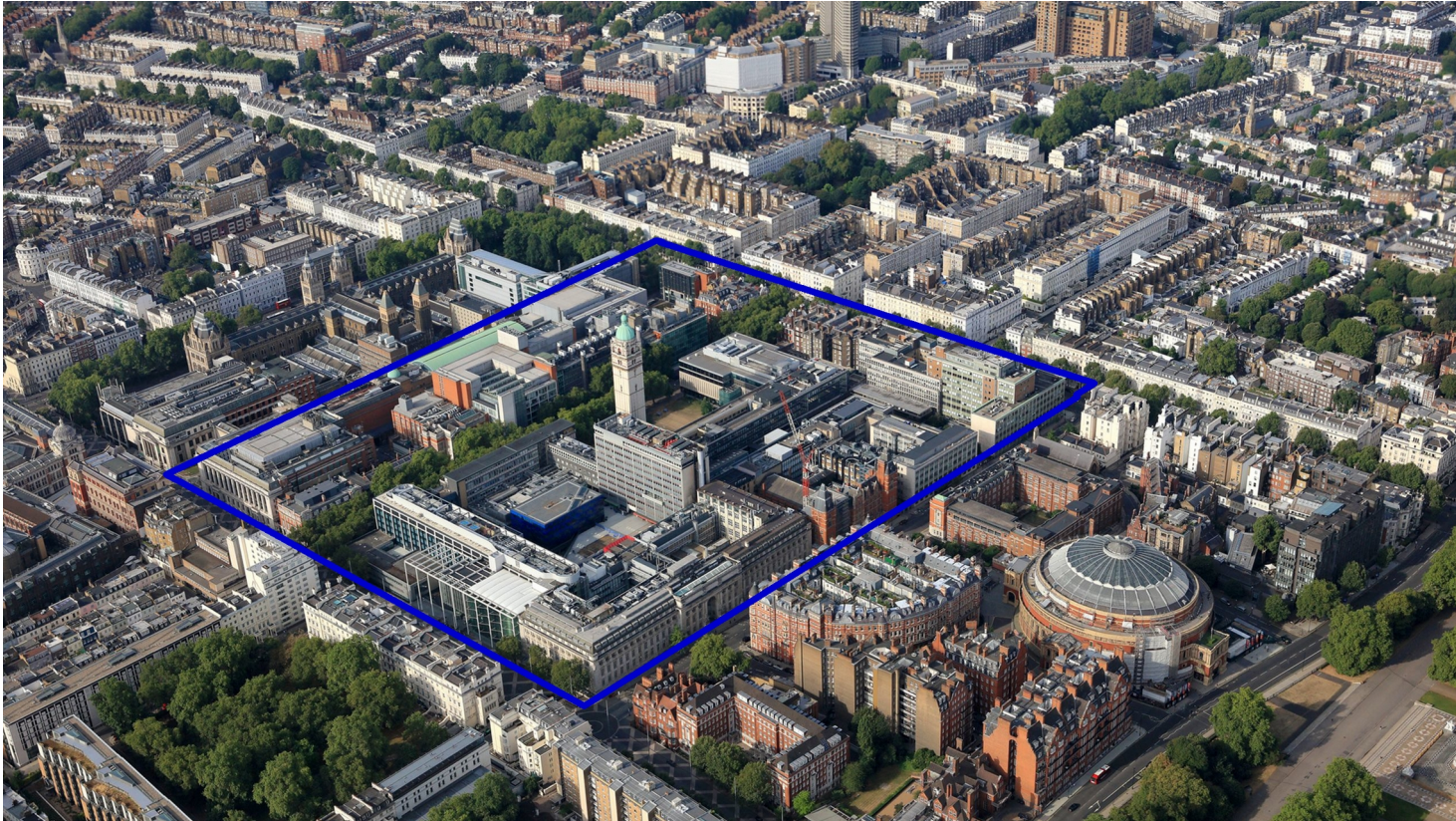
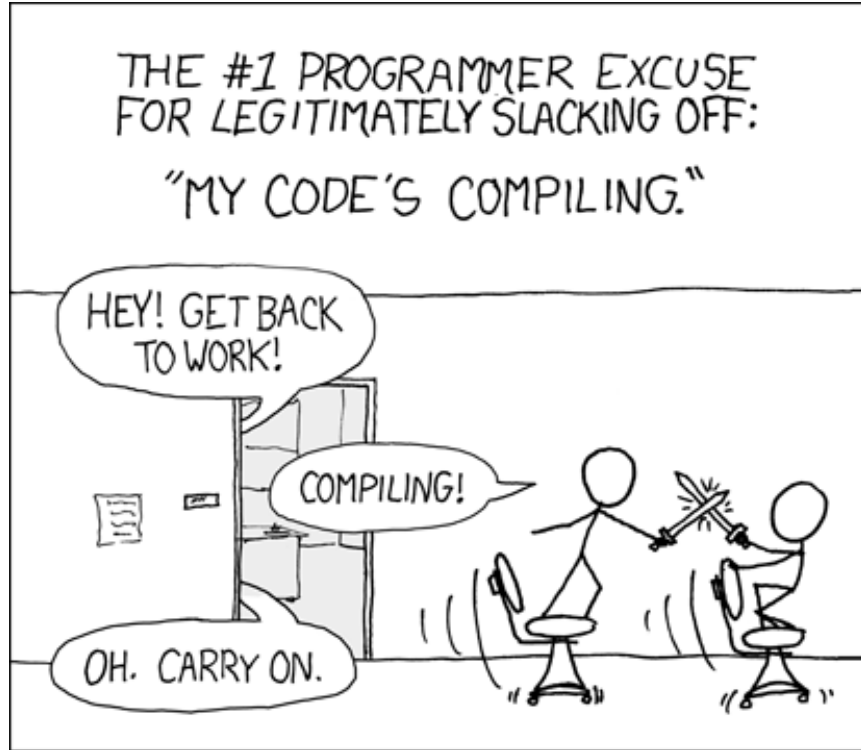
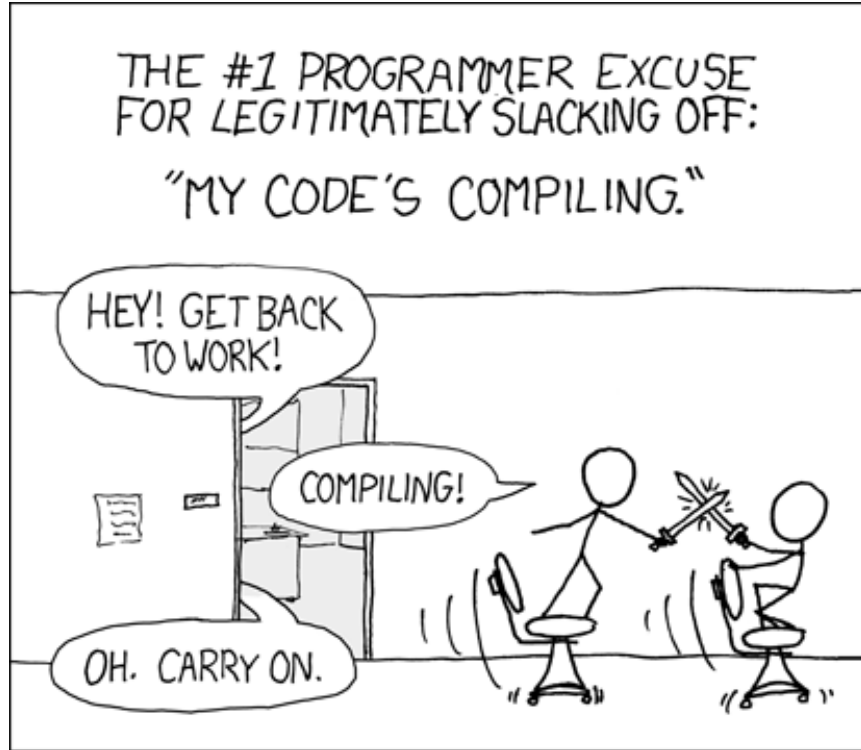


***EasyBob, the friendly software
installation bot***





Source: <https://xkcd.com/303/>



Source: <https://xkcd.com/303/>



CX3

325 compute nodes with 2x AMD EPYC 7742

34 compute nodes with 2x Intel Ice Lake Xeon Platinum 8358 (64 cores, 500GB RAM).

11 GPU nodes with 2x AMD EPYC 7742 (128 cores, 1TB RAM, 8 Quadro RTX 6000 per node)

10 GPU nodes with 2x Intel Skylake 4110 (16 cores, 192GB RAM, 8 Quadro RTX 6000 per node)

12 large mem compute nodes with 2x AMD EPYC 7742 (128 cores, 4TB RAM per node).

Interconnect: 100GbE

Total: 397 compute nodes, 46,880 cores, 403TB RAM, 168 RTX 6000

HX1

303 compute nodes:

Lenovo SD630v2 servers each with 2 x Intel Xeon Platinum 8358 (Ice Lake)

64 cores per node; 288 nodes; 18,432 compute cores; 512 GB RAM per node

GPU nodes:

Lenovo servers each with 4 x NVIDIA A100 80 GB RAM GPUs;

2 x Intel Xeon Platinum 8360Y (Ice Lake) 2.40GHz 36-core processors; 1 TB RAM per node; 15 nodes;

60 GPUs in total

Interconnect: Mellanox ConnectX-6 HDR200 (200 Gbit/s) InfiniBand

Total: 303 compute nodes, 19,512 cores, 159 TB of RAM, 60 A100 GPUs

Research Computing Platforms

The Research Computing Platforms team are responsible for maintaining the research computing platforms, such as the high performance computing and research data store, and for providing support to users in using these platforms.

The team draws together a wide range of research backgrounds (e.g. chemistry, biochemistry, engineering and physics) and research computing support experience (Kings, Oxford, Southampton, Leicester, Warwick and UCL in addition to Imperial). This diversity of experience informs our provision of a responsive, relevant and high quality service to our researchers.

Meet the team



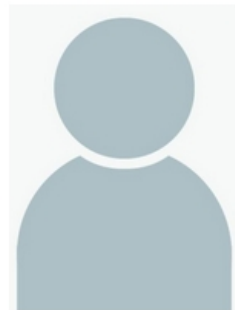
David Hempston
Principal Research
Infrastructure
Engineer



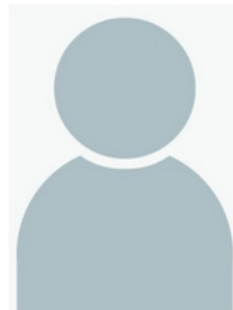
Dugan Witherick
Head of Research
Computing Platforms



**Jörg
Saßmannshausen**
Senior Research
Analyst



Lokesh Ragta
Senior Research
Software Engineer



Rayan Khan
Research Computing
Analyst



Santiago Lacalle Puig
Senior Research
Computing Analyst

Previously:

- installation was done manually
- updating software was tricky as procedure of how to install it was not always recorded
- software installation 'best effort' only due to small team

EasyBuild:

- provides build recipes and thus reproducibility

Problems:

- heterogeneous cluster
- software build needs to be done on the compute nodes → PBS-Pro submission scripts
- team too small to take on additional training required to use EasyBuild efficiently

Solution:

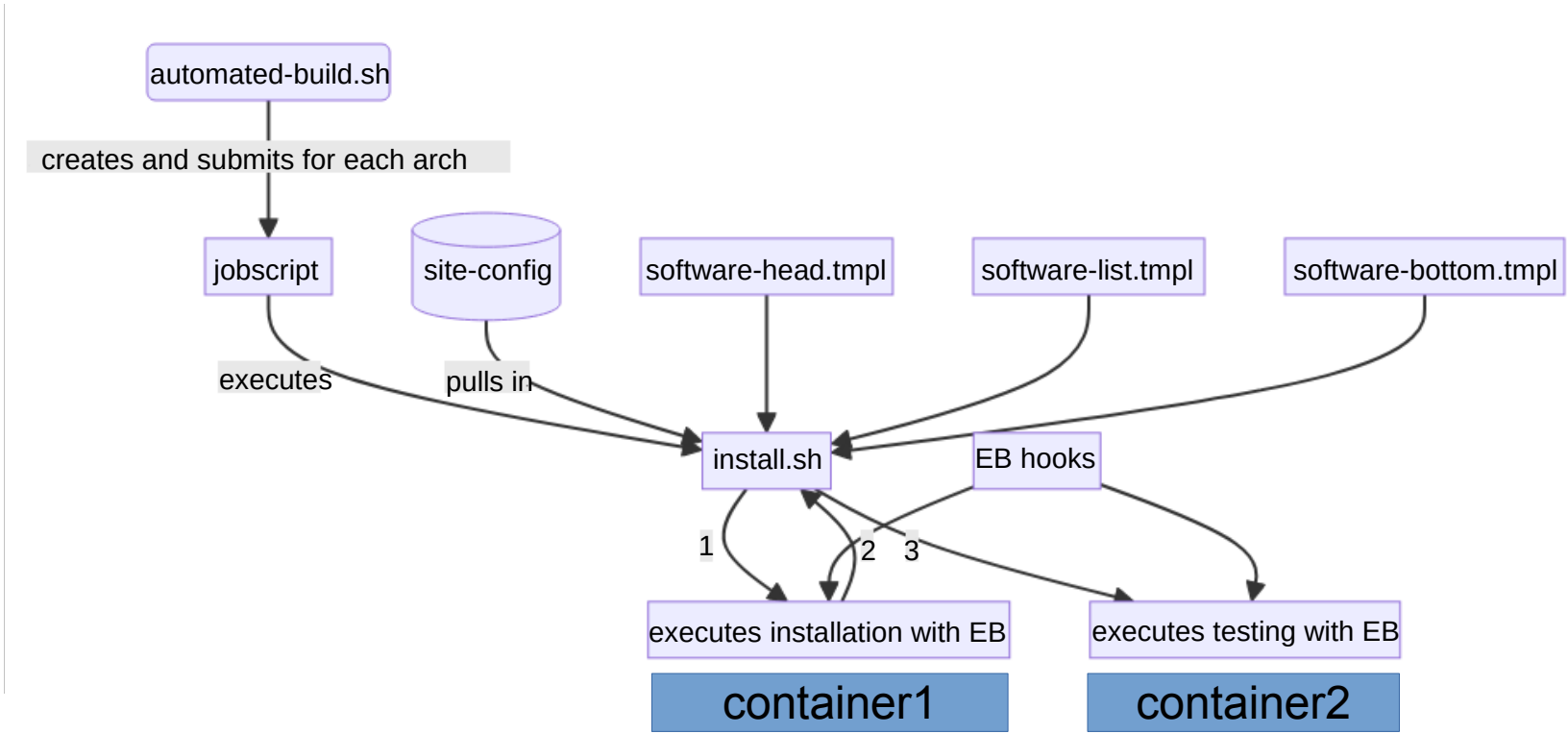
- need to find a way to make the software installation very simple so even 1st line support could use it
- automation!

EasyBob* is a set of scripts which means an operator, or a website, can install scientific software without knowing much more than the name of the EasyConfig file, and some very basic shell commands like how to change directory and create a text file.

It is based on EasyBuild, hence reproducibility etc. is 'build-in'. Due to the vast amounts of already existing EasyConfig files, the turnaround is high and fast.

As we build the software inside a container, we are quite agnostic of our underlying operating system.

The final goal is to use an user-interface, tied to our ticket system, so users can kick off the installation process semi-automatically where the team only has to press the green 'GO' button!



As demonstrated:

- a central configuration file “rules them all”
- updating to a newer version of EasyBuild require only a single change for the whole cluster
- software installation can be monitored live via the EasyBuild log files → good for debugging!
- the log files from the queue capture what usually is printed on the screen

As demonstrated:

- a user-facing GUI, integrated in the ticket system, allows a quick search of software packages, provided the module-class is known
- as a data-base is working in the background, other searches are possible, like keyword for example
- the data-base also allows users to quickly check for any references which should be quoted, provided they are in the EasyConfig file in the first place
- updating the data-base requires just two scripts:
 - one to get the new EasyConfig files from GitHub
 - the other to add them to the data-base

None of this work could have been done without the help and support of friends and colleagues. I am greatly in debt of the whole platforms team, help from the EasyBuild Slack channel, and various conversations over a meal or a beverage.

Special thanks to Bart Oldman for providing a Python script which reads out the EasyConfig file, Trevor Newbury for his SQL knowledge, Imperial for giving me the opportunity to speak at the EUM24 and of course: Bob!

EasyBob can be found on GitHub:

<https://github.com/sassy-crick/easybob>