

**SITE REPORT: THE
FRANCIS CRICK
INSTITUTE**

ADAM HUFFMAN

Senior HPC and Cloud Systems Engineer

The Francis Crick Institute

[@adamhuffman](#)

adam.huffman@crick.ac.uk

MY HISTORY WITH SCIENTIFIC SOFTWARE

Life Sciences at University of Manchester

Mainly RPM packaging of biology/bioinformatics
applications

Fedora packager

MY LATER HISTORY WITH SCIENTIFIC SOFTWARE

Imperial College London

CMS experiment, GridPP, WLCG

Highly domain-specific application provisioning

Mature software environment

CVMFS cf. Compute Canada, ALICE

THE FRANCIS CRICK INSTITUTE

Biomedical discovery institute dedicated to understanding
the fundamental biology underlying health and disease

Biggest biomedical research facility under a single roof in
Europe

1,500 staff, £650m investment

<http://www.crick.ac.uk>

MIGRATION

Moving from 3 sites to 1 new building

Diverse, devolved environments

Data

MY PRESENT WITH SCIENTIFIC SOFTWARE

Inheriting diversity of legacy clusters and experience

Users have very little experience with modules

Symbolic links

Hard-coded profile paths

THE SOLUTION?

...EasyBuild

EASYBUILD FOR EMEDLAB

Large biomedical research cloud

I used EasyBuild for several groups

EASYBUILD USAGE AT CRICK

New CAMP cluster

Mostly `foss-2016b`, some `intel`

683 modules in main area

119 modules in development area

COMMUNITY RESPONSE

Lack of familiarity

Have to introduce the concept and benefits of modules

Some groups took to it quickly, after initial complaints

Benefits and importance of 1:1 training with research groups

SPECIAL CHARACTERISTICS OF BIOINFORMATICS APPLICATIONS

Pipelines relying on specific old software versions

Similar pipeline, two different old versions

Reproducibility?

IMPACT OF EASYBUILD AT CRICK

Made one part of my work feasible

Thanks from research groups

Embedding better practices

EASYBUILD FINDS YOU OUT

Wish I'd been using `hide-deps` from the beginning...

Over-optimised code that won't run on legacy hardware
we've inherited (Intel)

Better training and documentation needed

EMPOWERED ENTHUSIASTS

Enabled fruitful cooperation between admins and research groups

Teething problems with `MODULEPATH` and imperfect knowledge sharing

Cf. work by Markus at JSC

ATTENTION-DEFICIT SCIENCE

Problems owing to upstream software/release quality

Special Interest Group, or revival of HPCBIOS?

Cf Damian's talk - sufficient familiarity for current hierarchical module naming scheme?

GENERAL EASYBUILD EXPERIENCE

Importance of community over strict technical merit

Very welcoming, extremely helpful

Shining example of a distributed open source project

Made my work feasible, and even enjoyable, as opposed to impossible

EASYBUILD OBSERVATIONS

Share tips of the founders

Speed adoption of contrib tools e.g. `easy_update`

Linting

Local testing

`.ebp` patch files

MORE OBSERVATIONS

Better metadata, cf. Markus' plan?

(Further) toolchain documentation improvements

Formalise site customisation methods

THANK YOU