Parallel Programming in R

- Generic embarrassingly parallel

  - **parallel** package: multi-core, socket clusters

  - **future** package (integrates with other backends)

- "Free" parallelism

  - Numerical libraries (e.g. LAPACK, BLAS, vecLib)

- Model-specific implementations (random forest, boosting)

Parallel Programming in R

- Assume tasks are fairly homogeneous

- Assume resources are fairly homogeneous

- Non-interruptible — if something happens all is lost

- Require pre-allocation of resources

- Okay for modern multi-core desktop systems and common use cases