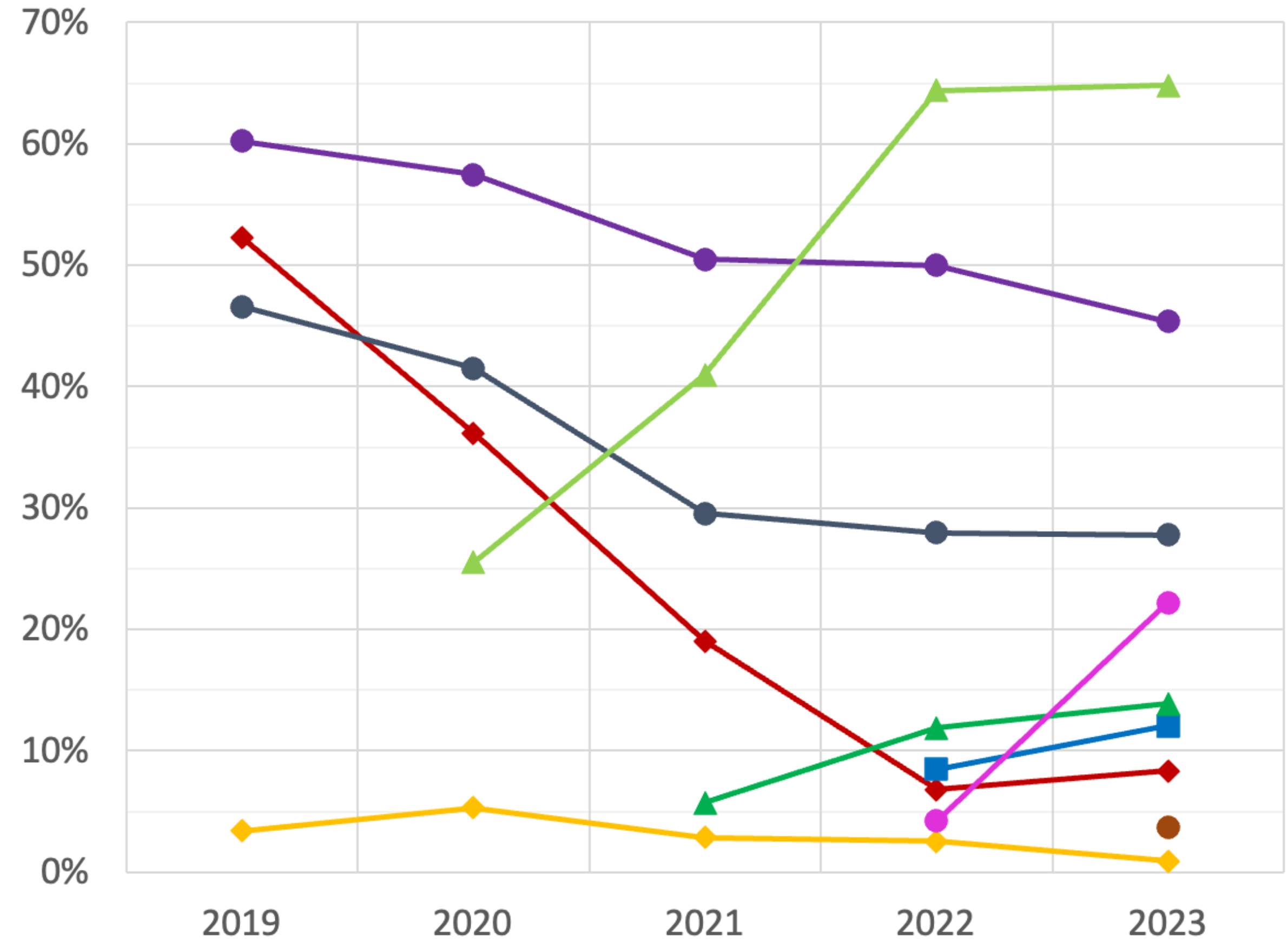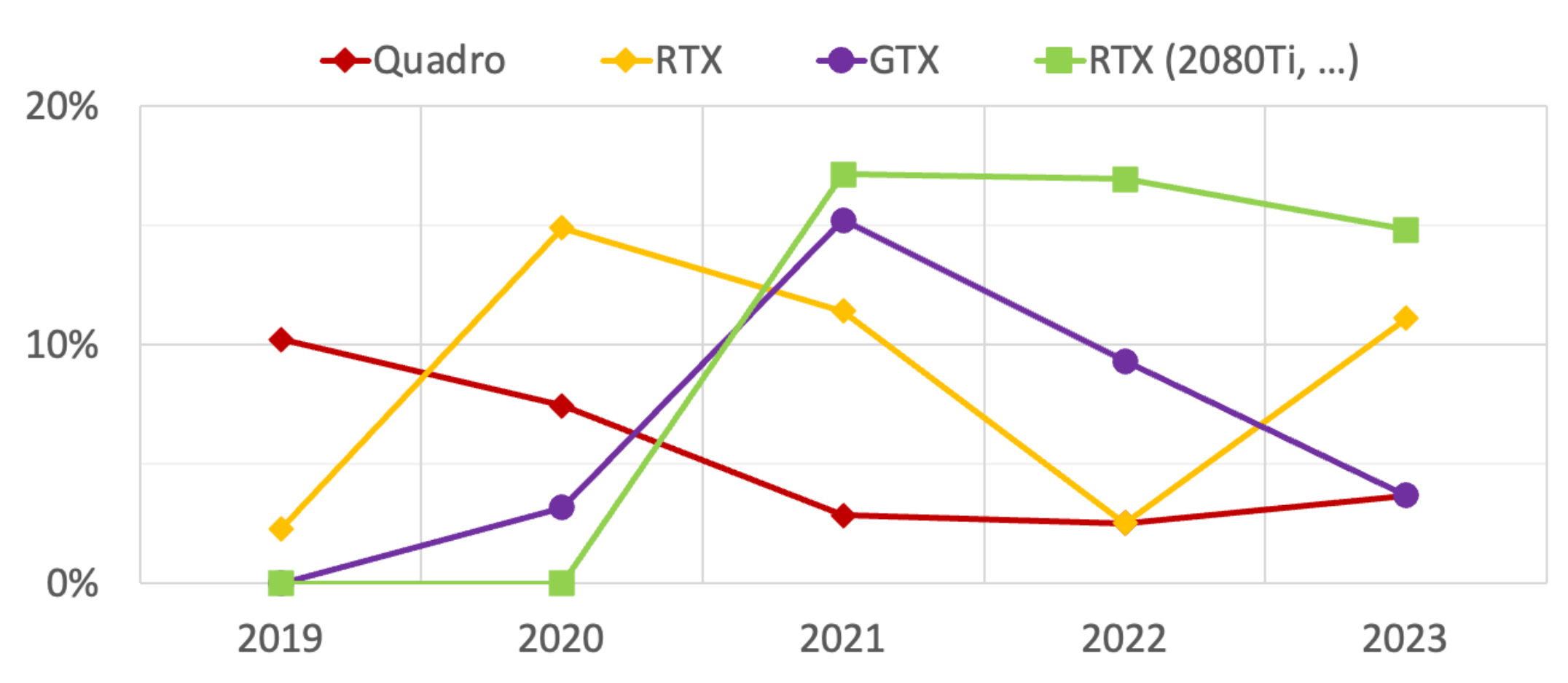
- No RISC-V (yet), but coming soon?

Accelerators (NVIDIA GPUs)

For which accelerators are you installing software using EasyBuild?
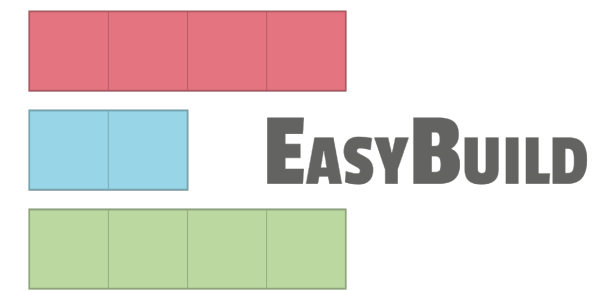

- Ampere (A100, A30) still most popular GPU family at ~64%

- V100 slowly decreasing, still ~45%

- Kepler almost fully retired, P100 still hanging on at 28%

- Moderate interest in RTX family and other Ampere GPUs

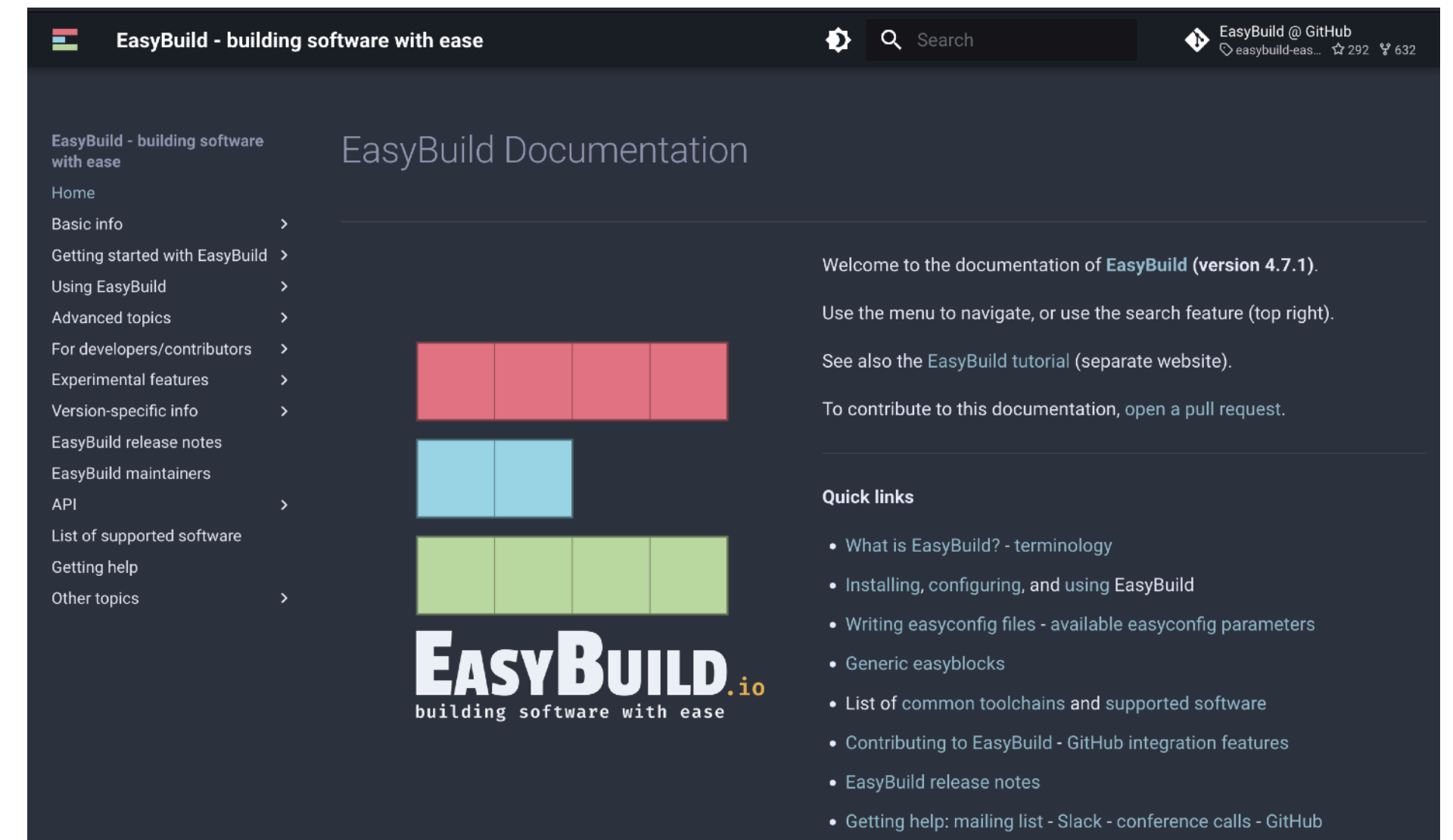- **AMD GPUs only ~5%**, Intel GPUs < 1% *(not shown in plot)*

Revamped EasyBuild documentation (since Jan'23)

https://docs.easybuild.io

How do you like the revamp of the EasyBuild documentation?

- Fantastic: faster search, better structure, etc. (20%)
- Great - I like it (36%)
- Significant improvement, but can be improved (20%)
- Don't see much difference (6%)
- Disappointed, was expecting more (3%)
- I like old docs better (1%)
- Haven't looked enough to form an opinion (14%)
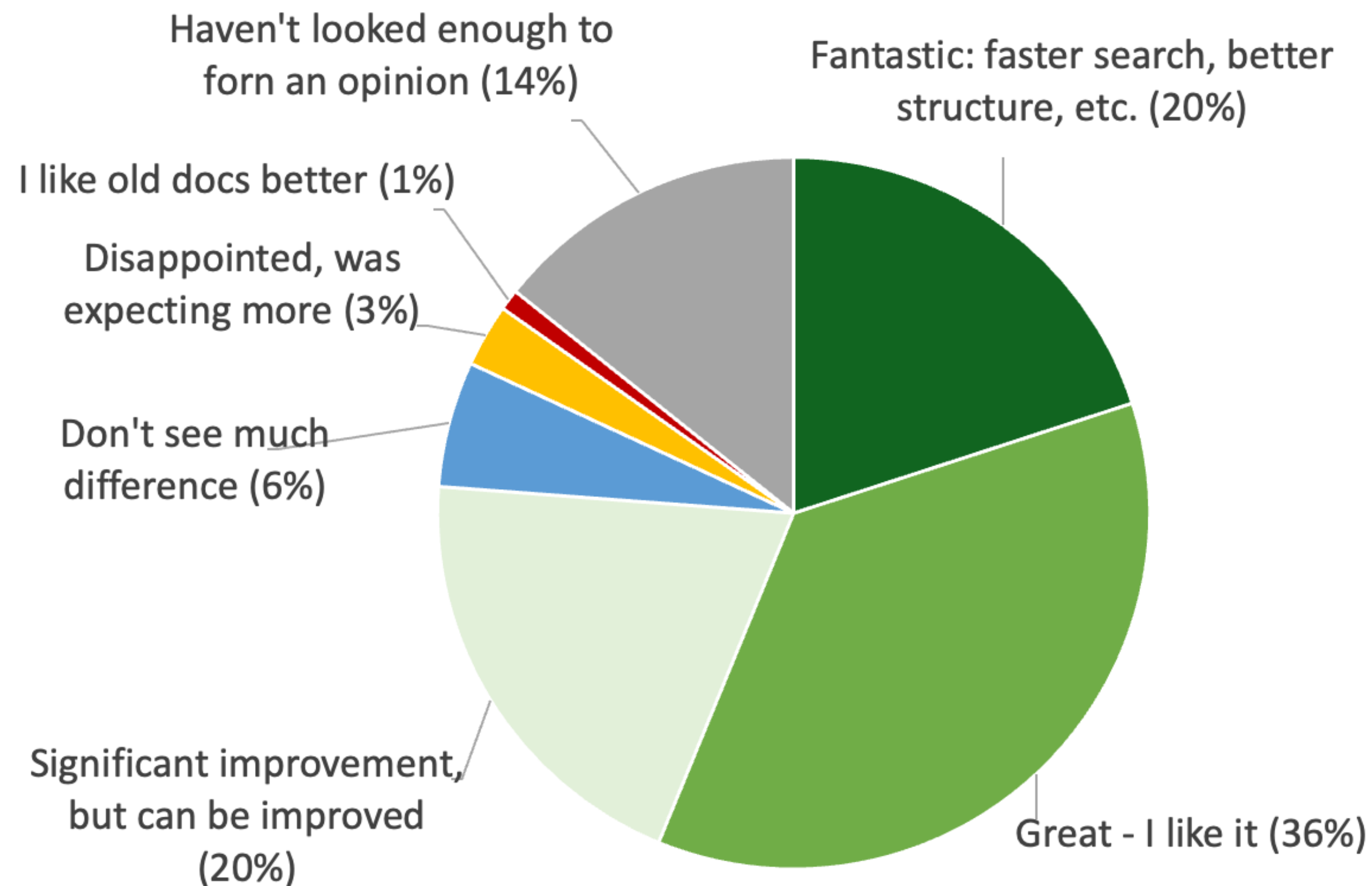
- Majority (~56%) likes revamped EasyBuild docs

- Increase in group that's critical (~10% → ~25%)

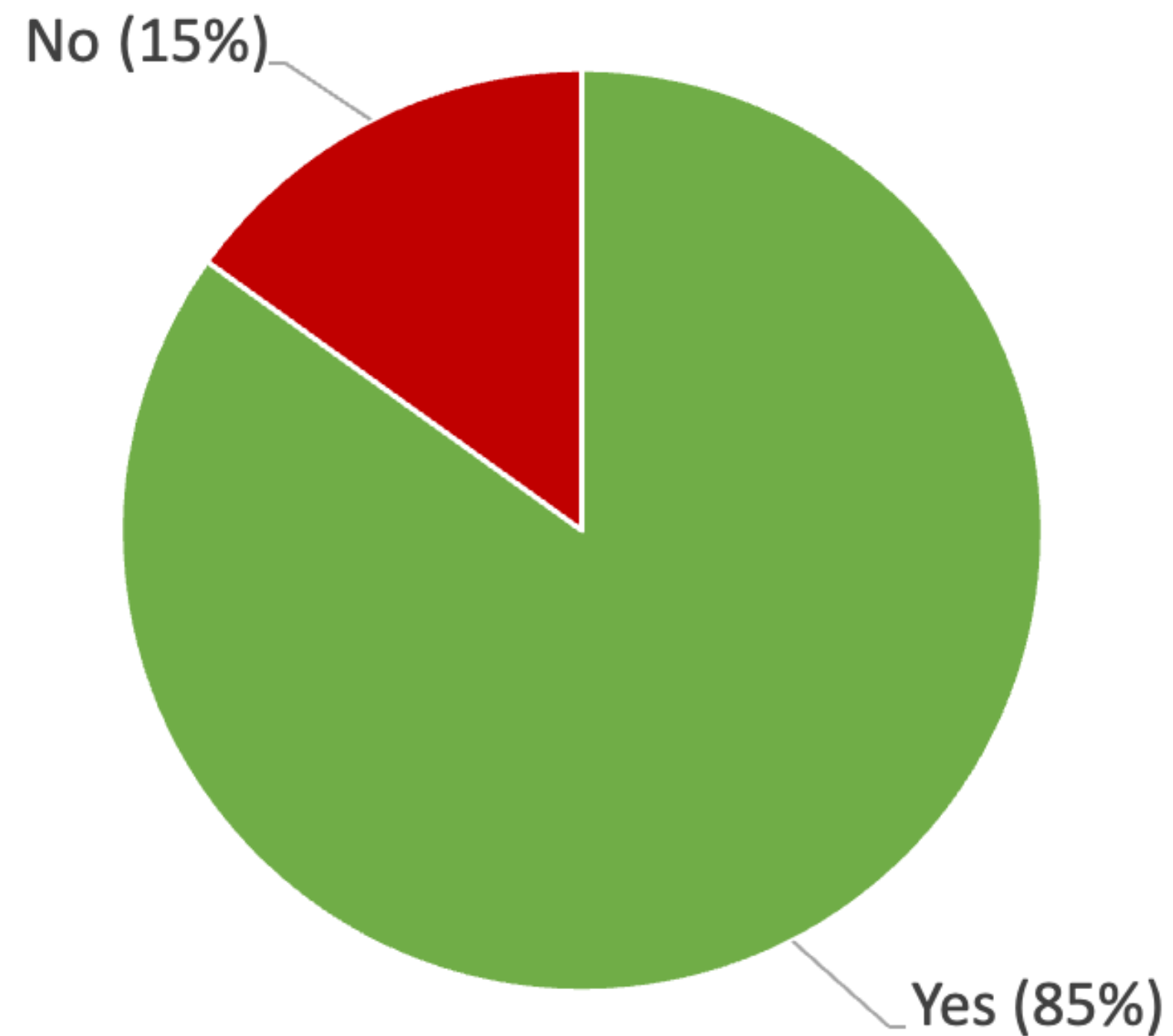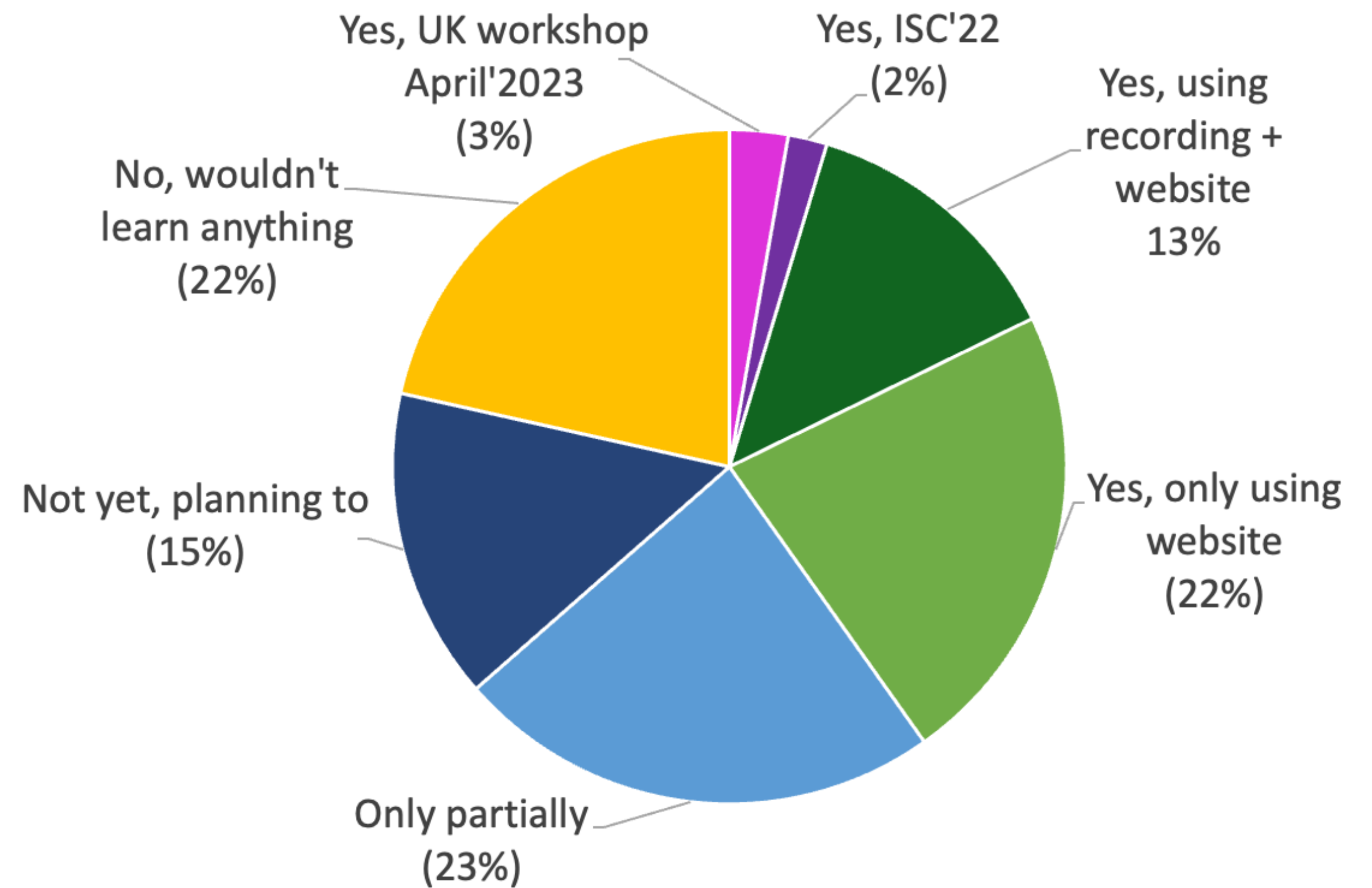# EasyBuild tutorial

https://tutorial.easybuild.io

Are you aware of the EasyBuild tutorial?

No (15%)

Yes (85%)

Have you worked through the EasyBuild tutorial?

Yes, UK workshop April'2023 (3%)

Yes, ISC'22 (2%)

Yes, using recording + website 13%

No, wouldn't learn anything (22%)

Yes, only using website (22%)

Not yet, planning to (15%)

Only partially (23%)

Significant increase in awareness of EasyBuild tutorial in 2023 (~66% → ~85%)

EasyBuild tutorial

https://tutorial.easybuild.io
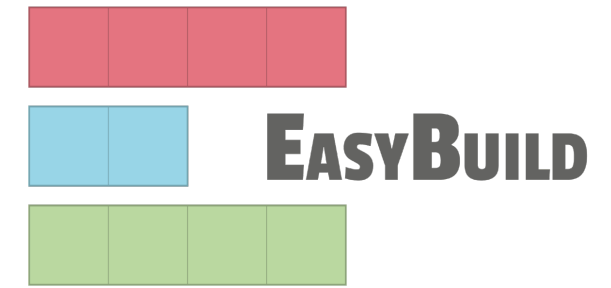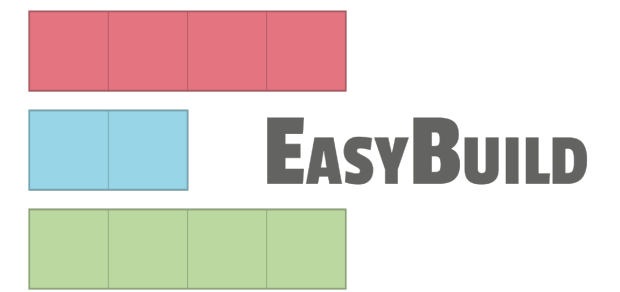
- Official EasyBuild tutorial at ISC'22 (half-day), like we had at ISC'21

- Tutorial submission for ISC'23 was not accepted :(

- For ISC'24 we focused on EESSI, but no success either (trying again for SC24…)

- Most recent tutorial: EasyBuild + EESSI UK workshop: 27+28 April 2023 (right after EUM'23)

    - **Slides + recording available** via https://tutorial.easybuild.io/2023-eb-eessi-uk-workshop

- Maybe it's time for another (extensive, online) EasyBuild tutorial?

    - Like we had in 2020, where we had over 100 registrations

    - Could go beyond the basics, for example covering implementing hooks, use of EasyBuild in EESSI, …

Supported software

Evolution of number of supported software packages in EasyBuild (without extensions)
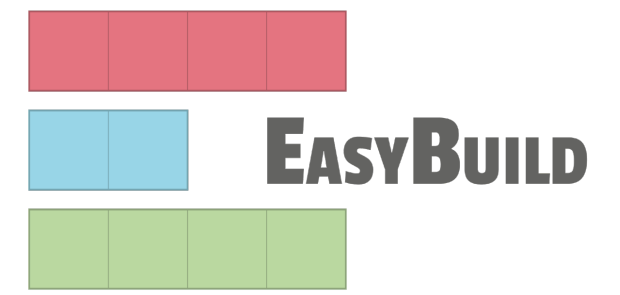

*old easyconfigs were
archived in EasyBuild v4.0*
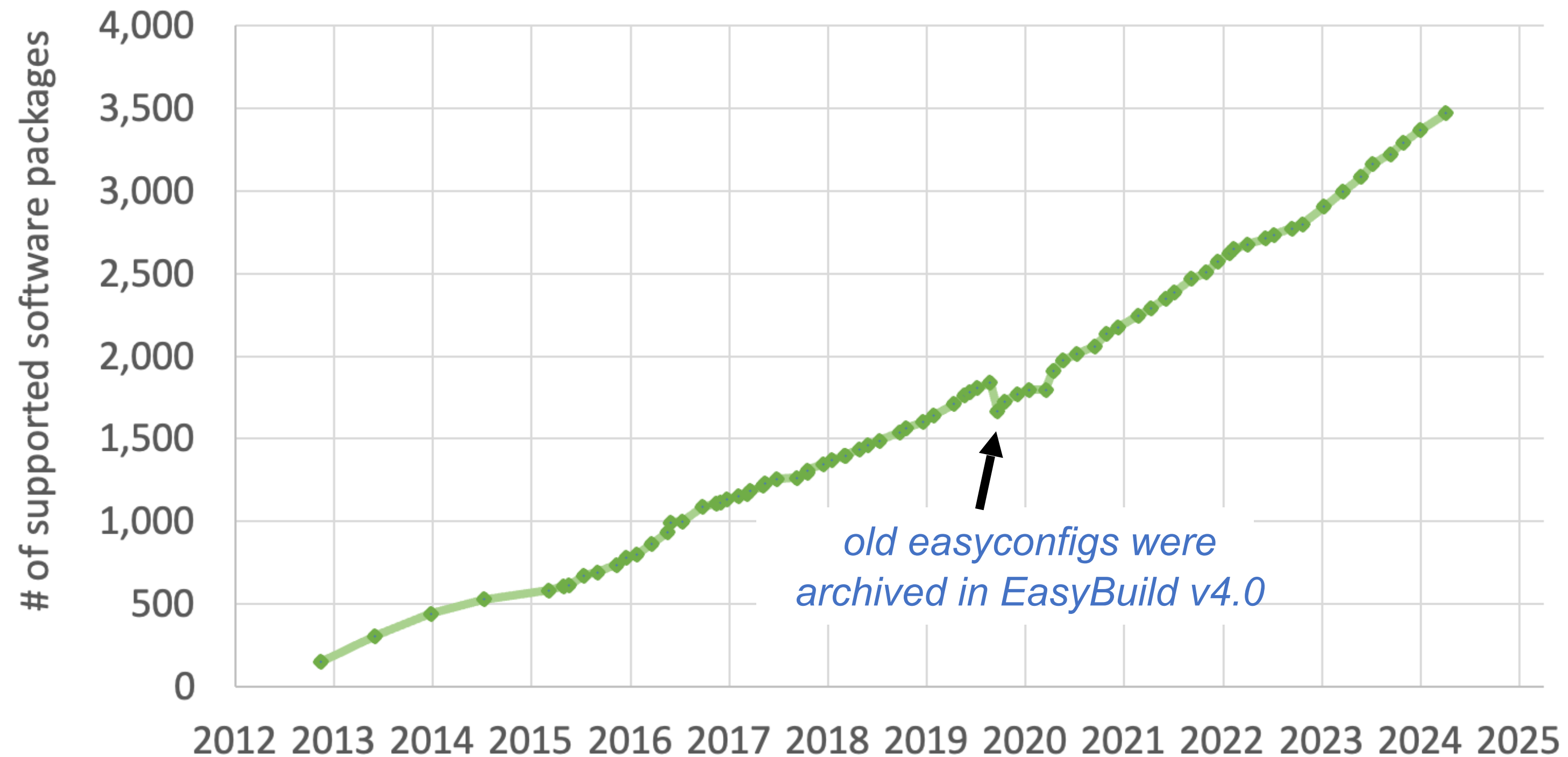
Close to 3,500 different software projects supported by EasyBuild v4.9.1 (plus ~2,000 extensions)

Stable upward trend, no sign of slowing down…

Another dip coming due to archiving of unsupported toolchains in EasyBuild 5.0

EasyBuild mailing list
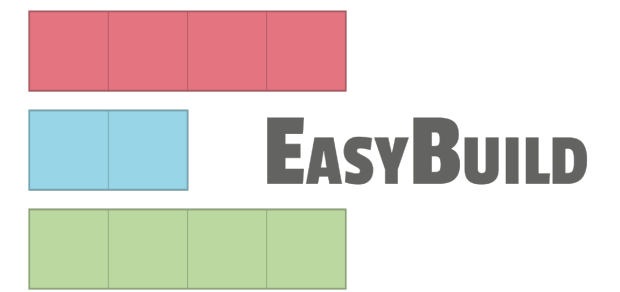

Are you subscribed to the EasyBuild mailing list?

- Yes, but not reading
- Yes, but only reading
- Yes, occasionally participate
- Yes, frequently participate
- No, but I should join
- No, not interested

- No, not interested (15%)
- Yes, but not reading (9%)
- Yes, but only reading (33%)
- Yes, occasionally participate (15%)
- Yes, frequently participate (1%)
- No, but I should join (27%)
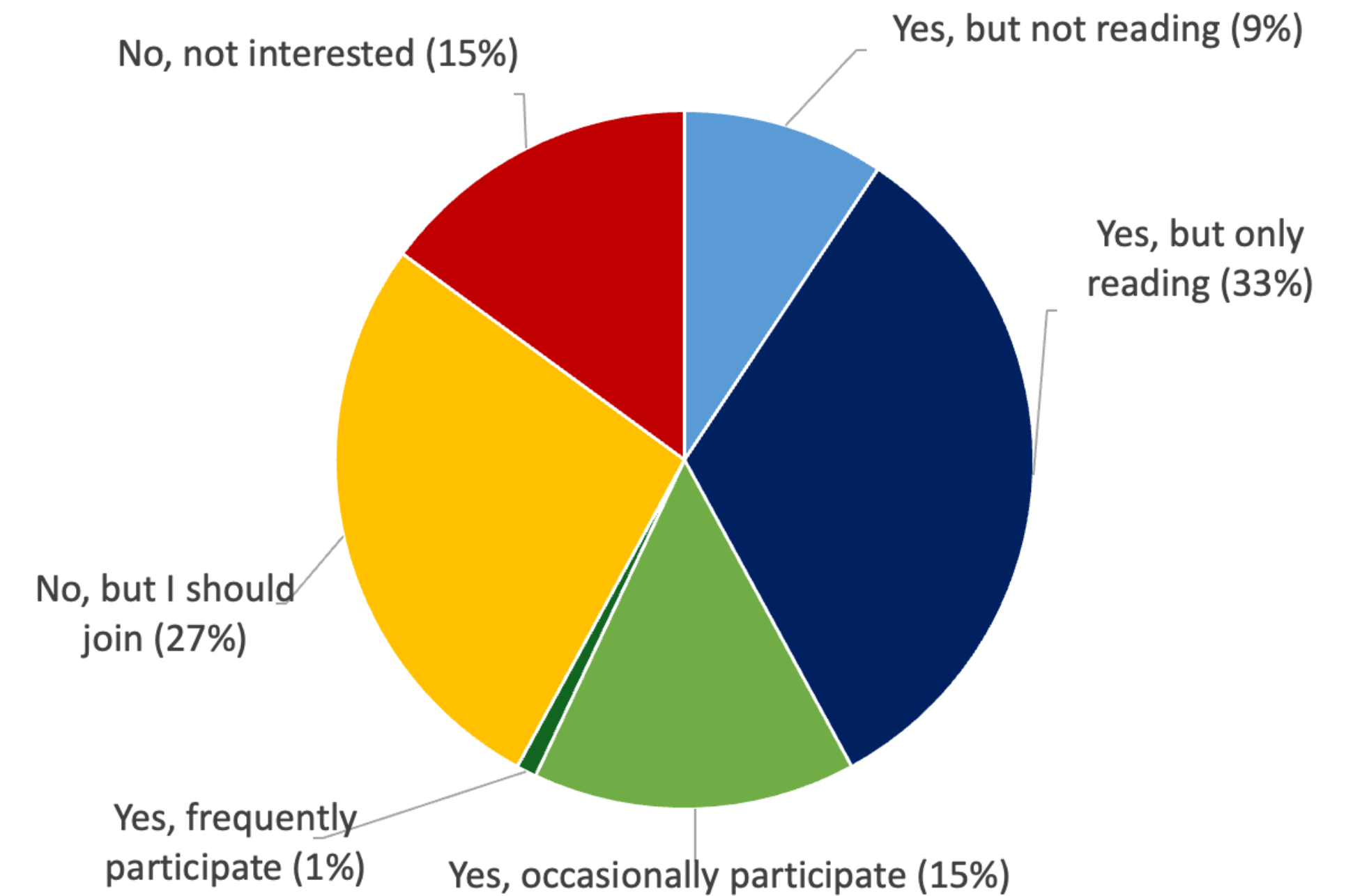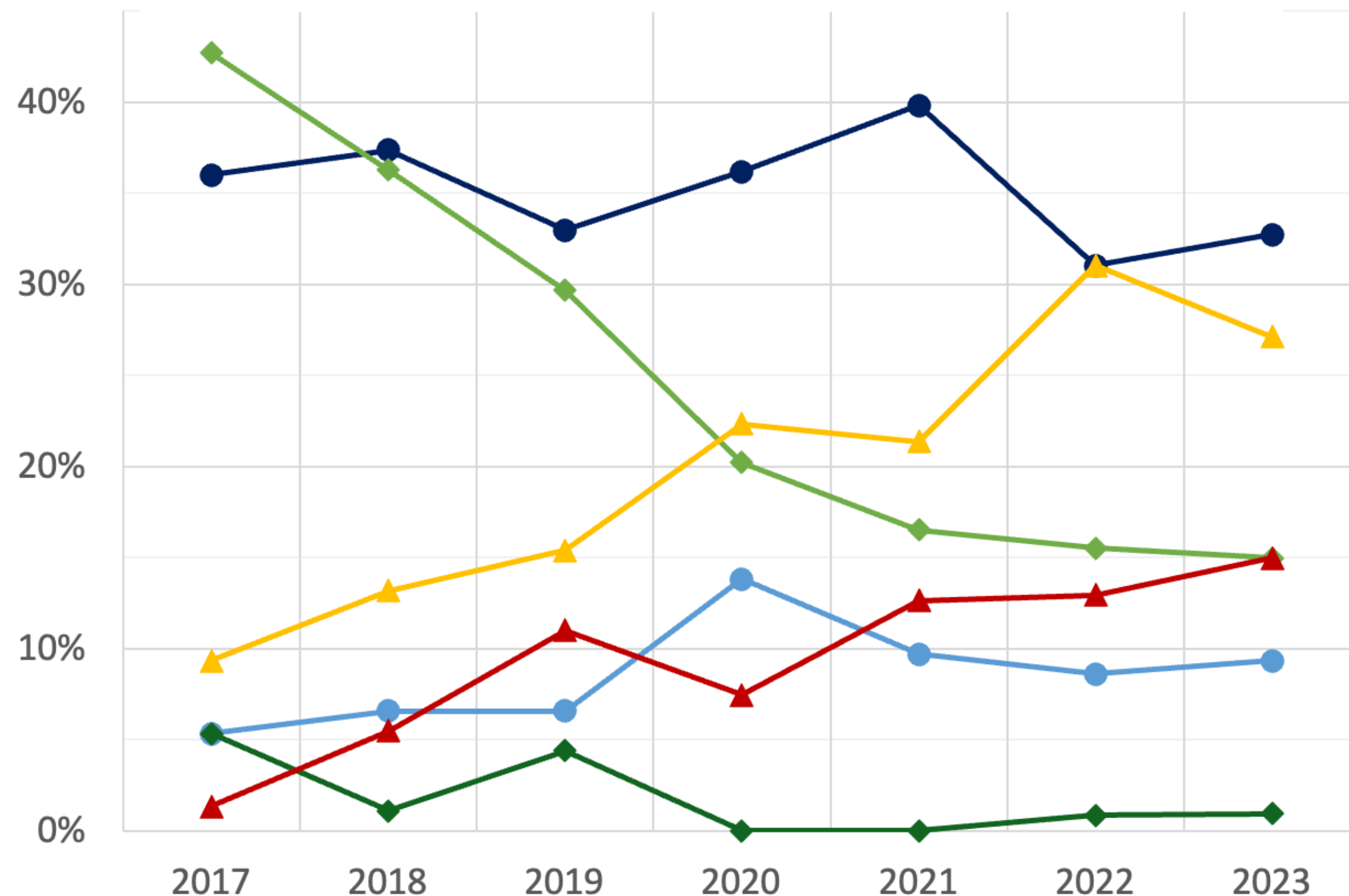
- Active participation remains quite limited, just ~15%

- Still significant interest in joining the mailing list (~27)

- Mailing list remains an important communication channel!

Subscribe to EasyBuild mailing list via
https://lists.ugent.be/wws/subscribe/easybuild
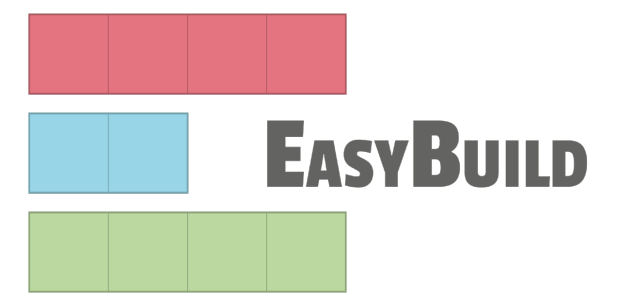
EasyBuild mailing list (traffic)

Traffic on EasyBuild mailing list peaked in 2017,

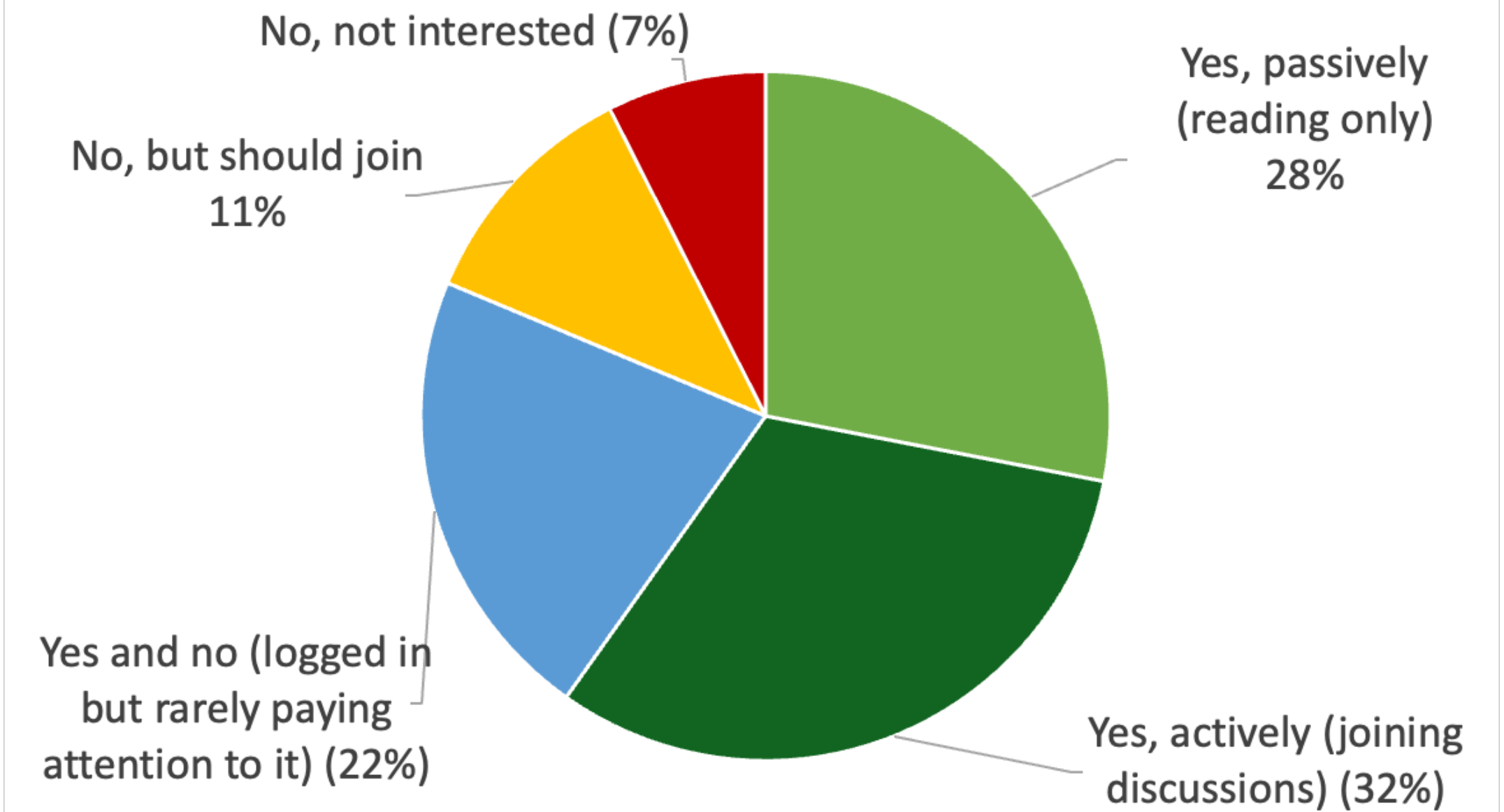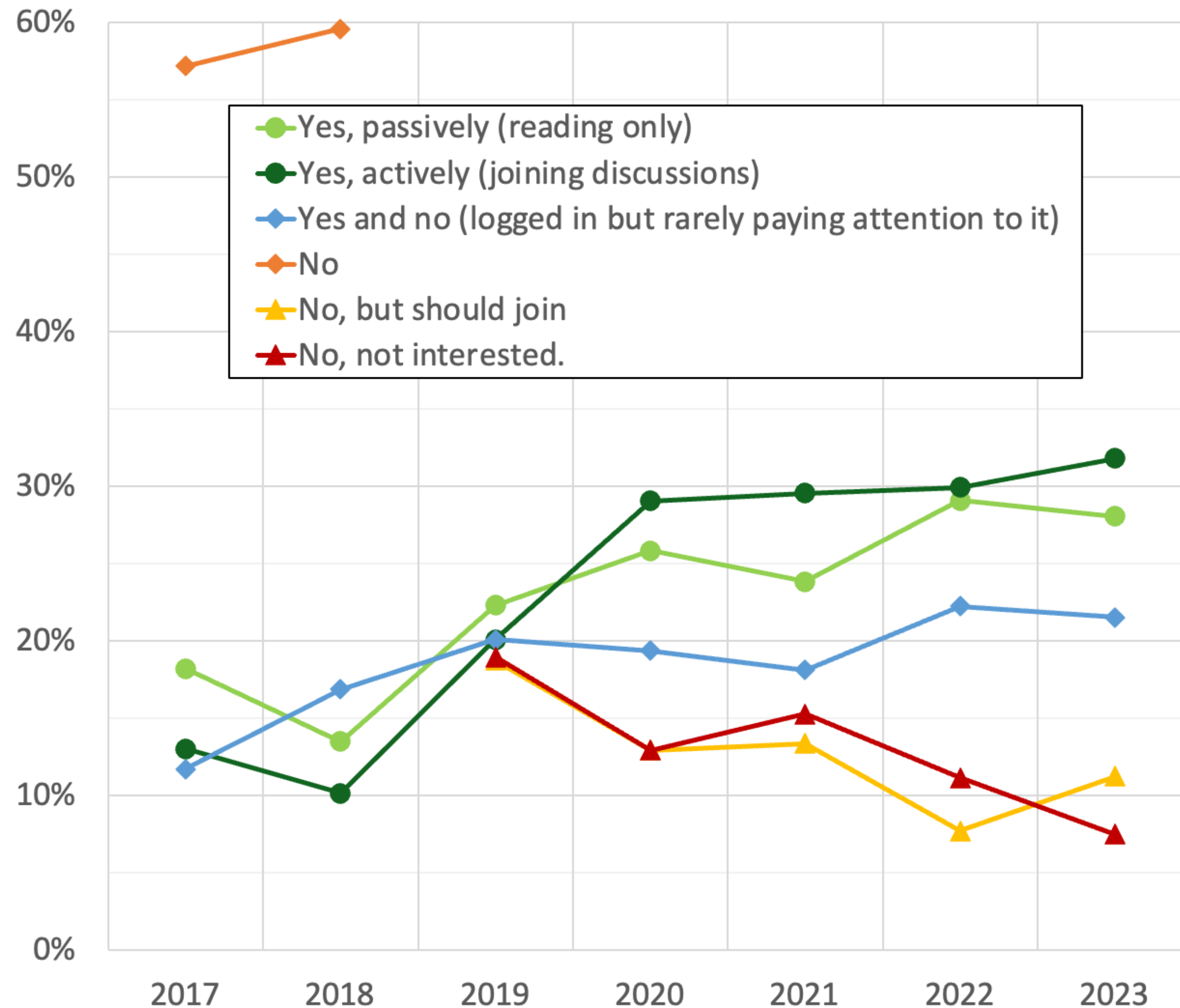due to start of EasyBuild Slack channel, but remains active.

slack effect
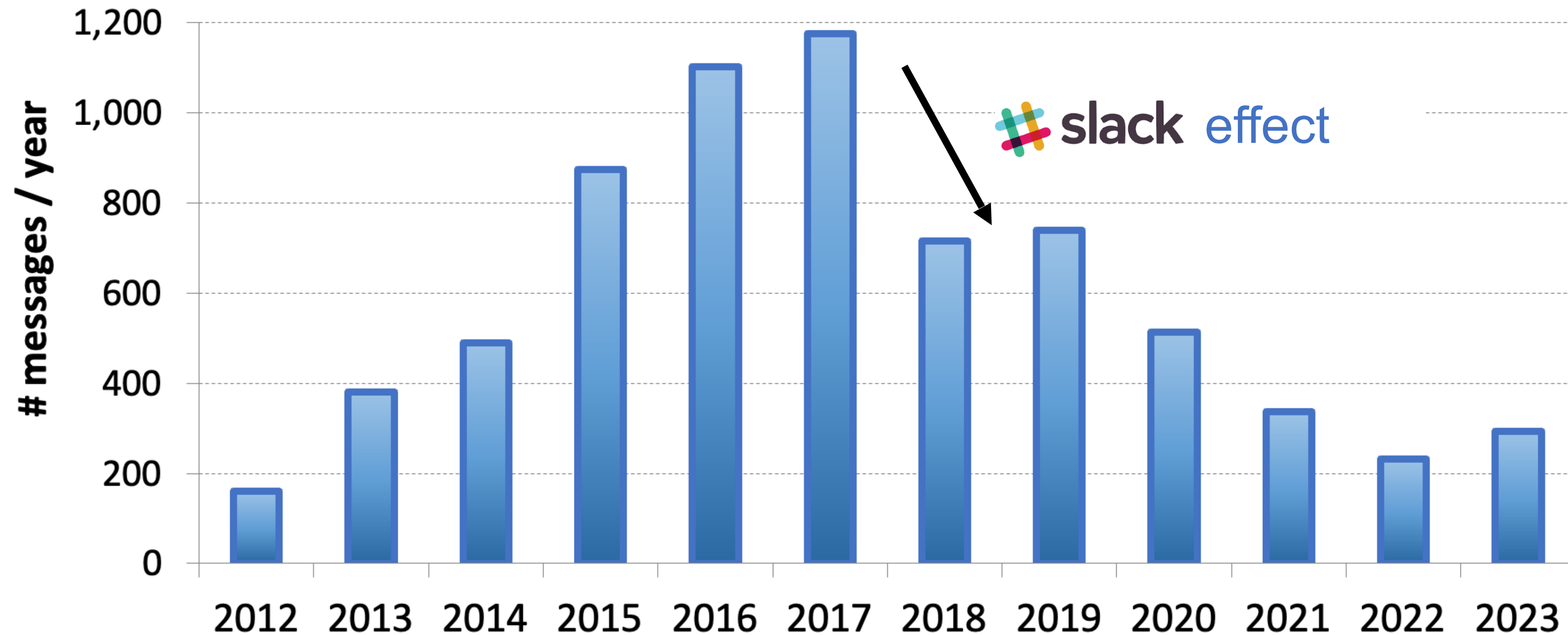
325 subscribers to EasyBuild mailing list currently, small uptick in activity in 2023…

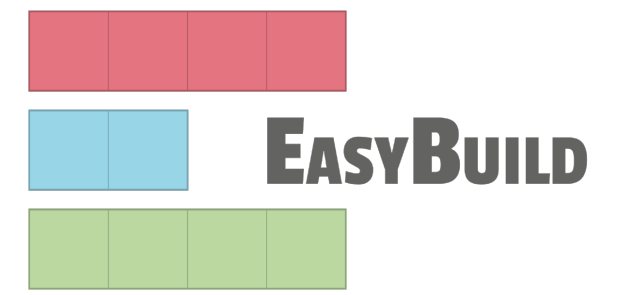# EasyBuild Slack

Do you use the EasyBuild Slack channel?

Legend:
- Yes, passively (reading only)
- Yes, actively (joining discussions)
- Yes and no (logged in but rarely paying attention to it)
- No
- No, but should join
- No, not interested.

Pie chart:
- No, not interested (7%)
- No, but should join 11%
- Yes, passively (reading only) 28%
- Yes and no (logged in but rarely paying attention to it) (22%)
- Yes, actively (joining discussions) (32%)

Only ~32% of community *actively* participates in Slack.

Significant part of community (~19%) is not on Slack.

Join EasyBuild Slack via
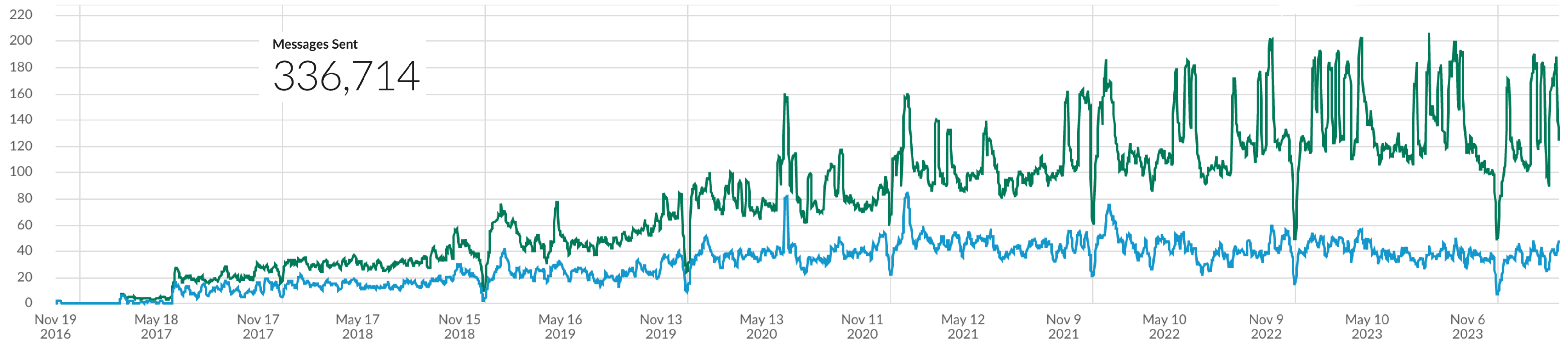*https://easybuild.io/join-slack*

31

EasyBuild Slack

- Steady growth in activity since start of EasyBuild Slack in 2017

- 875 members on Slack channel, steady increase of 150-160 per year since 2020

- Total messages sent increases quickly (since 2019: 105k ☞ 192k ☞ 260k ☞ 311k ☞ 336k)
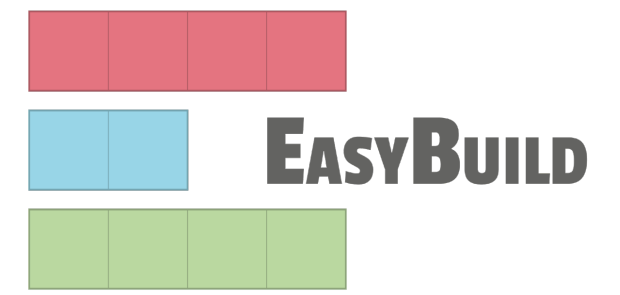
Join EasyBuild Slack via *https://easybuild.io/join-slack* !
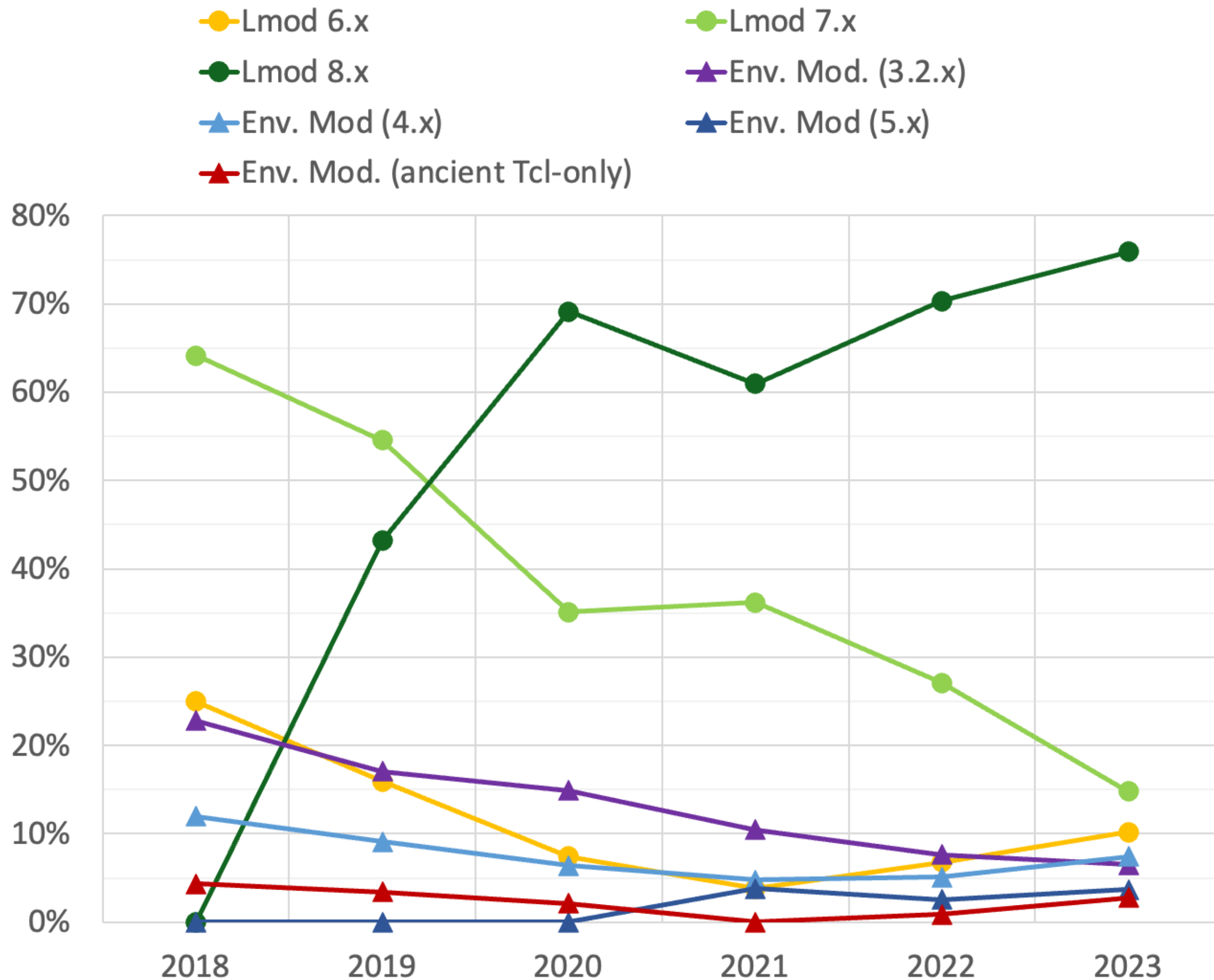
Weekly Daily

Messages Sent
336,714

● Weekly active members ● Members who posted

32

Environment modules

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

Legend:
- Lmod 6.x
- Lmod 7.x
- Lmod 8.x
- Env. Mod. (3.2.x)
- Env. Mod (4.x)
- Env. Mod (5.x)
- Env. Mod. (ancient Tcl-only)

- Lmod remains most popular modules tool (by far)

- ~76% is using Lmod 8.x, ~15% is still using Lmod 7.x

- ~10% still using Lmod 6.x (deprecated in EasyBuild!)
  - Partially due to outdated Lmod package in Debian/Ubuntu
  - New maintainer for Lmod package in Debian/Ubuntu: Alexandre Strube (JSC)

- Lmod 6 + 7 will no longer be supported by EasyBuild 5.0 !

- Tcl-based Environment Modules is small minority (~7%)

Other tools & projects

Do you use any other tools/projects in combination with EasyBuild?
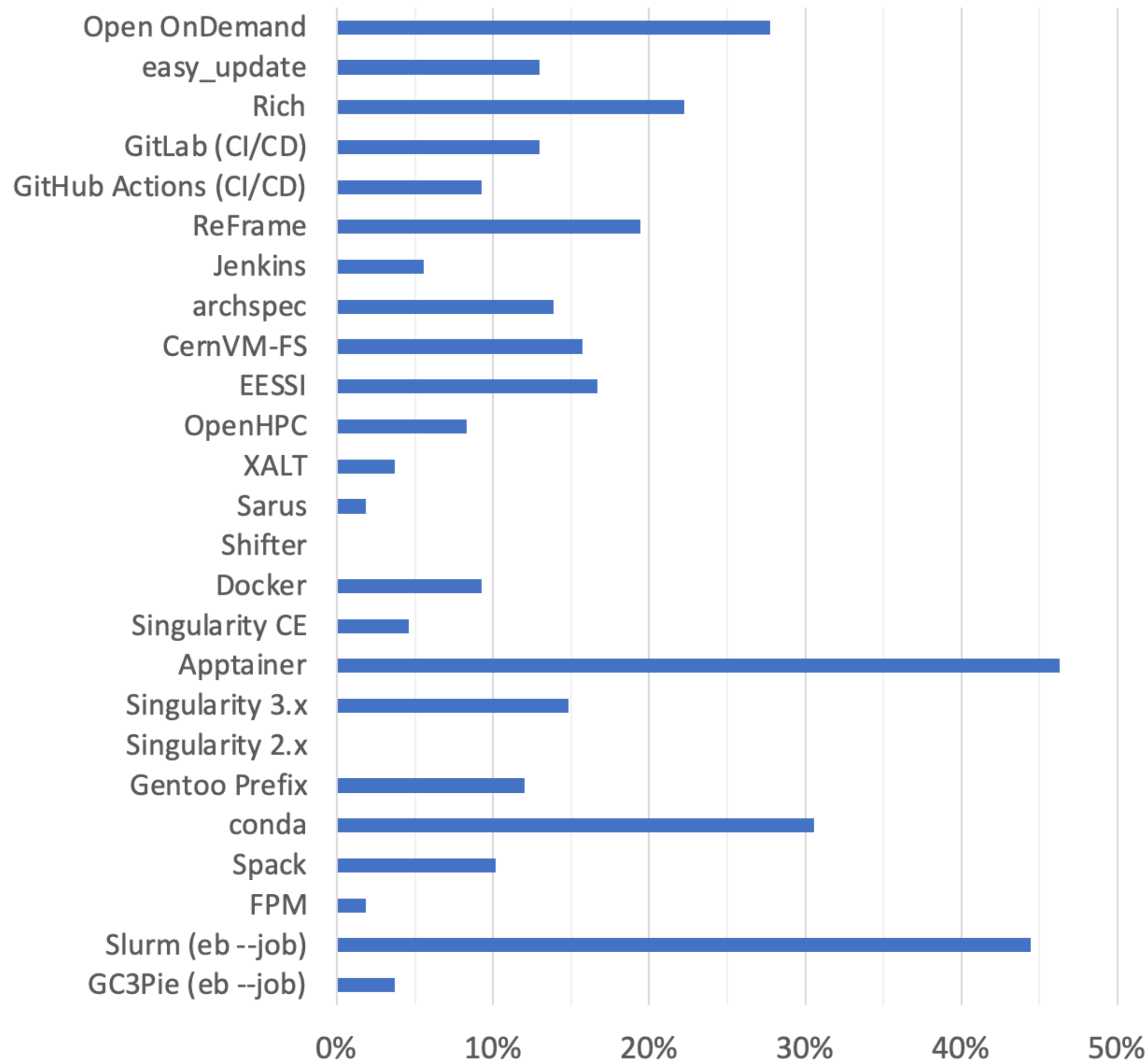
- Slurm way more popular than GC3Pie as job backend, will become default job backend in EasyBuild 5.0

- Apptainer is most postill pular 3rd party tool used with EasyBuild

- Conda is used by over 30% (combined with EasyBuild?!)

- CernVM-FS is steady at (only) 16%

- Small stable minority (~10%) uses both EasyBuild and Spack?

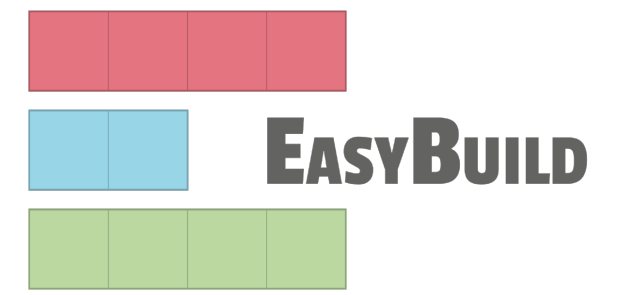- Increase for Open OnDemand (18% ☞ 28%), Rich (17% ☞ 22%), ReFrame (11% ☞ 19%), EESSI (10% ☞ 17%)

European Environment for Scientific Software Installations

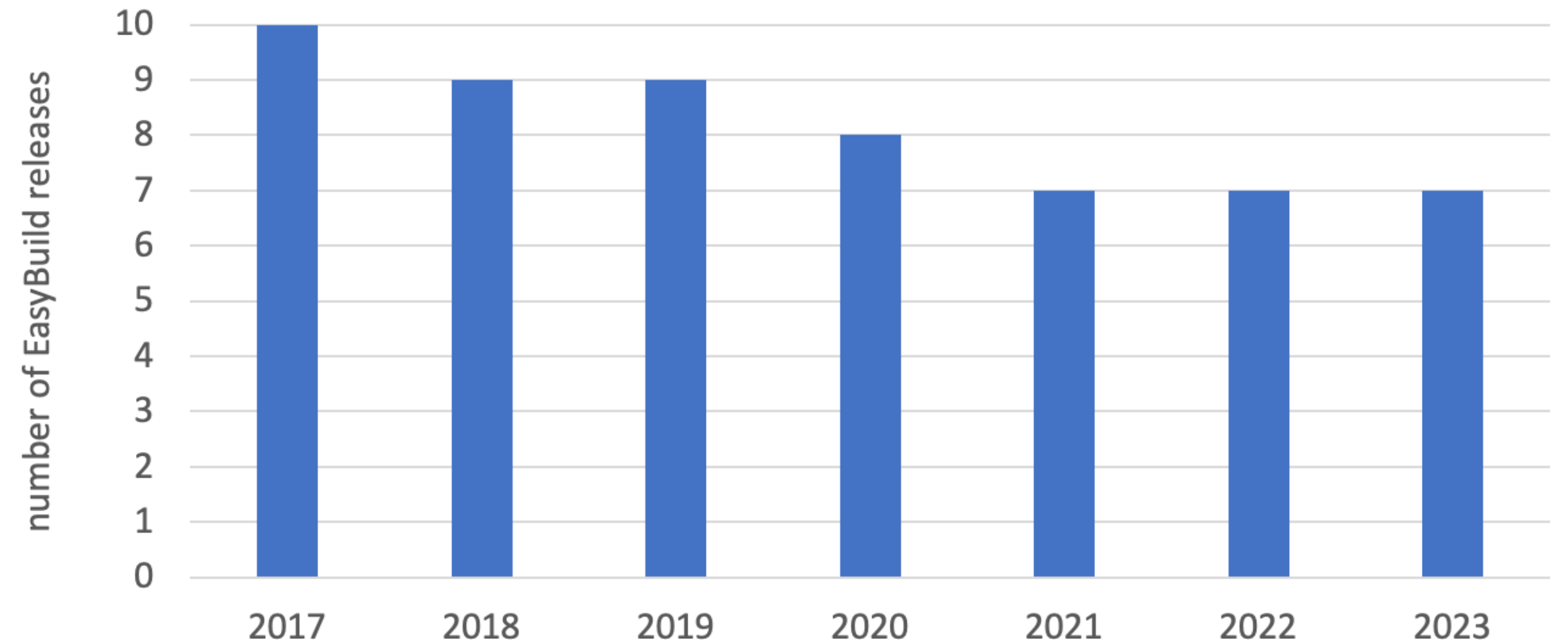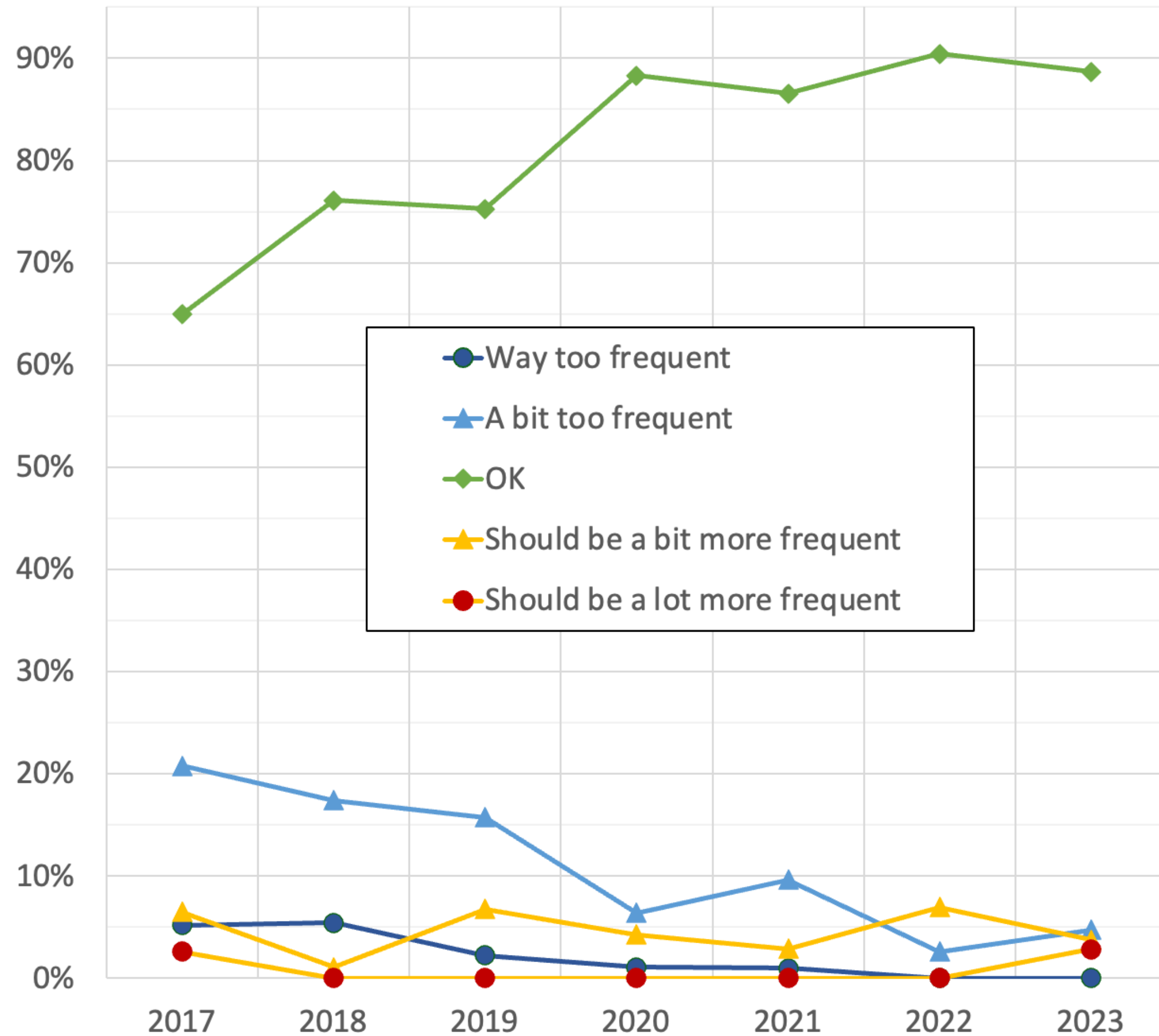

Have you heard about the EESSI project?
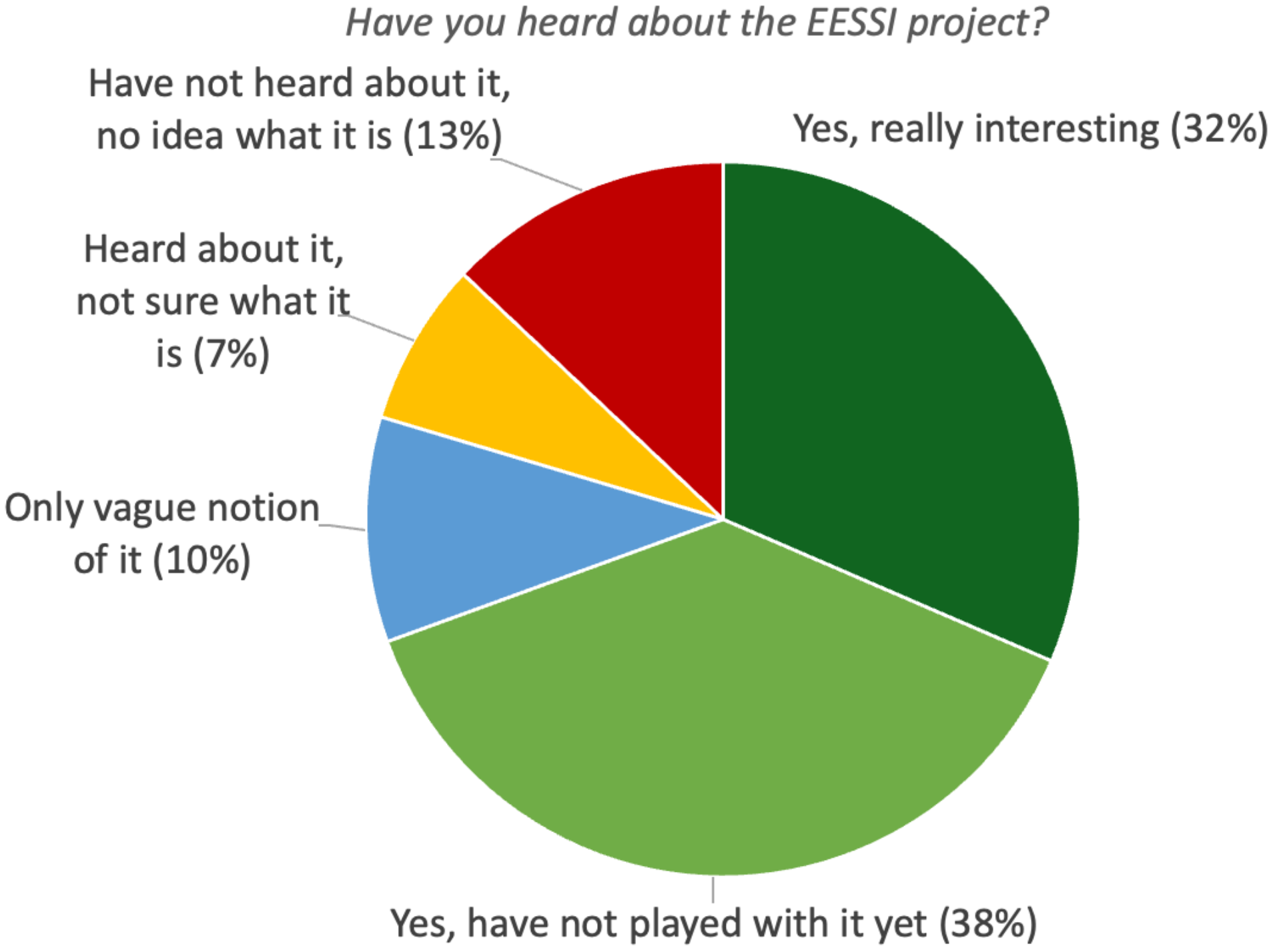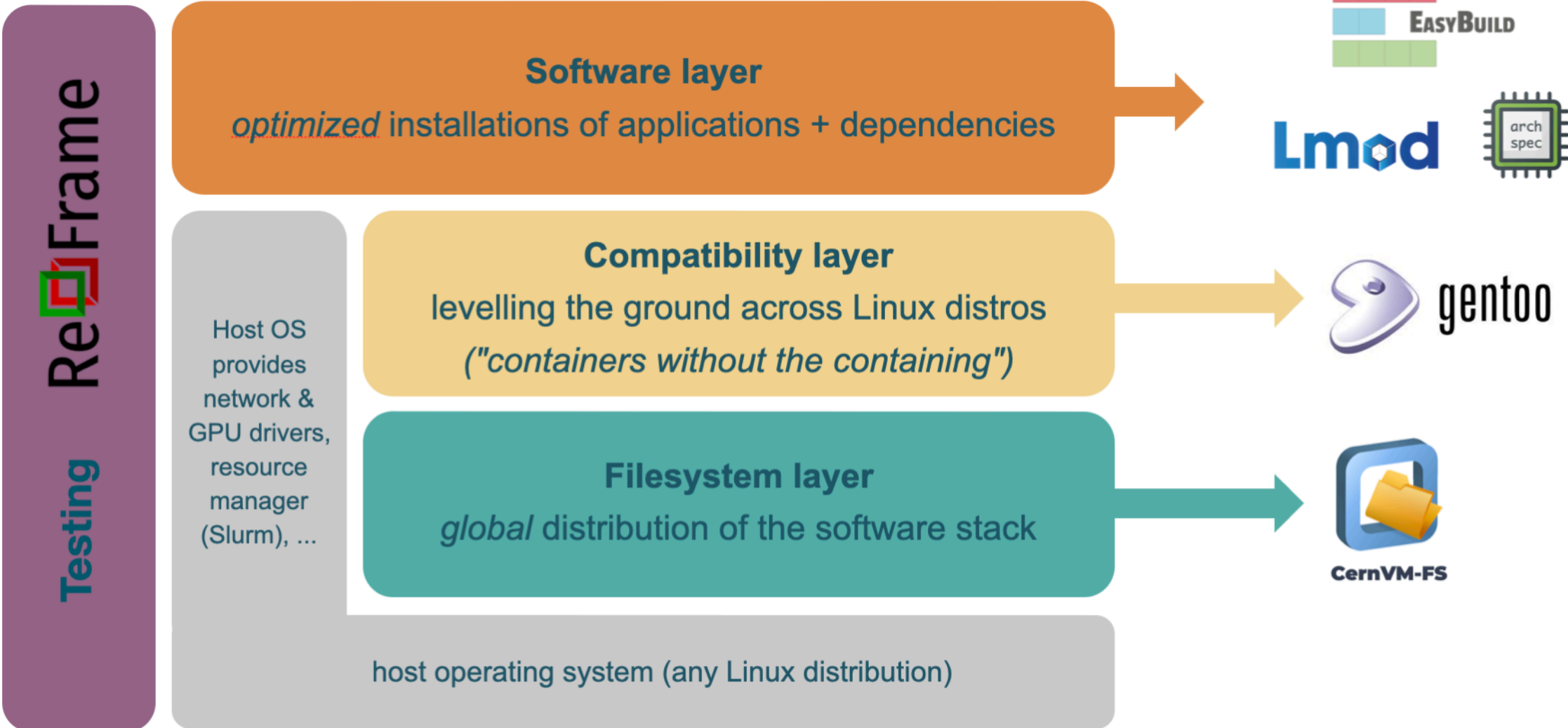
- Slight increase in EasyBuild community of people who have hands-on experience with EESSI (~23% ☞ 31%)

- It is now a funded project, via the MultiXscale EuroHPC Centre-of-Excellence (2023-2026)

- Significant involvement of EasyBuild community, but does not replace EasyBuild!

- **More on EESSI on 3rd day of EUM'24**

EasyBuild releases

How do you like the frequency of EasyBuild releases?

How do you like the frequency of EasyBuild releases?

Legend:
- Way too frequent
- A bit too frequent
- OK
- Should be a bit more frequent
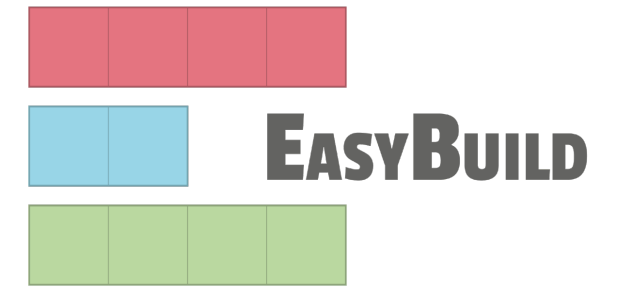- Should be a lot more frequent

Number of EasyBuild releases has been steady in recent years

Currently: new EasyBuild release every 7-8 weeks
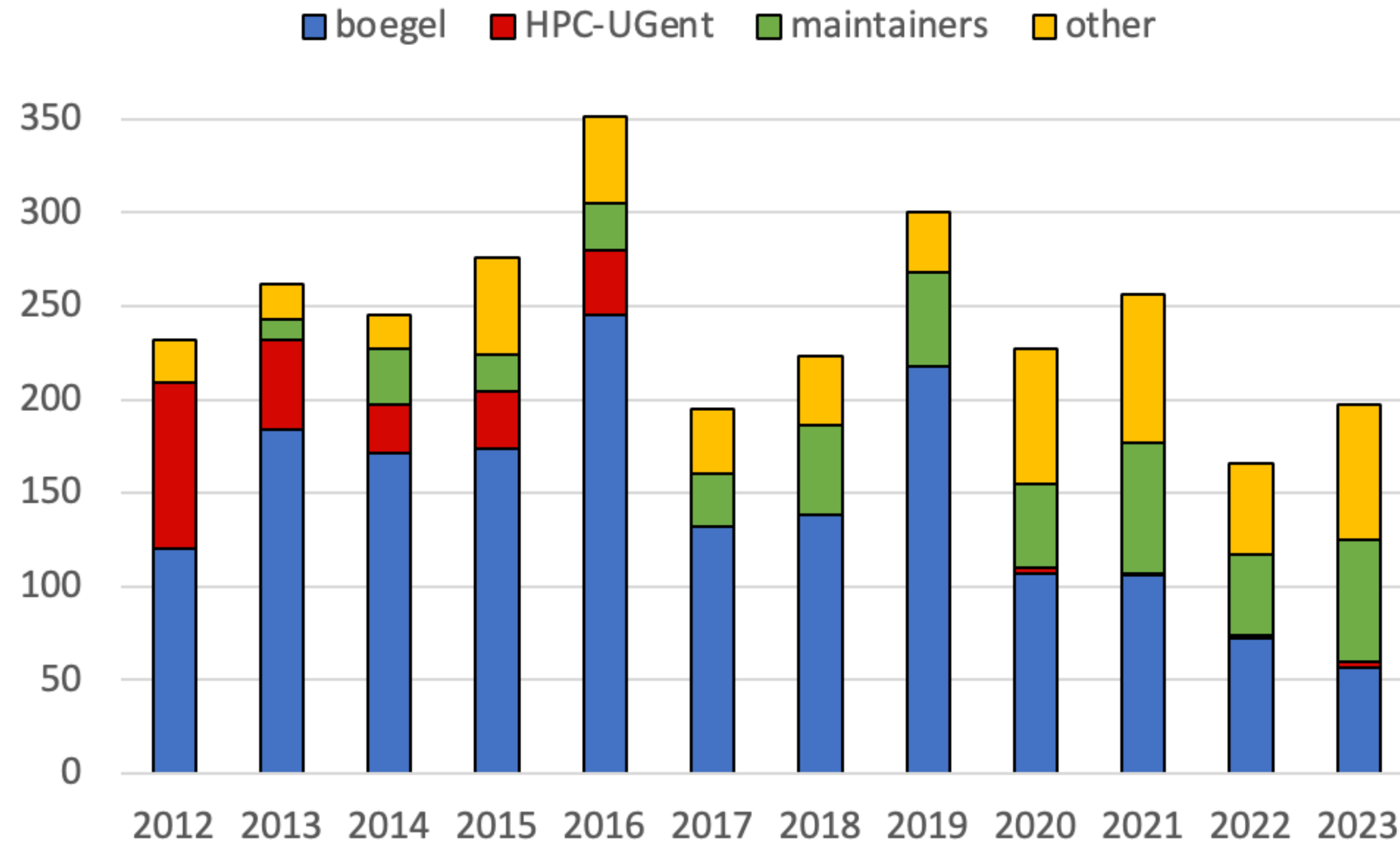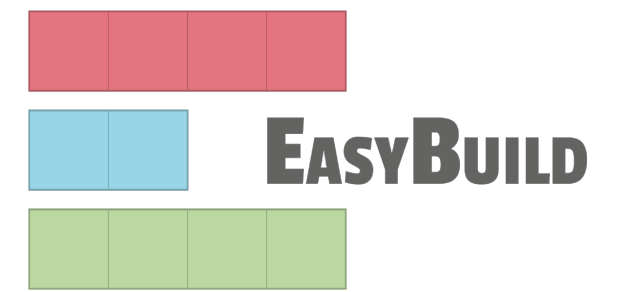
Current frequency is well appreciated 👍

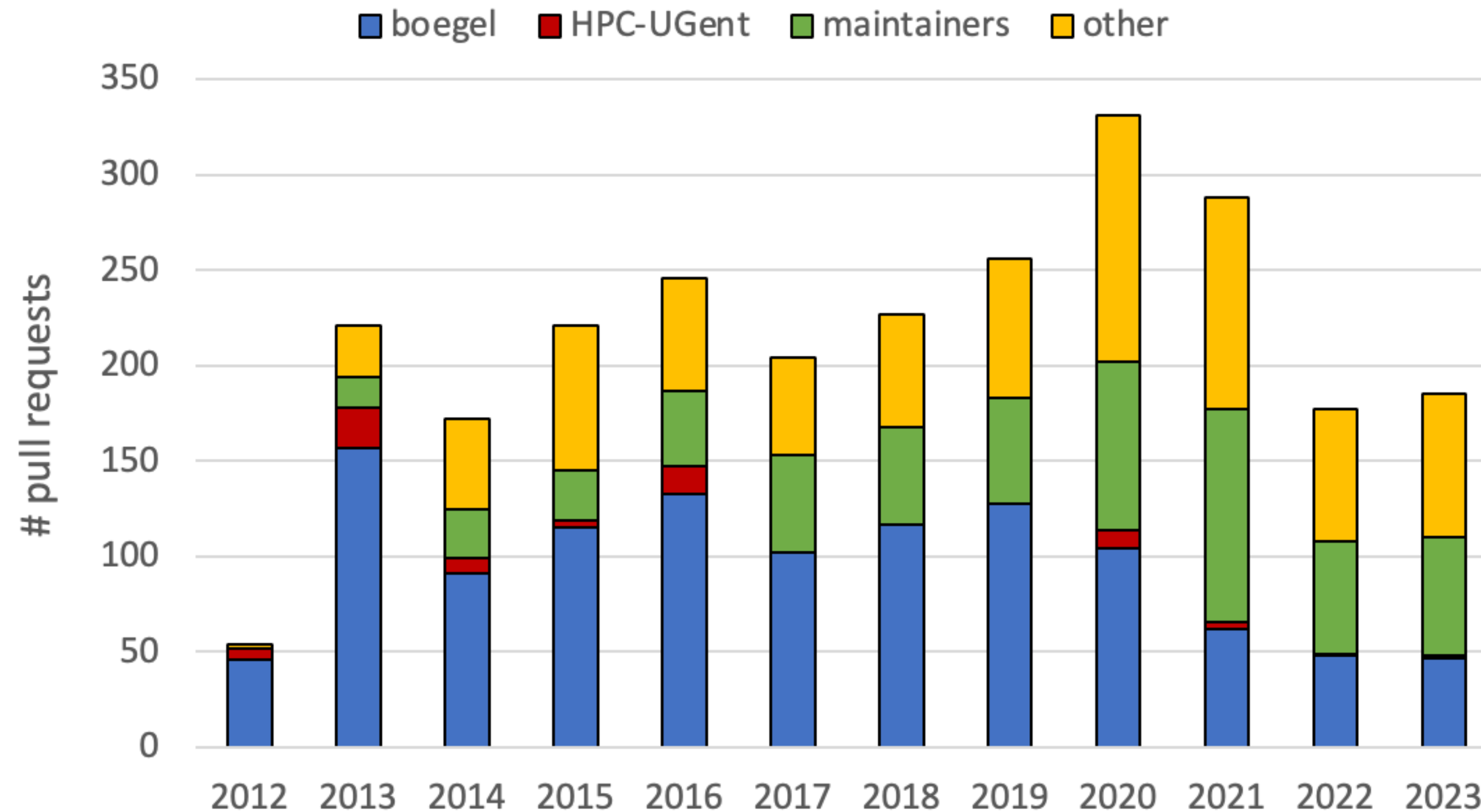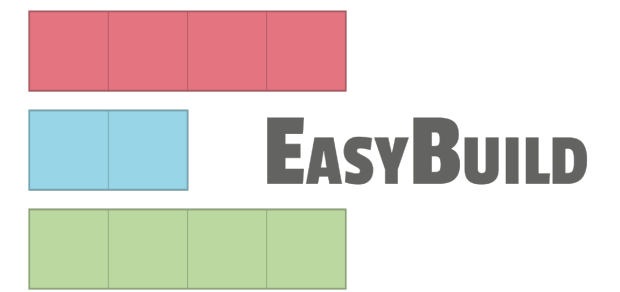# Contributors, contributors, contributors, …

Overview of contributions (framework)

- ~71% of framework PRs in 2023 by contributors outside HPC-UGent (was ~55% in 2020-2022)

- Growing share of framework PRs by non-maintainers (~33% in 2023)

Overview of contributions (easyblocks)

- Reduced activity in easyblocks in 2022-2023 (shift to using generic easyblocks?)

- ~75% of PRs by contributors outside HPC-UGent, ~40% by people by non-maintainers
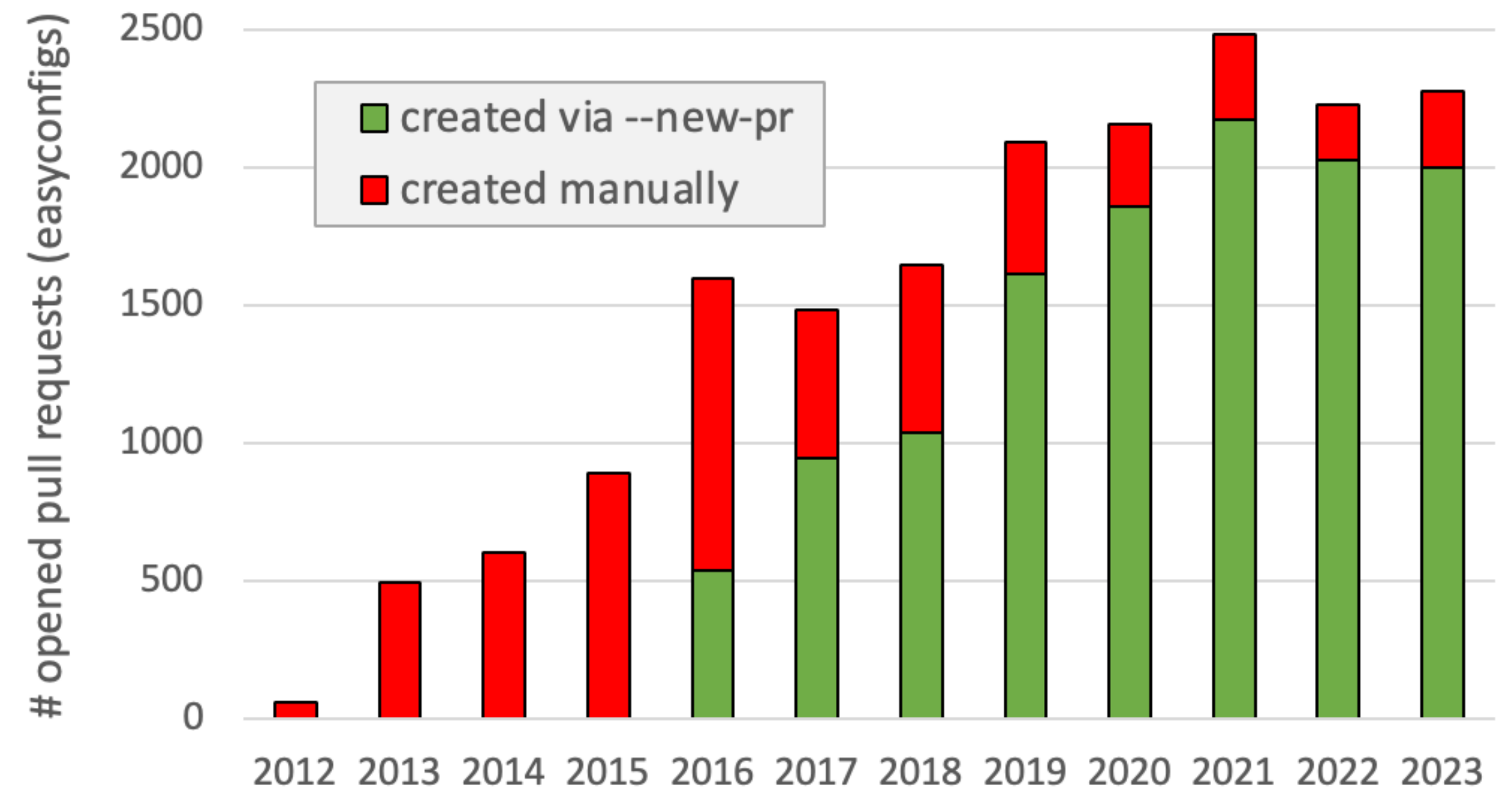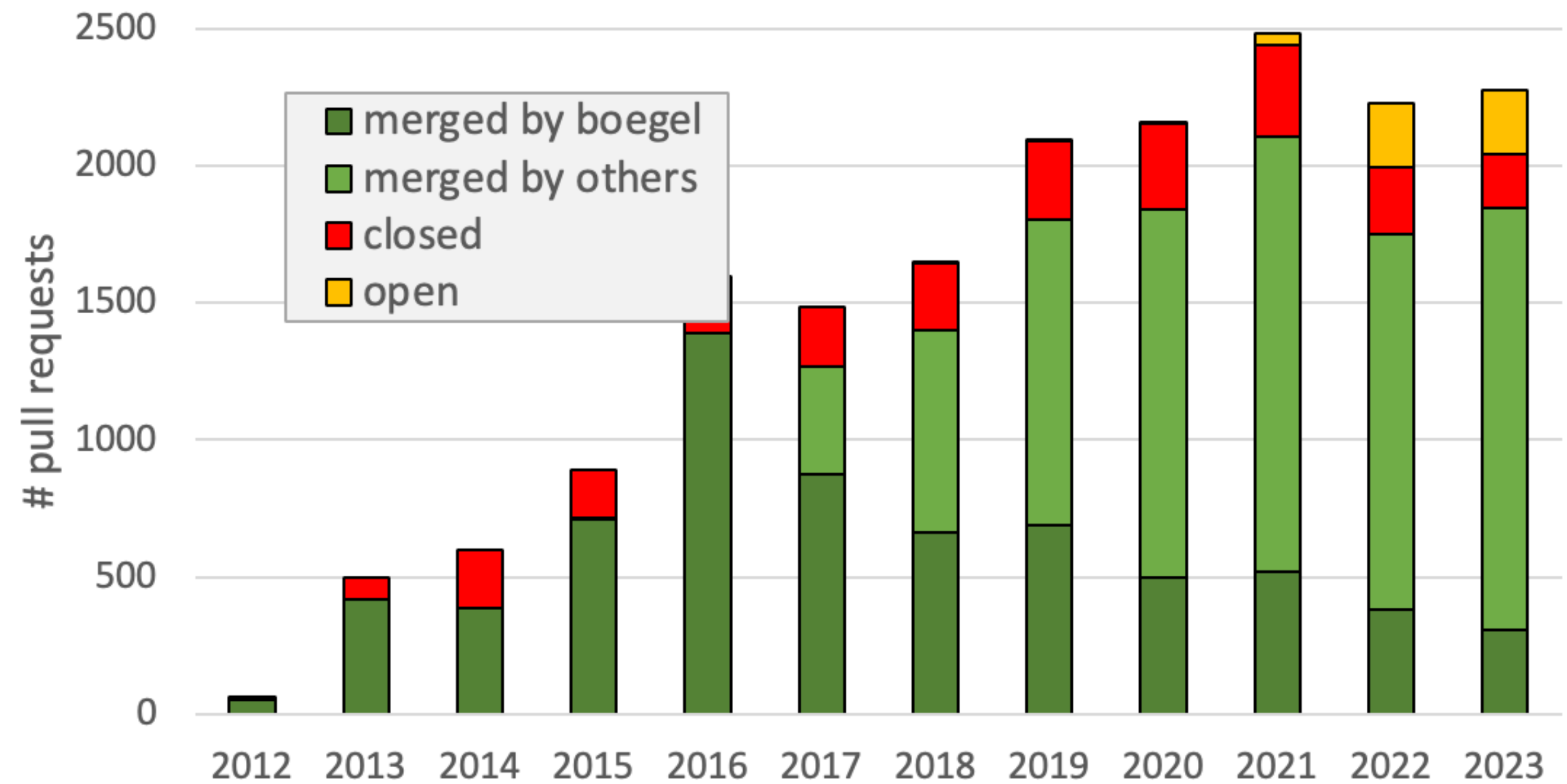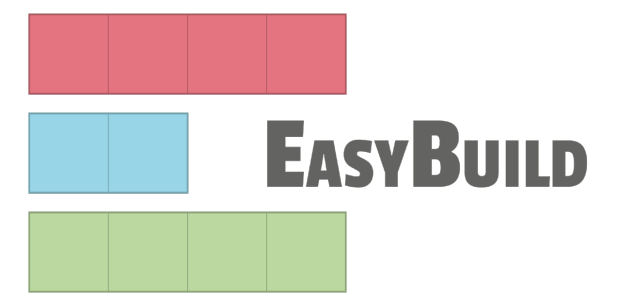
Overview of contributions (easyconfigs)
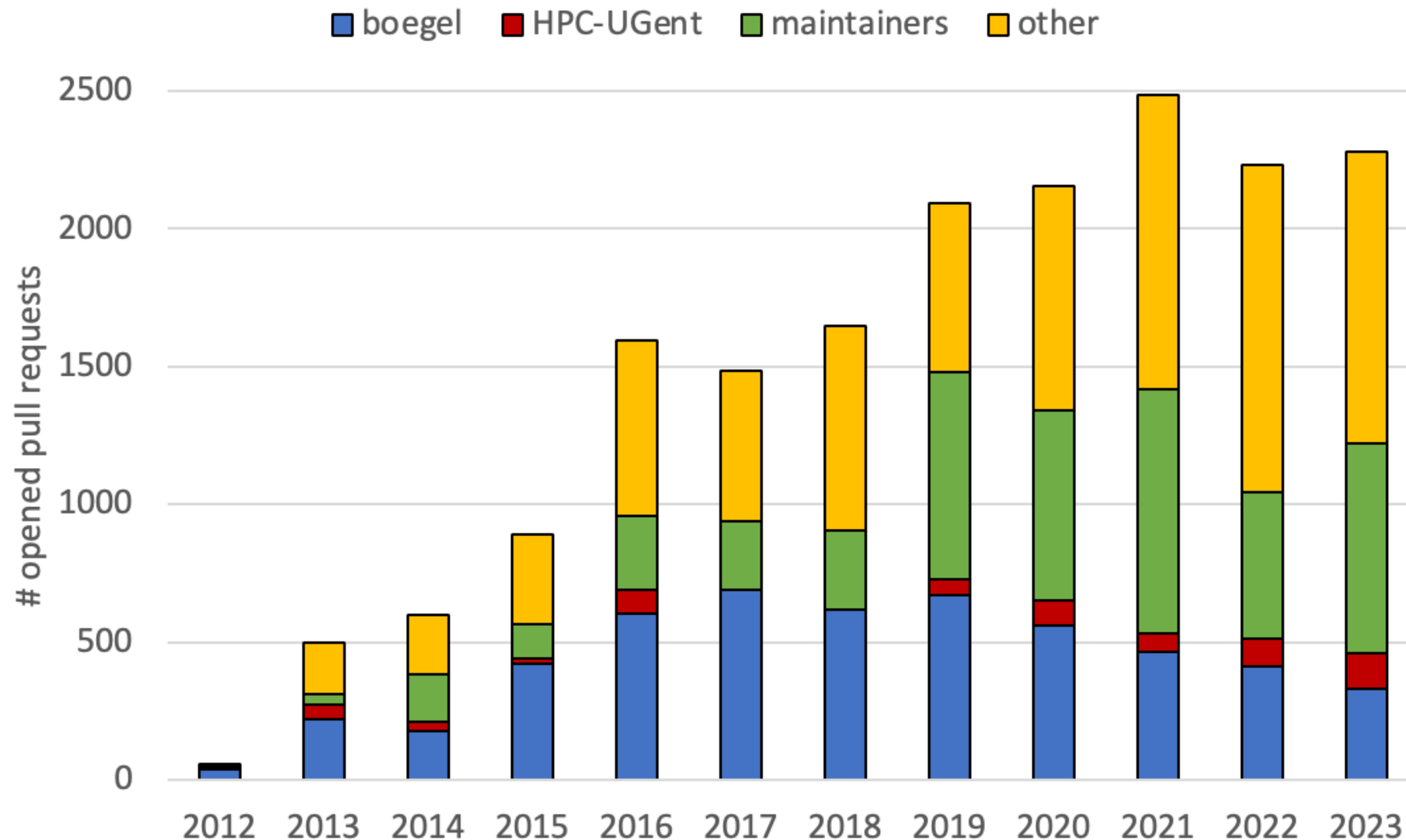
- Reasonably stable number of easyconfig PRs per year since 2019 (2.1k-2.5k)

- At the limit of what current EasyBuild maintainers can process

- Over 85% of PRs from contributors outside of HPC-UGent (still increasing year by year)
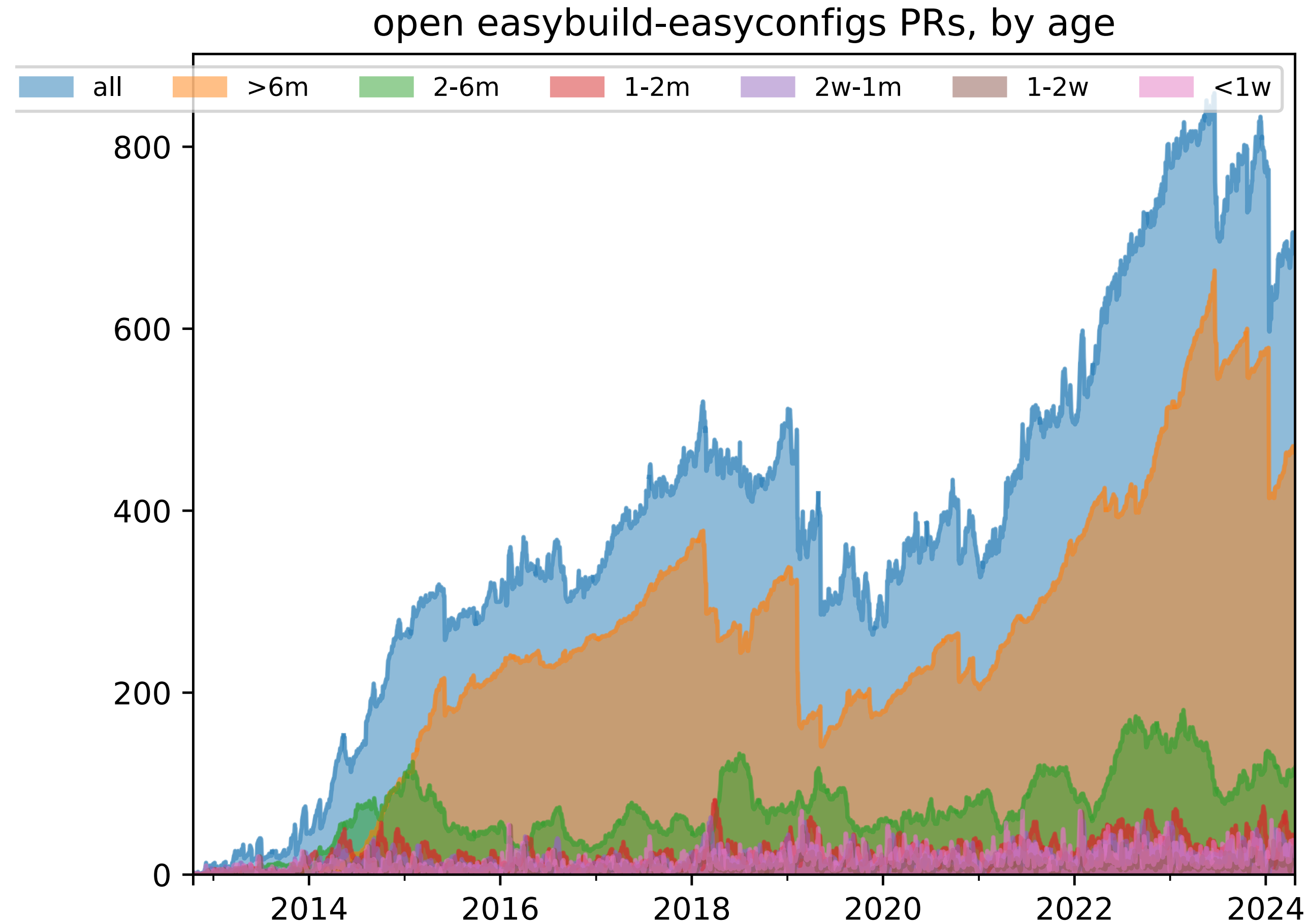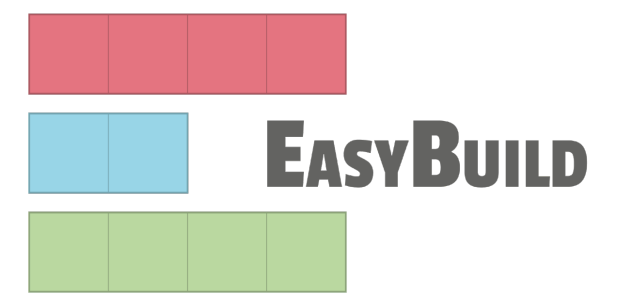
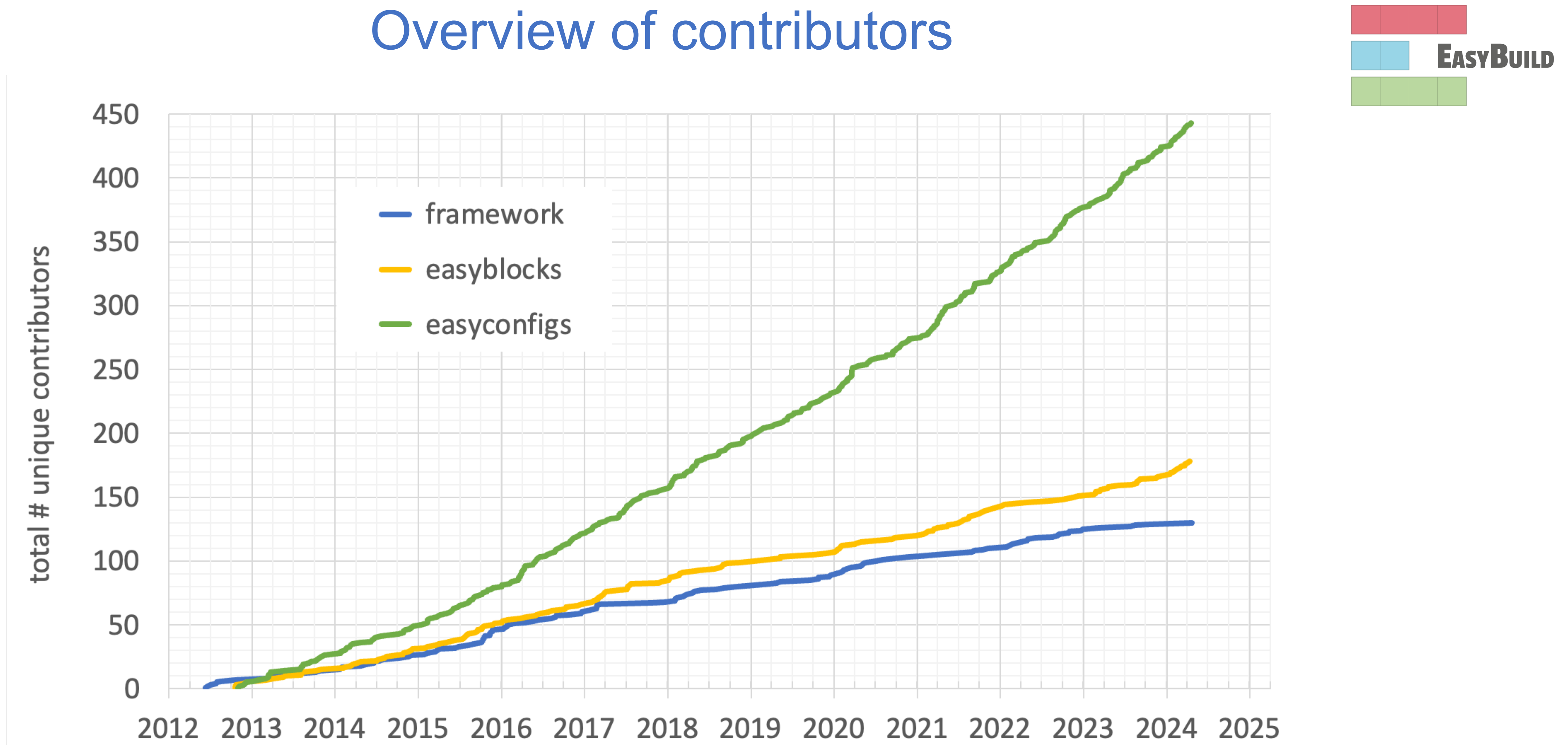# Overview of contributions (easyconfigs)

- Growing number of easyconfig PRs merged by other maintainers (>85% in 2023)

- Significant number of easyconfigs PRs from 2022-2023 are still open…

- Ratio of PRs opened with "`eb --new-pr`" is stable at ~88%

Overview of contributions (open easyconfigs PRs)

open easybuild-easyconfigs PRs, by age

- Backlog of open easyconfigs PRs has been pushed down a bit, thanks to toolchain support policy

- Lack of sufficient time from active EasyBuild maintainers to process contributions *and* tackle backlog
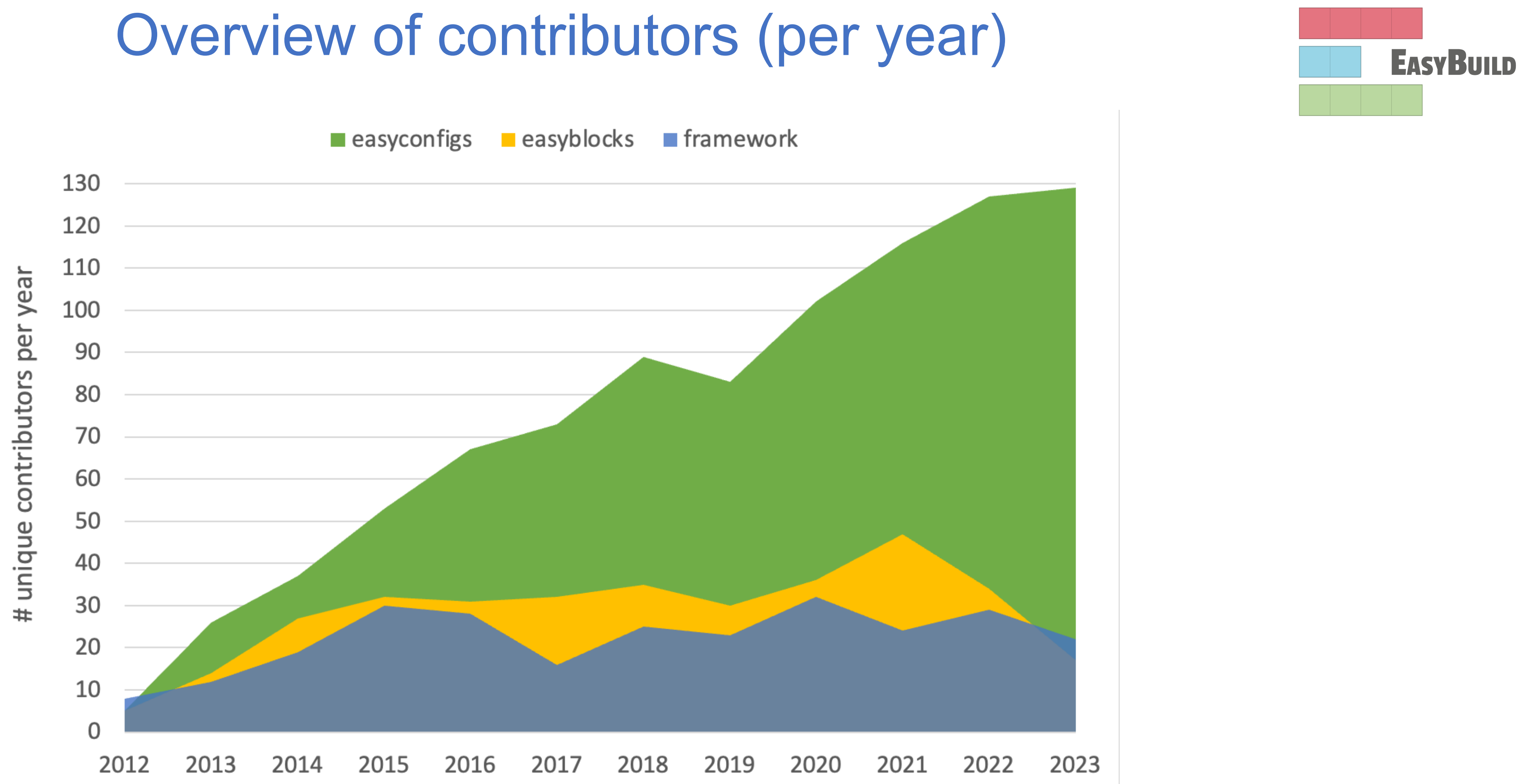
Overview of contributors

- Close to 50 new contributors to easyconfigs repo in 2023, now almost 450 in total

- Continued growth in unique contributors for all 3 major EasyBuild GitHub repos
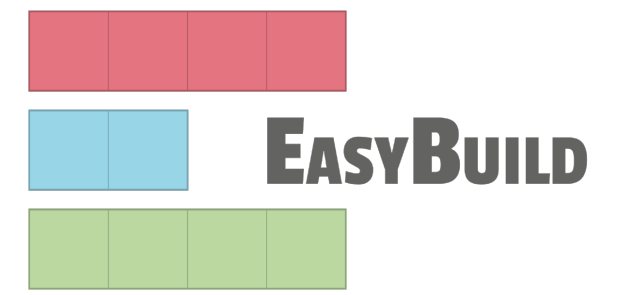
Overview of contributors (per year)

- 20-30 unique contributors per year to EasyBuild framework/easyblocks

- Only 17 unique contributors to easyblocks in 2023 (but in 2024Q1 already > 25)

- Close to 130 unique contributors **per year** for easyconfigs (keeps growing…)
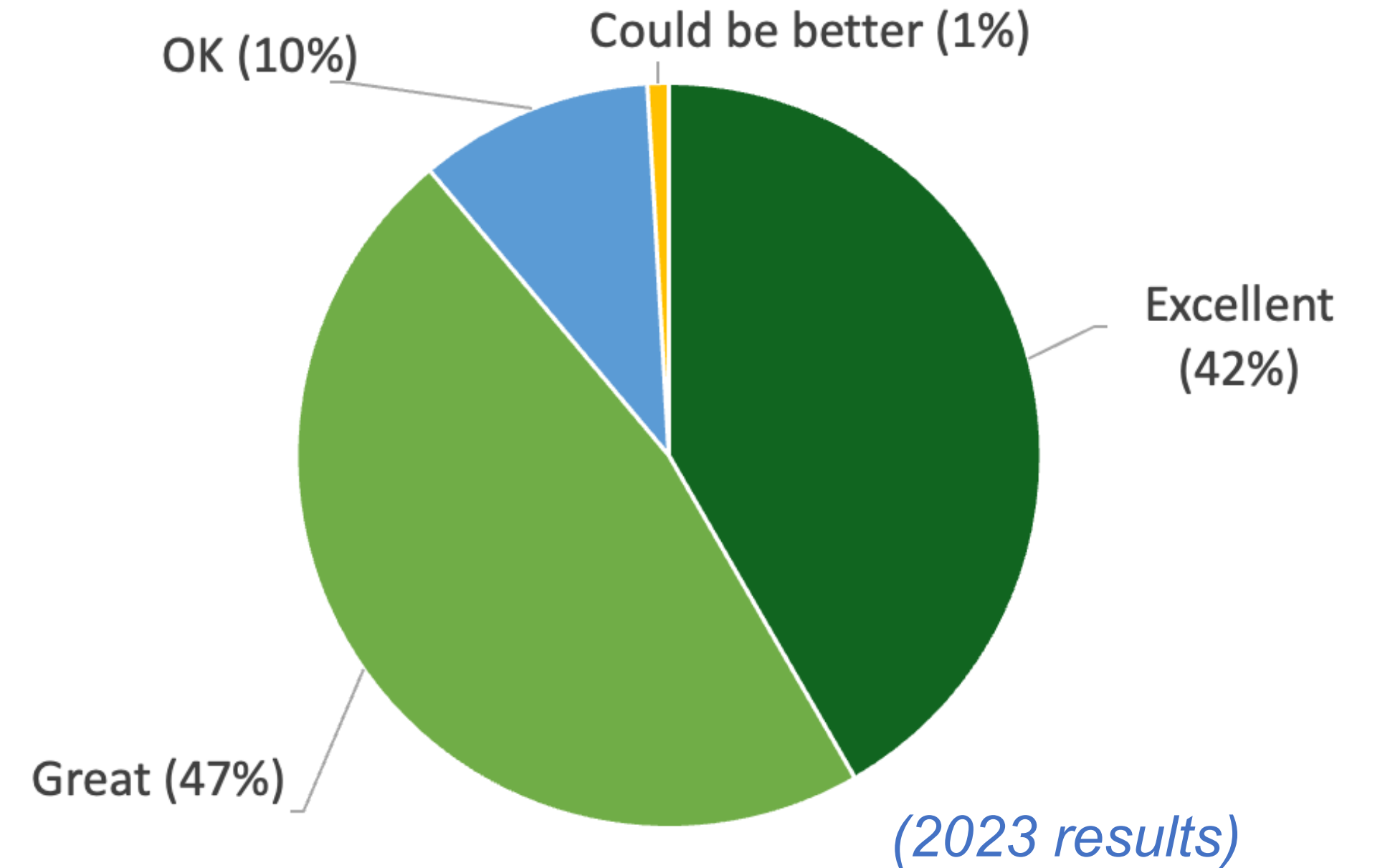
Overall quality of EasyBuild

How would you rate the overall quality of EasyBuild?

- Excellent
- Great
- OK
- Could be better
- Pretty bad

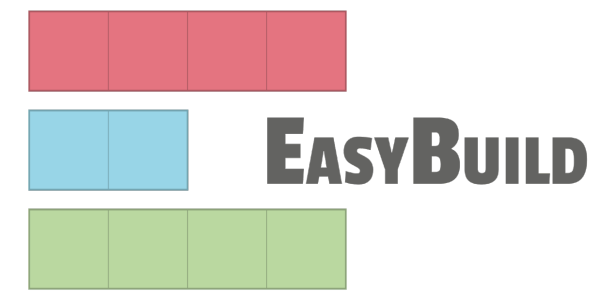*How would you rate the overall quality of EasyBuild?*

OK (10%)
Could be better (1%)
Excellent (42%)
Great (47%)

(2023 results)
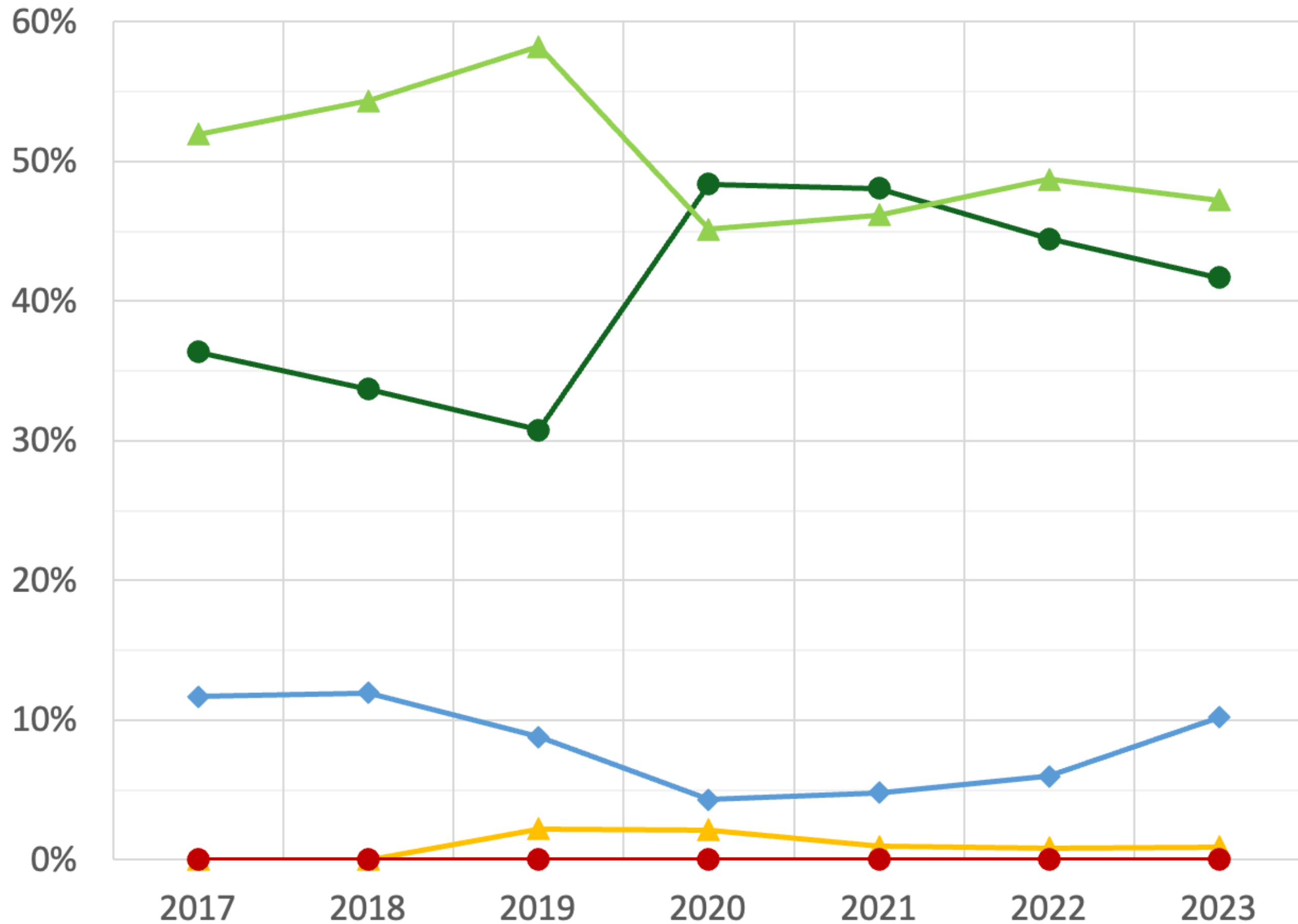
Positive evaluation of overal quality is stable

*(nobody ever answered "pretty bad")*

Goals & challenges (short term)

- **Wrap up efforts towards EasyBuild 5.0, and release it**

- Keep up with increasing number of software installation requests

Software installation requests for HPC-UGent systems

- Continued integration of EasyBuild in EESSI + related efforts

- Support for RHEL9-based systems (incl. OpenSSL 3)

- Revisit environment in which easyconfigs are being tested

  - Currently mainly "bare" Rocky 8 + Rocky 9 (generoso + jsc-zen3)

  - We should definitely also test in EESSI build environment (non-blocking?)

 - Also check whether `--module-only` and `--sanity-check-only` works
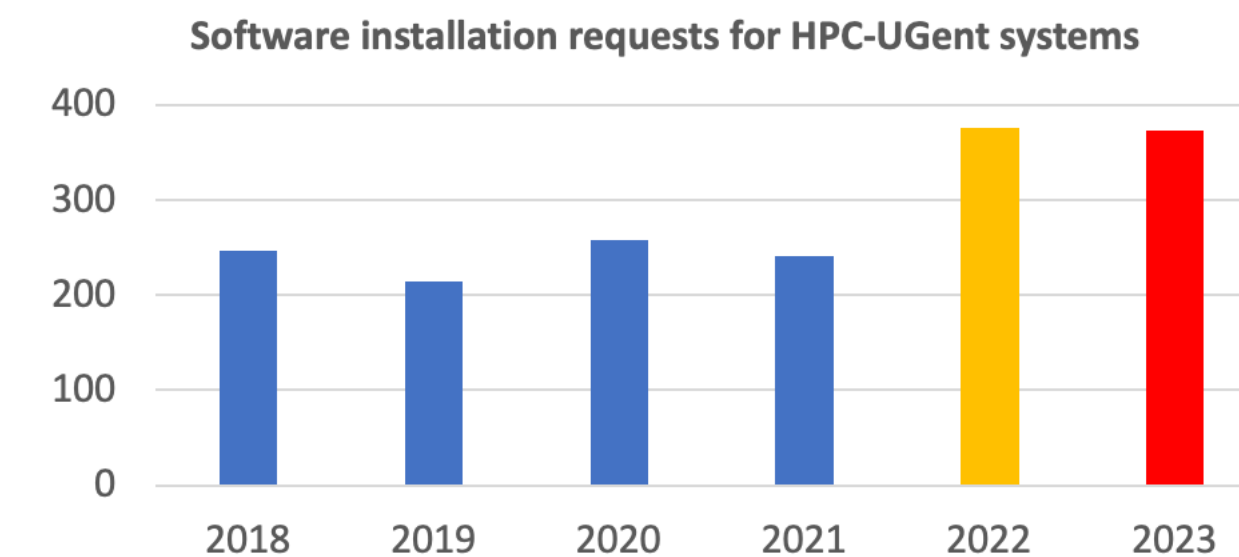
EasyBuild 5.0 (spoilers)
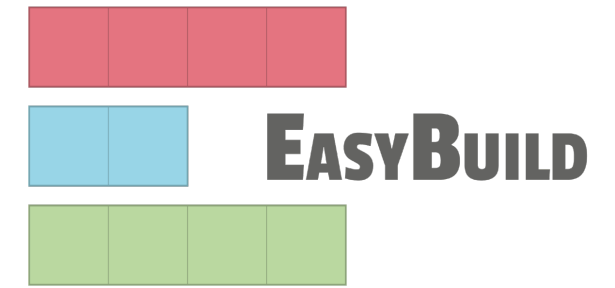
- Overview of changes: docs.easybuild.io/easybuild-v5 (currently still incomplete)

- Effort started in April 2023, still ongoing (but a lot of progress!)

- Merged PRs: ~75 for framework, 170 for easyblocks, 85 for easyconfigs

- Main changes:

  - Drop support for running on top of Python 2.7, require Python 3.6+ and Lmod 8+

  - Deprecate `run_cmd` + `run_cmd_qa`, replaced by `run_shell_cmd`

  - Changed defaults: use RPATH linking, enable `--trace`, use `pip`, …

  - Improve consistency in naming of easyconfig parameters, configuration options, …

  - Adopt toolchain support policy: archive easyconfigs, close PRs touching old easyconfigs/toolchains

Goals & challenges (longer term)
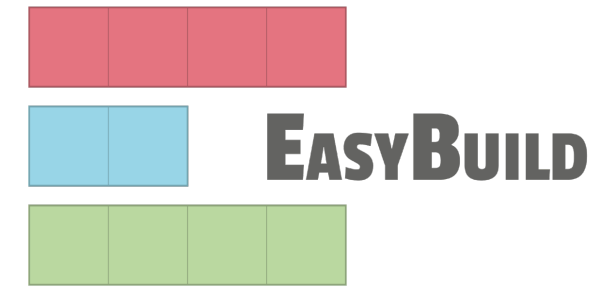
- **Keep up with incoming contributions** (mainly easyconfig PRs, but also beyond)

- Catch up on backlog of PRs (and probably doing some semi-aggressive cleanup)

- Support for installing software on Arm systems (cfr. MultiXscale Task 1.2)
  - Rise of Arm-based HPC systems in Europe (Deucalian, Jupiter, …)
  - Fueled by NVIDIA Grace being adopted

- Support for AMD GPUs & ROCm software stack (cfr. MultiXscale Task 1.2)
  - Fueled by lack of + pricing of NVIDIA GPUs

- What else? Let us know what's important to you!
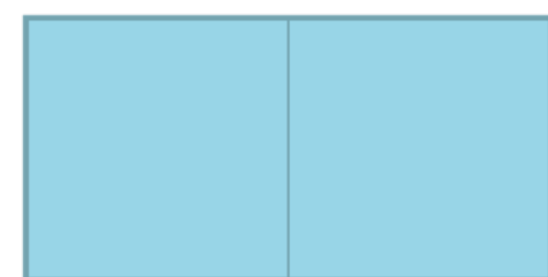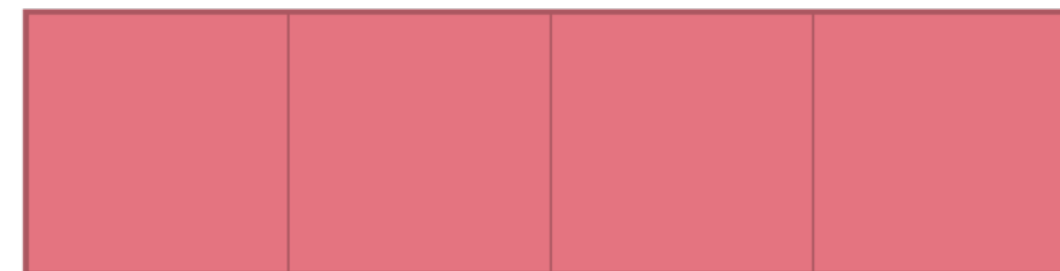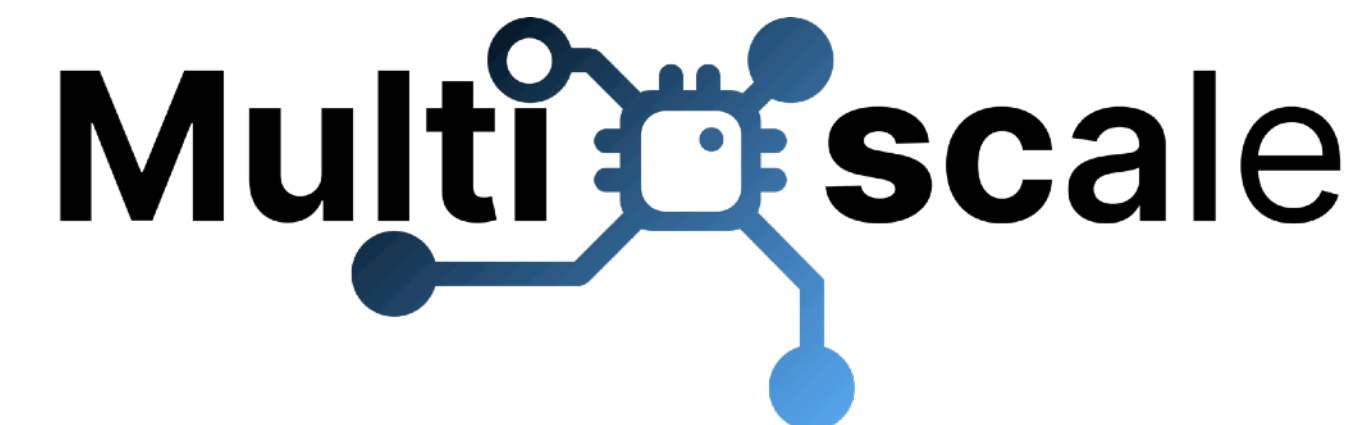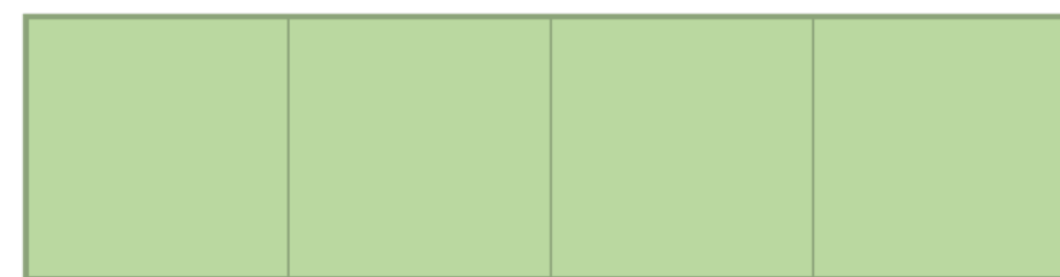
Questions?

@boegel (GitHub)

@kehoste (Twitter)

kenneth.hoste@ugent.be

easybuild@lists.ugent.be

easybuild.io

docs.easybuild.io

tutorial.easybuild.io