UNIVERSITY OF BIRMINGHAM | RESEARCH SOFTWARE GROUP

EasyBuild at the University of Birmingham

Thursday 28th January 2021

Simon Branford

University of Birmingham and BEAR

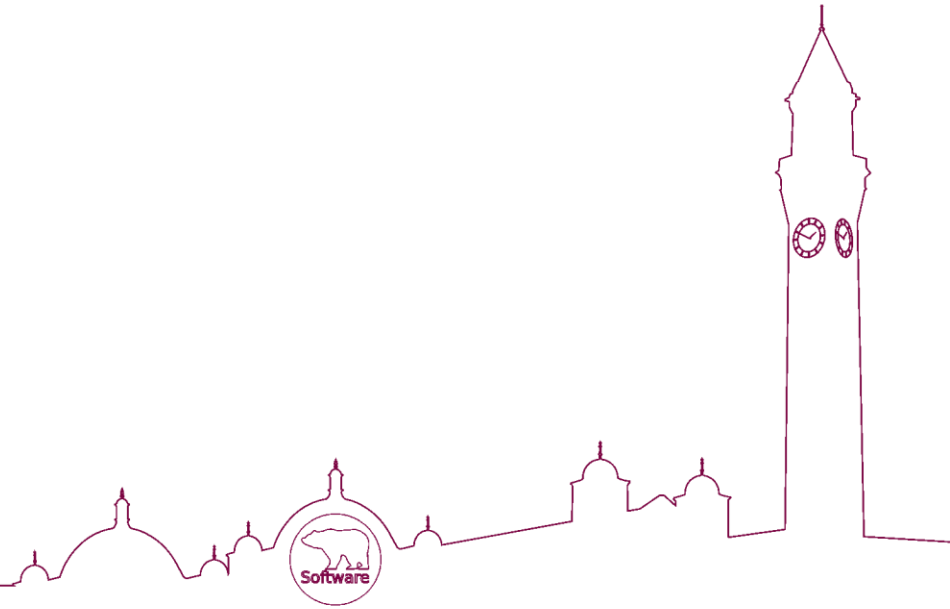
- Birmingham Environment for Academic Research
  - HPC, incl. POWER9 AI Cluster
  - storage, high speed networking, GitLab

BlueBEAR and Our Users

- BlueBEAR is our Linux supercomputer
- Available for free to all researchers at the University
  - We have lots of bioinformaticians
  - Lots of HTC, some HPC
- Groups can purchase additional resources
- BlueBEAR v3
  - started in 2012
  - We do rolling updates – both hardware and software

Hardware and OS

- BlueBEAR
  - Intel (Centos 7)
    - ~~Sandybridge~~
    - Haswell / Broadwell + GPUs (/ SkyLake + GPUs)
    - CascadeLake
  - IBM (RHEL 7)
    - POWER9 + GPUs
- BEARCloud / CaStLeS VMs
  - Intel CPUs (Centos 7 or Ubuntu 16.04)
    - NFS and no-Infiniband

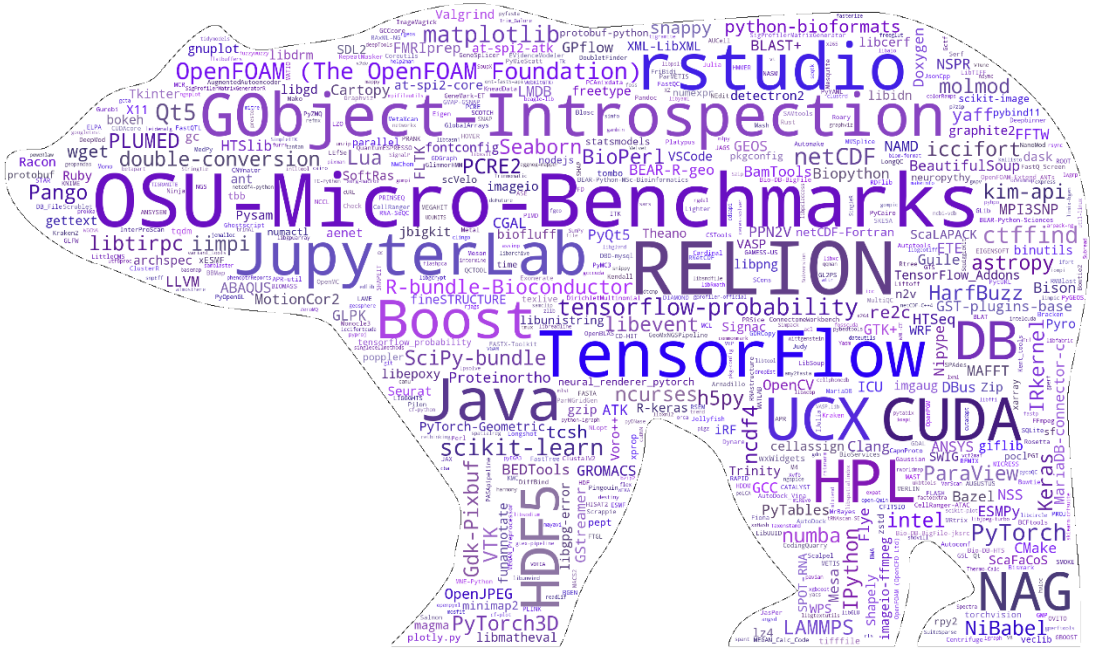


2020 Software Installs

- Over 1000 installations in 2020

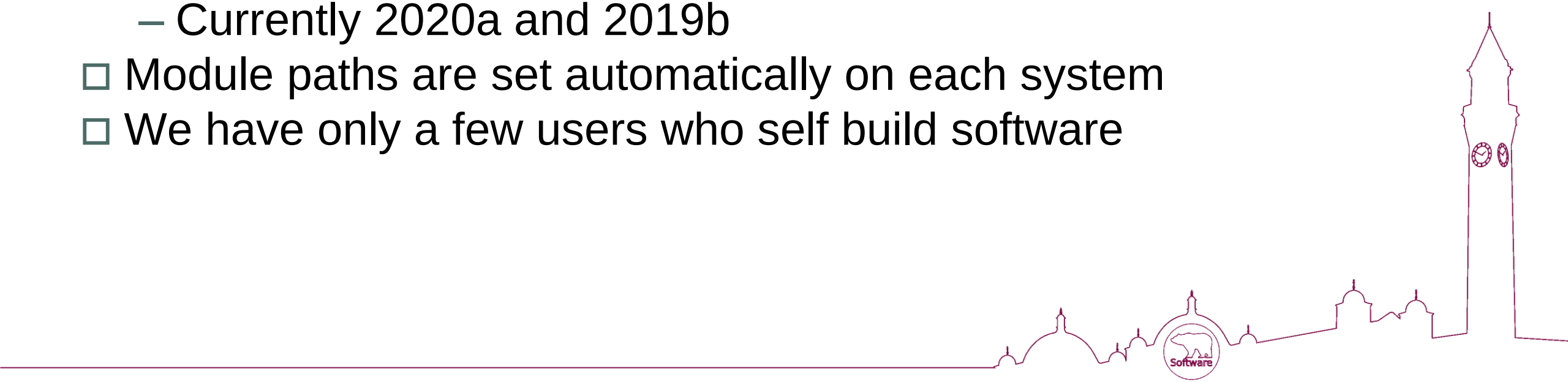
- GCCcore 343
- foss 485
- fosscuda 106
- intel 29
- intelcuda 3

- 2018b 28
- 2019a 92
- 2019b 586
- 2020a 254


rstudio
GObject Introspection
OSU-Micro-Benchmarks
Jupyterlab
RELION
TensorFlow
Boost
Java
UCX
CUDA
HPC
HDF5
NAG
DB
matplotlib
OpenFOAM (The OpenFOAM Foundation)
Seaborn
netCDF
BioPerl
PCRE2
Lua
Pango
wget
double-conversion
libtirpc
imimpi
archspec
LLVM
ABAQUS
MotionCor2
R-bundle-Bioconductor
tensorflow-probability
SciPy-bundle
Proteinortho
scikit-learn
PyTorch-Geometric
tcsh
ncurses
h5py
ncdf4
cellassign
Clang
BEDTools
GROMACS
GCC
PyTables
numba
intel
PyTorch
ScaFaCoS
CMake
ESMPy
NSS
Bazel
ParaView
Keras
TRKernel
MAFFT
Nipype
DBus
Zip
imgaug
ICU
OpenCV
GTK+
WRF
HTSeq
GST-plugins-base
Bison
HarfBuzz
TensorFlow_Addons
astropy
binutils
ctffind
ETE3
Guile
libpng
VASP
PPN2V
netCDF-Fortran
ScaLAPACK
MPI3SNP
kim-api
BeautifulSoup
dask
iccifort
NAMD
graphite2
FFTW
yaff
molmod
NSPR
Doxygen
libcerf
BLAST+
XML-LibXML
GPflow
snappy
python-bioformats
protobuf-python
detectron2
libidn
scikit-image
pybind11
Qt5
X11
bokeh
PLUMED
HTSlib
gc
wget
Racon
protobuf
Ruby
Pango
gettext
Tkinter
libgd
Cartopy
at-spi2-core
at-spi2-atk
FMRIprep
SDL2
libdrm
gnuplot
freetype
LMDB
fontconfig
SoftRas
scVelo
imageio
Pysam
CGAL
numactl
aenet
jbigkit
biofluff
PyQt5
Theano
BEAR-R-geo
BamTools
Biopython
statsmodels
GEOS
pkgconfig
VSCode
ngsjs
tombo
Parallel
Valgrind
Ghostscript
tbb
ant
Eigen
Automake
MISpice
neuropythy
Yasm
libxml2
libxslt
Cython
Mako
Jinja2
MarkupSafe
pytz
Pygments
Sphinx
docutils
alabaster
babel
snowballstemmer
imagesize
packaging
pyparsing
six
requests
certifi
chardet
idna
urllib3
pytest
attrs
more-itertools
pluggy
py
wcwidth
setuptools
wheel
pip
Python
Tcl
Tk
SQLite
libreadline
ncurses
bzip2
zlib
libffi
OpenSSL
XZ
GMP
MPFR
MPC
GCCcore
binutils
GNU
GCC
Clang
LLVM
CMake
Ninja
Meson
pkg-config
Autotools
M4
Bison
flex
help2man
gperf
Perl
Perl-bundle-CPAN
Perl-bundle-Bioconductor
Java
Ant
Maven
Gradle
Scala
Groovy
Kotlin
Clojure
Leiningen
Boost
Eigen
Armadillo
Blitz++
GSL
FFTW
OpenBLAS
LAPACK
ScaLAPACK
BLACS
MPI
Open MPI
MPICH
MVAPICH2
UCX
UCC
libfabric
PMIx
hwloc
libevent
libpciaccess
libxml2
libxslt
libyaml
libpng
libjpeg-turbo
libtiff
libwebp
libgif
giflib
freetype
fontconfig
Pango
Cairo
GTK+
GTK2
GTK3
GLib
GObject Introspection
ATK
at-spi2-core
at-spi2-atk
Gdk-Pixbuf
librsvg
libxml2
libxslt
libyaml
libpng
libjpeg-turbo
libtiff
libwebp
libgif
giflib
freetype
fontconfig
Pango
Cairo
GTK+
GTK2
GTK3
GLib
GObject Introspection
ATK
at-spi2-core
at-spi2-atk
Gdk-Pixbuf
librsvg
VTK
OpenJPEG
minimap2
PyTorch3D
magma
plotly.py
libmatheval
GStreamer
libgpg-error
pept
Voronoi
Seurat
Rtools
PyTorch-Geometric
tcsh
scikit-learn
SciPy-bundle
libepoxy
Proteinortho
neural_renderer_pytorch
tensorflow-probability
GLPK
fineSTRUCTURE
R-bundle-Bioconductor
MotionCor2
ABAQUS
LLVM
archspec
libtirpc
imimpi
gettext
Racon
wget
double-conversion
Lua
PCRE2
Pango
Ruby
protobuf
SoftRas
scVelo
imageio
Pysam
CGAL
numactl
aenet
jbigkit
biofluff
PyQt5
Theano
PPN2V
netCDF-Fortran
ScaLAPACK
MPI3SNP
kim-api
BeautifulSoup
dask
iccifort
NAMD
graphite2
FFTW
yaff
molmod
NSPR
Doxygen
libcerf
BLAST+
XML-LibXML
GPflow
snappy
python-bioformats
protobuf-python
detectron2
libidn
scikit-image
pybind11
Qt5
X11
bokeh
PLUMED
HTSlib
gc
wget
Racon
protobuf
Ruby
Pango
gettext
Tkinter
libgd
Cartopy
at-spi2-core
at-spi2-atk
FMRIprep
SDL2
libdrm
gnuplot
freetype
LMDB
fontconfig
SoftRas
scVelo
imageio
Pysam
CGAL
numactl
aenet
jbigkit
biofluff
PyQt5
Theano
BEAR-R-geo
BamTools
Biopython
statsmodels
GEOS
pkgconfig
VSCode
ngsjs
tombo
Parallel
Valgrind
Ghostscript
tbb
ant
Eigen
Automake
MISpice
neuropythy
Yasm
libxml2
libxslt
Cython
Mako
Jinja2
MarkupSafe
pytz
Pygments
Sphinx
docutils
alabaster
babel
snowballstemmer
imagesize
packaging
pyparsing
six
requests
certifi
chardet
idna
urllib3
pytest
attrs
more-itertools
pluggy
py
wcwidth
setuptools
wheel
pip
Python
Tcl
Tk
SQLite
libreadline
ncurses
bzip2
zlib
libffi
OpenSSL
XZ
GMP
MPFR
MPC
GCCcore
binutils
GNU
GCC
Clang
LLVM
CMake
Ninja
Meson
pkg-config
Autotools
M4
Bison
flex
help2man
gperf
Perl
Perl-bundle-CPAN
Perl-bundle-Bioconductor
Java
Ant
Maven
Gradle
Scala
Groovy
Kotlin
Clojure
Leiningen
Boost
Eigen
Armadillo
Blitz++
GSL
FFTW
OpenBLAS
LAPACK
ScaLAPACK
BLACS
MPI
Open MPI
MPICH
MVAPICH2
UCX
UCC
libfabric
PMIx
hwloc
libevent
libpciaccess

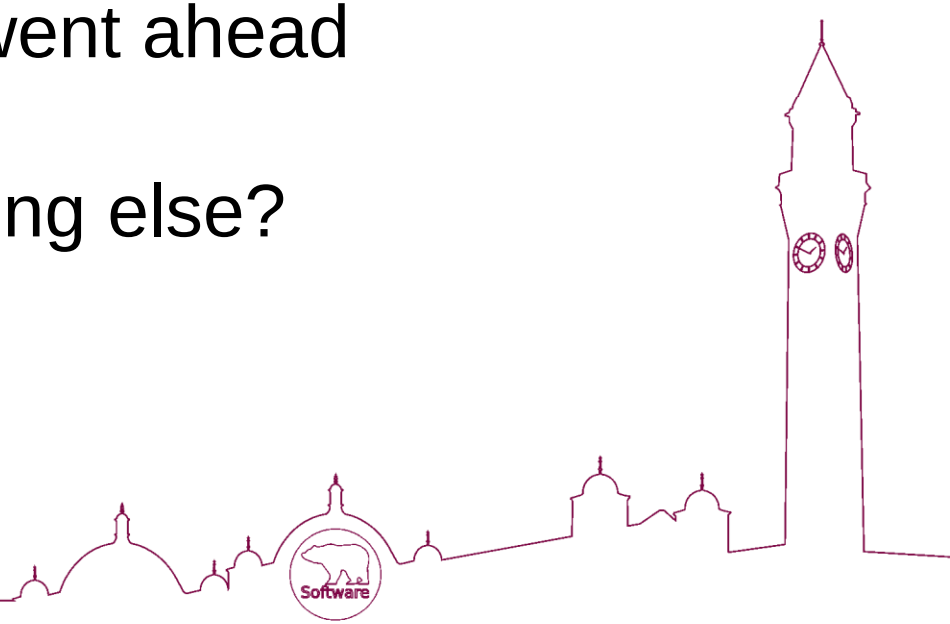
EasyBuild at Birmingham

- Started using EasyBuild in 2016
- All installations are by request
  - We decide on which toolchain to use for the installation
- Generally we have two active toolchains generations for installation
  - Currently 2020a and 2019b
- Module paths are set automatically on each system
- We have only a few users who self build software



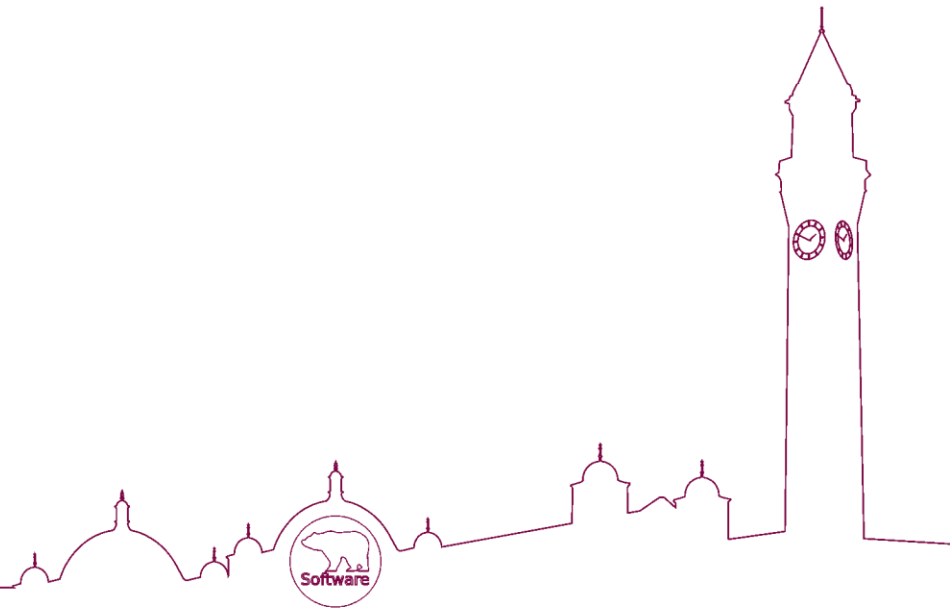
OS Upgrade

- Upgrade to CentOS 8 / RHEL 8
  - Installed (nearly) all existing software
    - 2018b (x86_64) and 2019a (POWER9)
- CentOS announcement ~1 week before planned upgrade
  - Upgrade of GPU nodes and POWER9 went ahead
  - CPU-only nodes remain on CentOS 7
- SLES, RHEL, OpenSUSE, Rocky, something else?
  - We've not yet decided…



MPI

- Intel MPI issues
  - Crashes on CentOS 8 (#11762)
  - MLX provider in libfabric not working (#10213)
- Intel MPI: 2% of our 2020 installs use it
- Lots of time for very few applications
- Decided that we'd go with iomkl
  - Already using OpenMPI



Open OnDemand

- “Open OnDemand is an NSF-funded open-source HPC portal based on OSC’s original OnDemand portal. The goal of Open OnDemand is to provide an easy way for system administrators to provide web access to their HPC resources”
  - https://openondemand.org/

- GUI Applications: ANSYS, ABAQUS, MATLAB, ParaView, …
- JupyterLab, RStudio Server, VS Code
- Shell access, File browser

Baskerville

- National Tier 2 Accelerated Compute Facility
- Funded following the 2019 Tier 2 HPC round
- “Fundable” but insufficient funding at the time
- Led by University of Birmingham – Professor Iain Styles
- Housed in our dedicated research data centre

Research Partners

University of Birmingham

diamond

The Rosalind Franklin Institute

The Alan Turing Institute

UKRI Engineering and Physical Sciences Research Council

Technology Partners

NVIDIA

Lenovo

OCF

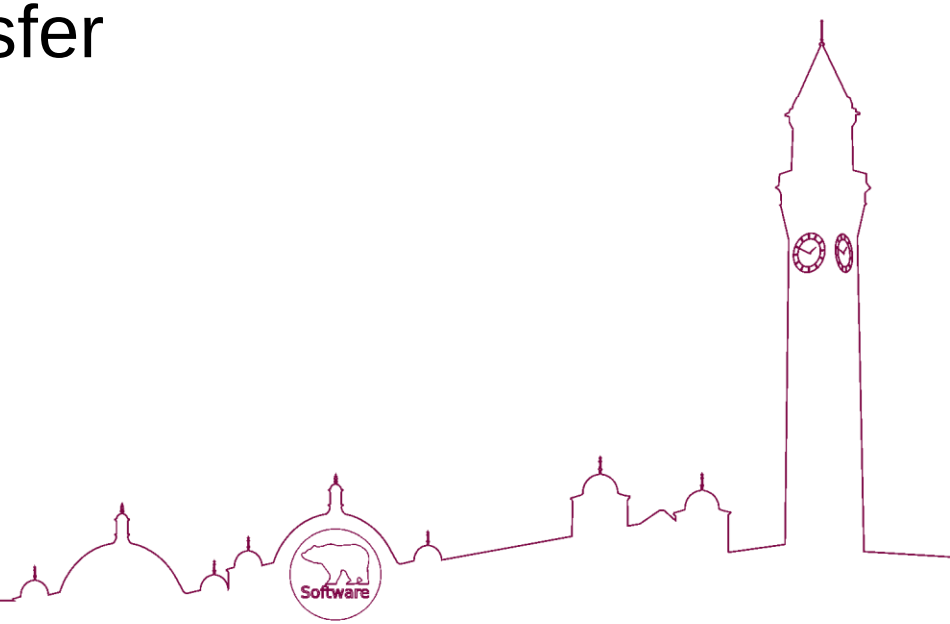
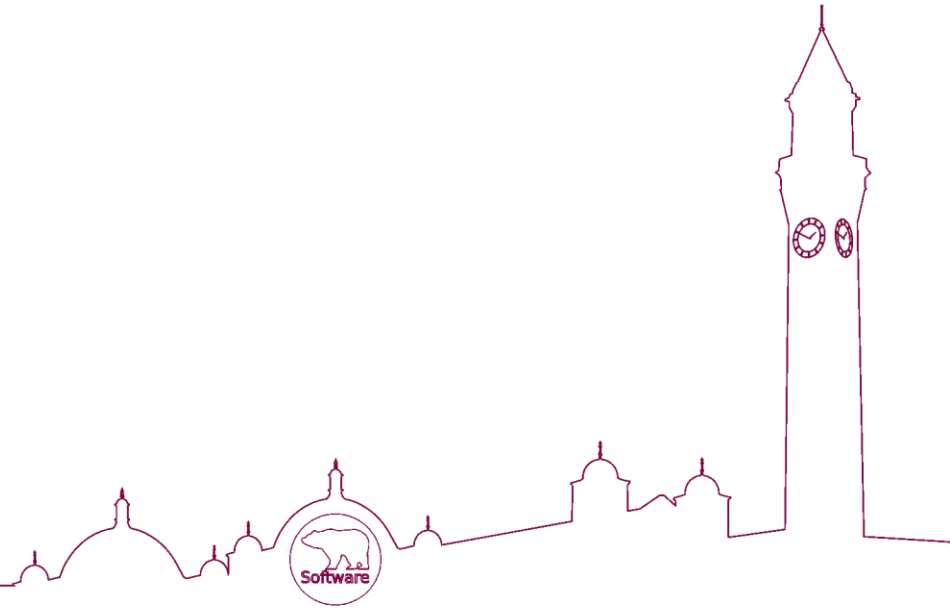
Baskerville: Compute Nodes

- 46 1U systems, each with:
  - 2x Intel Xeon next generation CPUs
  - 512GB RAM
  - 960GB NVMe SSD
  - HDR Infiniband (200Gbps)
  - 4x NVIDIA A100-40 GPUs, with NVLINK
  - Lenovo Liquid cooled
  - Plate rating: 53.4 TFlops per node

Baskerville: Storage and Data Transfer

- Lenovo DSS-G – running IBM Spectrum Scale
  - ~500TB usable flash (96 drives)
  - ~5.2PB usable spinning disk (418 drives)

- Will include Globus end point for data transfer



Thanks