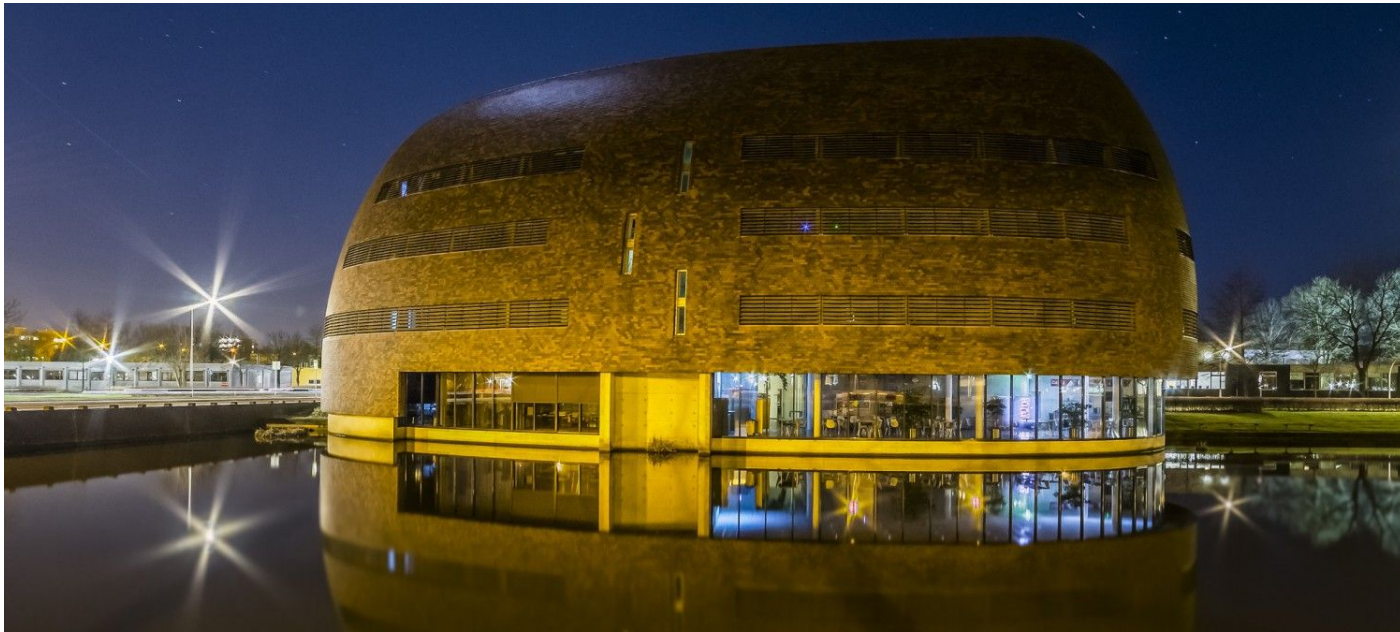




EasyBuild @ RUG

Fokke Dijkstra
Center for Information Technology
University of Groningen
The Netherlands



Peregrine cluster

- > 222 nodes
 - . Different CPU releases:
 - Ivy Bridge, Haswell, Broadwell
 - . 6 GPU (K40) nodes (GPU extension planned)
 - . 7 big memory nodes
- > Lustre file system
- > SLURM scheduler
- > CentOS 7 (upgraded from 6)



Background

- > Started in 2015
- > Before 2015:
 - manual installation, documented on a wiki
- > Growing amount of software installation requests
 - Wide range of users (Physics, chemistry, biology, medicine, arts, sociology, economics)
 - Most users don't know how to install software
- > Limited resources
 - Installation
 - Testing/benchmarking



Toolchains

- > First using locally updated goolfc toolchain
 - . Problem: out of sync with rest of world
 - . Advantage: CUDA included
- > Moved to 2016a toolchains
 - . More configs available
 - . Focused on foss/2016a
- > Lmod used
 - . Flat module tree, grouped by application domain



Updating toolchains

- > Users are confused by toolchains
 - . They just want to use certain software
 - . Mixing toolchains will give issues
 - . Forces us to install more software under multiple toolchains
- > Took us some time to get rid of goolfc
 - . Monitoring usage through Lmod useful
- > Planning to move to 2018a now

Contributing back

- > Contributed multiple easyconfigs to 2016a
- > Not yet looked at easyblocks



Version/dependency hell

- > Toolchain: foss / intel
 - > CUDA: yes or no?
 - > Python: 2 or 3?
 - > Tensorflow: Many versions
 - > Keras: Depending on tensorflow version
-
- > Versioning of underlying libraries unstable



Python

- > Too many packages, versions
- > Outside world converging on Anaconda?
- > Why bother?



Other issues

- > Heterogeneous cluster
 - . Rebuild software for different nodes?
 - . Optimize for lowest common denominator?
 - . Use dynamic architecture for packages that support this?
- > Logging
 - . Finding back reason of compilation errors can be hard

Wishlist/ideas

- > Easier way to handle heterogeneous hardware
- > Easier way to handle different Linux distros
- > Could scale out to University Linux desktops and other clusters.
 - . Distribution channel needs to scale too
- > Make software from different toolchains compatible

Acknowledgments

- > Easybuild team!

- > Locally:
 - Ruben van Dijk
 - Bob Dröge
 - Rest of HPC team