Parallel Computing in R using NetWorkSpaces
N Carriero, J Lai, M Schultz, S Weston and G Warnes

Supported by:
Yale Center for High Performance Computation in Biology and Biomedicine and NIH grant: RR19895-02
Pfizer

Shared Workspaces

- Variation on the theme of a workspace.
- The NetWorkSpace object encapsulation uses an Internet-based server to hold the workspace.
- A given NetWorkSpace can be accessed by multiple processes: *Any process capable of instantiating an appropriate NetWorkSpace object may retrieve the value of a variable.* (Or store (name, value) pairs for that matter.)
Coordination via NetWorkSpaces

- **Shared Access:** Communication.
- **Blocking References:** Synchronization.
- Coordination provided within the context of the existing, familiar concept of a "workspace".
- Coordination data has independent existence

**Benefits**

- **Simplifies development:**
  - Familiar conceptual foundation
  - Uncoupling in space and time
  - Anonymity
- **Promotes flexibility:**
  - Dynamic processing ensembles
  - Cross platform
  - Cross environment
Sleigh

- Inspired by snow (Tierney, Rossini, Li, Sevcikova), but snow and sleigh differ in many ways.
- Supports “parallel” apply.
- Implemented on top of NetWorkSpaces.
- Vehicle for launching codes that explicitly use NetWorkSpaces for coordination.
• MATLAB, octave, python, perl, ruby, …

• Software available from: http://nws-r.sourceforge.net
(open source for open source systems; commercial for commercial systems: www.lindaspaces.com)

• API used in this talk is a “teaser”. More serious projects use a richer, but more verbose, API.