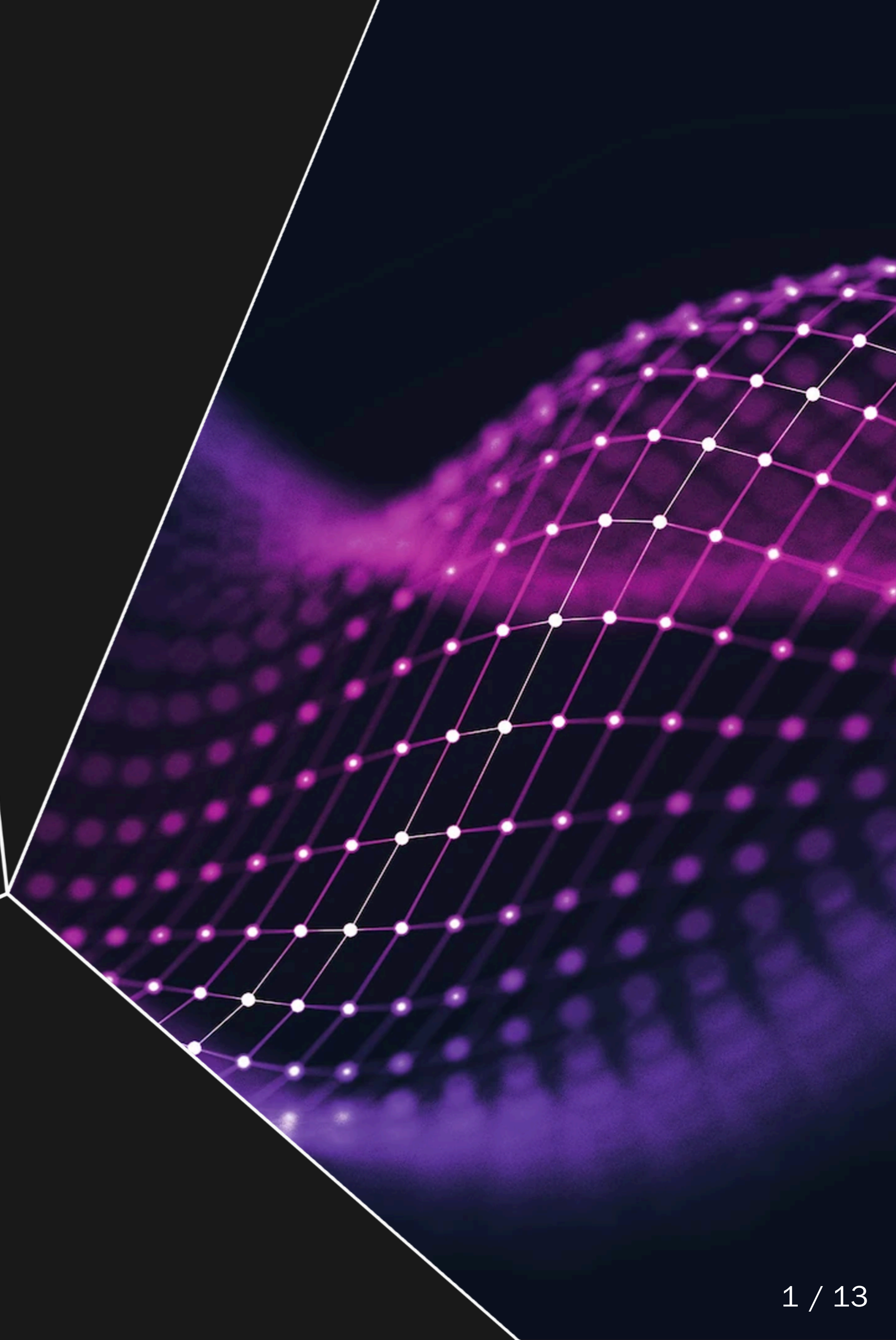
Optimising R code for BlueBEAR

Research Software Group
Advanced Research Computing



UNIVERSITY OF BIRMINGHAM

R User Group, May 2026

Optimising R code for BlueBEAR

Topics

- R packages on BlueBEAR
- R parallel workflows

What is BlueBEAR?

- UoB’s “tier 3” HPC (high-performance computing) system
  - Also HTC workflows: high-throughput computing
- It’s heterogeneous:
  - What does this mean?
  - Why is this important for R package installs?
- BEAR Portal ↗ = graphical frontend to BlueBEAR (sort of).

R Package Installs: part 01

Do you need to install?

- Check if we’ve already installed package via a module.
    - https://bear-apps.bham.ac.uk [demo]
    - N.B. changes `2023a` and above.
- RStudio (with “meta” modules).
- JupyterLab – R kernels with modules tool. [demo]

R Package Installs: part 02

R package install: process overview

> Don’t use RStudio!

- Start an interactive job.
  - Where to run? (See “Heterogeneous issues” below.)
  - How long for?
  - Alternatively, use a batch job to run the install. (Reproducible!)
- Follow the documentation here: https://docs.bear.bham.ac.uk/bluebear/software/self_installs_r/

> Heterogeneous issues
>
> - Pay close attention to the node type on which you’re running the session.
> - You may need to install the R packages multiple times.

R Package Installs: part 03

Other options⋯

- R Conda environments
  - https://astrobiomike.github.io/R/managing-r-and-rstudio-with-conda
  - The same heterogeneous issues apply!
- Python dependency stack
  - *reticulate* package.
  - The same heterogeneous issues apply!

R Parallel Jobs: part 01

- What speeds-up a job?
  - More resource…
    - … if the script is written to parallelise.
    - … if the workflow is not storage-bound.
  - Putting the data in the correct location.
    - RDS (network storage) is relatively slow.
    - Local (`/scratch`) storage is much faster.
    - … but memory is also fast, although has less capacity.

R Parallel Jobs: part 02

- Note that array jobs may be a good solution for some types of workflow.
  - https://docs.bear.bham.ac.uk/bluebear/array_jobs/
- Our recommendation: single-node parallelisation is best for R.
- Accurate detection of cores: https://docs.bear.bham.ac.uk/bluebear/software/guides/R/#accurate-detection-of-cores

```
library(parallelly)
library(parallel)
# Detect how many cores are available for R to use
n <- availableCores()
cat("Using", n, "cores\n")
# Create a PSOCK cluster using all available cores
cl <- makeClusterPSOCK(n)
# Function each worker will run
hello_fun <- function(i) {
paste("Hello from worker", i, "on", Sys.info()[["nodename"]])
}
# Run in parallel
results <- parLapply(cl, seq_len(n), hello_fun)
print(results)
# Shut down the cluster
stopCluster(cl)
```

R Parallel Jobs: part 03 – existing tools

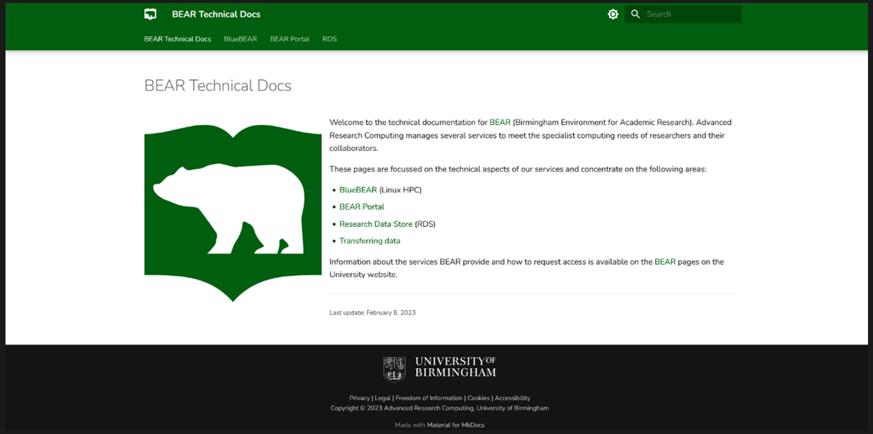
- https://mirai.r-lib.org/ ↗
  - Cluster config: https://mirai.r-lib.org/reference/cluster_config.html ↗
- future ↗
  - Slurm example: https://future.batchtools.futureverse.org/ ↗
- BiocParallel
  - Available on BlueBEAR: https://apps.bear.bham.ac.uk/applications/R-bundle-Bioconductor/ ↗

Miscellaneous

- Batch jobs over interactive RStudio / JupyterLab.
- Why has my R Studio session died?!
  - OOM-kills.
    - Accumulation of variables.
    - Cores and memory.
- GPU jobs:
  - Batch job or interactive job only. No RStudio.
  - We don’t have any R modules built with GPU capability.
    - Historically, much of the GPU functionality was provided via Python packages. (See again *reticulate*).

Support

BEAR Technical Docs



BEAR Technical Docs

Welcome to the technical documentation for BEAR (Birmingham Environment for Academic Research). Advanced Research Computing manages several services to meet the specialist computing needs of researchers and their collaborators.

These pages are focussed on the technical aspects of our services and concentrate on the following areas:

- BlueBEAR (Linux HPC)
- BEAR Portal
- Research Data Store (RDS)
- Transferring data

Information about the services BEAR provide and how to request access is available on the BEAR pages on the University website.

Last update: February 8, 2023

UNIVERSITY OF BIRMINGHAM

Privacy | Legal | Freedom of Information | Cookies | Accessibility
Copyright © 2023 Advanced Research Computing, University of Birmingham

Made with Material for MkDocs

BlueBEAR Technical Documentation Site: https://docs.bear.bham.ac.uk

- The BEAR Technical Docs website should be your first resource for BEAR questions
- Continually updated…
  - … and please let us know if anything is unclear or needs further explanation

IT Service Desk

- https://www.birmingham.ac.uk/bear/sd/other-bear-request
  - Questions and issues running jobs
  - Advice
- https://www.birmingham.ac.uk/bear/sd/new-bear-software
  - Use to request new software on BlueBEAR with a link to the software and preferred version

> **Visibility for Masters students**
>
> Masters students cannot see some research tickets so should submit a general ticket, clearly stating that it’s for BEAR.

Any Questions??