Parallel Computing with R using GridRPC

Junji Nakano[1,*], **Ei-ji Nakama**[2]

1. The Institute of Statistical Mathematics, Tokyo, Japan
2. COM-ONE Ltd. Ishikawa, Japan

*Contact author: nakanoj@ism.ac.jp

Keywords: Grid computing, Heterogeneous clusters, Ninf-G

Parallel computing becomes popular presently to achieve massive calculations. We have several techniques to perform parallel computing with R, for example, **Rmpi** and **snow**. These packages provide flexible and stable parallel computing mechanisms, and are especially suitable for a cluster of homogeneous computers in an intra-network. Grid computing, on the other hand, appeared recently to use simultaneously several heterogeneous computer clusters which are located far away and connected by Internet. **GridR** is one package to use R in such a grid environment. We propose to use GridRPC, a remote procedure call API for grid computing, for adding parallel computing functions to R. We provide a package to realize **snow**-like functions by utilizing Ninf-G, a reference implementation of the GridRPC API.

References

Ninf administration group (2008). Ninf: A Global Computing Infrastructure, `http://ninf.apgrid.org/`.

Schmidberger, M, Morgan, M., Eddelbuettel, D., Yu, H., Tierney, L. and Mansmann, U. (2009). State of the Art in Parallel Computing with R. *Journal of Statistical Software*, August 2009, Vol. 31, Issue 1.