Guide to the **pbdPROF** Package

Gaurav Sehrawat[1], Wei-Chen Chen[2], Drew Schmidt[3], Pragneshkumar Patel[3], and George Ostrouchov[2,3]

[1]Jaypee Institute of Information Technology
Uttar Pradesh, India

[2]Computer Science and Mathematics Division,
Oak Ridge National Laboratory,
Oak Ridge, TN, USA

[3]Remote Data Analysis and Visualization Center
University of Tennessee,
Knoxville, TN, USA

July 17, 2013

Contents

Acknowledgement

Sehrawat was generously supported by Google for Google Summer of Code 2013.

Chen and Ostrouchov were supported in part by the project "Visual Data Exploration and Analysis of Ultra-large Climate Data" funded by U.S. DOE Office of Science under Contract No. DE-AC05-00OR22725. Ostrouchov, Schmidt, and Patel were supported in part by the project "NICS Remote Data Analysis and Visualization Center" funded by the Office of Cyberinfrastructure of the U.S. National Science Foundation under Award No. ARRA-NSF-OCI-0906324 for NICS-RDAV center.

Warning: The findings and conclusions in this article have not been formally disseminated by the U.S. Department of Energy and should not be construed to represent any determination or policy of University, Agency, and National Laboratory.

This document is written to explain the main functions of **pbdPROF** (Sehrawat *et al.*, 2013), version 0.1-0. Every effort will be made to ensure future versions are consistent with these instructions, but features in later versions may not be explained in this document.

Information about the functionality of this package, and any changes in future versions can be found on website: "Programming with Big Data in R" at `http://r-pbd.org/`.

1 Introduction

The main features of **pbdPROF** are:

1. ...

1.1 System Requirements

pbdPROF requires an MPI installation and an MPI-using package, such as **pbdMPI** (Chen *et al.*, 2012a) or **Rmpi** (Yu, 2010). For information regarding how to install MPI or **pbdMPI**, please see the **pbdMPI** vignette (Chen *et al.*, 2012b) or the pbdR website `http://r-pbd.org/`.

2 Installation

WCC: Need references and quick installations.

The **pbdPROF** currently is by default using **fpmpi** (**?**) library internally, i.e., a source copy of **fpmpi** is located at **pbdPROF/src/fpmpi** and built in a static library at **pbdPROF/lib/libfpmpi.a**. However, external profiler libraries such as **fpmpi**, **mpiP** (**?**), and **TAU** (**?**) can be also linked by **pbdPROF** via suitable `--configure-args` to `R CMD INSTALL`. We explain the whole procedure in Section 2.1 using **fpmpi** as an example and leave some keys steps for **mpiP** and **TAU** in Sections 2.2 and 2.3.

2.1 fpmpi

Using internal **fpmpi** library, via

Shell Command

```
R CMD INSTALL pbdPROF_0.1-0.tar.gz
```

By default, this compiles `src/fpmpi/*`, generates a static library `libfpmpi.a`, and installs the library to `pbdPROF/lib/`. No shared library is generated or needed, so the directory `pbdPROF/libs/` is empty (no need to build `pbdPROF.so`.) The linking argument is saved in `Makeconf` and installed to `pbdPROF/etc/` for further linking such as **pbdMPI** is reinstalled with `--enable-pbdPROF`.

Linking with external **fpmpi** library, via

Shell Command

```
R CMD INSTALL pbdPROF_0.1-0.tar.gz \
  --configure-args="--with-fpmpi='-L/path_to_fpmpi/lib -lfpmpi'"
```

This only provides the linking arguments `-L/path_to_fpmpi/lib -lfpmpi` which is saved in `Makeconf` and installed to `pbdPROF/etc/` for further linking such as **pbdMPI** is reinstalled with `--enable-pbdPROF`.

Reinstall **pbdMPI**, **pbdSLAP**, and **pbdNCDF4**, via

Shell Command

```
R CMD INSTALL pbdMPI_1.0-0.tar.gz --configure-args="--enable-pbdPROF'"
```

Note that the `pbdMPI/R/get_conf.r` and `pbdMPI/R/get_lib.r` are used in `pbdMPI/configure.ac` or `pbdMPI/configure` to determine an appropriate linking flag `PROF_LDFLAGS` based on preset flags in `pbdPROF/etc/Makeconf`.

If the internal library is used in **pbdPROF**, then the path to the `pbdPROF/lib/libfpmpi.a` is set in the flag `PKG_LIBS` of `pbdMPI/src/Makevars.in`. If the external library is used in **pbdPROF**, then the linking arguments `-L/path_to_fpmpi/lib -lfpmpi` is set in the flag `PKG_LIBS` of `pbdMPI/src/Makevars.in`. Therefore, the **pbdMPI** can be intercepted by the **fpmpi** library when MPI function calls are evoked.

No mater the external or internal library is used, the `PROF_LDFLAGS` in `pbdMPI/etc/Makefile` provides the linking information to the profiler library. It is also used in `PKG_LIBS` which will be export to other pbdR packages at installation via the flag `SPMD_LDFLAGS`, therefore, no need to add further flags to `R CMD INSTALL` when reinstall packages for further profiling.

For further profiling, such as **pbdSLAP** and **pbdBASE**, one may reinstall both packages, via

Shell Command

```
R CMD INSTALL pbdSLAP_0.1-6.tar.gz
R CMD INSTALL pbdBASE_0.2-2.tar.gz
```

Note that since both packages have MPI C functions involved, it is necessary to link with profiler library in order to profile communications evoked by both packages.

For profiling **pbdNCD4**, one may need to recompile **HDF5** and **netCDF4** libraries and link with profiler library, since those are the MPI functions involved such as MPI-IO. Also, as recompile finish, one may reinstall the package, via

Shell Command

```
R CMD INSTALL pbdNCDF4_0.1-1.tar.gz
```

2.2 mpiP

2.3 TAU

2.4 Test Script

Below we provide sample scripts to test that the installation of **pbdPROF** was successful. For **pbdMPI**, use:

Test script for pbdMPI

```
library(pbdMPI)
init()

set.seed(comm.rank())
x <- allreduce(rnorm(100))

finalize()
```

and for **Rmpi**, use:

Test script for pbdMPI

```
library(Rmpi)

# ...
```

3 FAQ

4 Profiling Examples with fpmpi

5 References

Chen WC, Ostrouchov G, Schmidt D, Patel P, Yu H (2012a). "pbdMPI: Programming with Big Data – Interface to MPI." R Package, URL http://cran.r-project.org/package=pbdMPI.

Chen WC, Ostrouchov G, Schmidt D, Patel P, Yu H (2012b). *A Quick Guide for the pbdMPI package*. R Vignette, URL http://cran.r-project.org/package=pbdMPI.

Sehrawat G, Chen W-C Schmidt D, Patel P, Ostrouchov G (2013). "pbdPROF: Programming with Big Data – MPI Profiling Tools." R Package, URL http://cran.r-project.org/package=pbdPROF.

Yu H (2010). "Rmpi: Interface (Wrapper) to MPI (Message-Passing Interface))." R Package (v:0.5-9), URL http://cran.r-project.org/package=Rmpi.