

The sfCluster/snowfall system: convenient parallel computing in R

Jochen Knaus

Institute of Medical Biometry and Medical Informatics, University of Freiburg

DFG Forschergruppe FOR 534

jo@imbi.uni-freiburg.de

December 12, 2008



Situation / Intention

- We wanted a solution for a heterogeneous infrastructure with many users with different knowledge levels running parallel R programs at the same time.
- Although there are many working cluster solutions for R, most of them need manual setup or are running in “traditional” batch mode.
- Especially cluster setup and handling can be too difficult for users and therefore a barrier to get them into parallel computing.
- We did not want to develop an “industrial strength” solution, our target audience are small to medium sized institutions.

Our solution: snowfall and sfCluster

snowfall

R package based on snow.

Can be used without sfCluster, but benefits from sfCluster environment.

sfCluster

Unix tool for automatic cluster management and monitoring.

snowfall R package

Design goals

- All snow supported cluster types: **MPI, PVM, Socket, NWS**.
- Connector to **sfCluster**.
- Easy access.
- Fully supporting **sequential execution** without any code changes (all wrappers work in sequential mode, too) – also enable development/debugging on Windows laptops.
- Directly **runnable everywhere** (even without snow): programs are distributable inside packages.
- Extended **error checks** and implicit cluster handler.
- Wrappers to mostly all **snow** functions. snow functions are also callable from snowfall.
- Function API equivalent to snow – porting is easy.

snowfall runtime configurations

Exchange parameters with snow

- Mostly all cluster parameters can be set using the command line. No code changes are needed to change e.g. the amount of wanted processors:

```
R --no-save --args --cpus=4 --type=PVM < myRPrg.R
```

```
R --no-save --args --hosts=localhost:3,other.xy.de:2 \  
--type=NWS < myRPrg.R
```

Through this feature, **snowfall** can be used as an R frontend for other cluster management tools, as well.

snowfall R package (2)

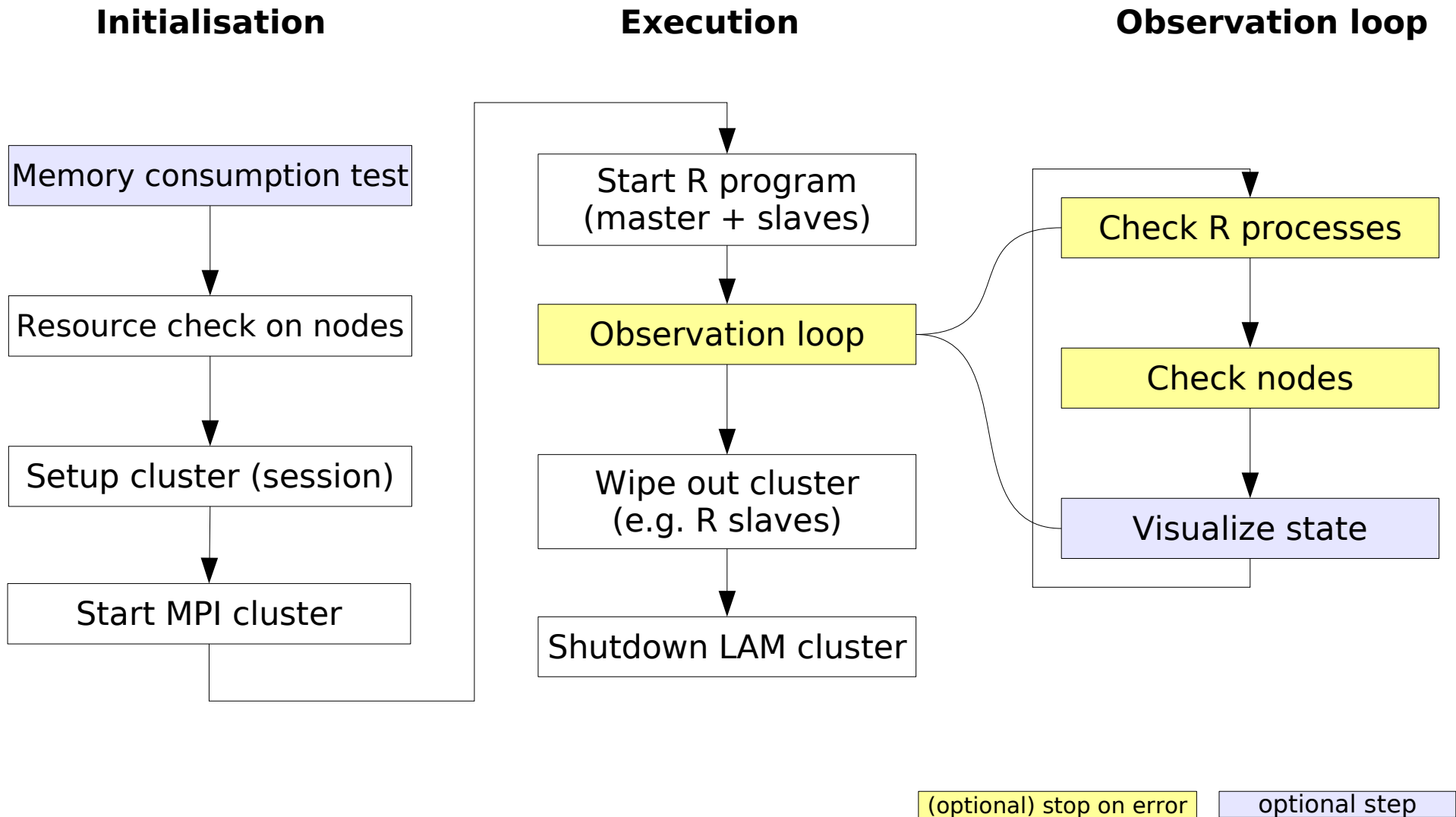
Simpler functions for common tasks

- **Loading libraries** and sources in the cluster.
- Extended variable handling over the cluster (with exporting and removal).
- Additional: parallel call with **intermediate result save and restore** (results are not lost on single node shutdowns/crashes) – this can also be used for “dynamical” cluster resizing.

sfCluster management tool

- Hide cluster handling, setup and shutdown from user.
- Implementation as Unix command line tool (written in Perl).
- Using only open source tools.
- Build upon MPI (currently LAM, OpenMPI in the future).
- Automatic resource allocation, depending on current usage of universe. Partly usage of machines is possible.
- One LAM cluster per program (means: multiple clusters per user): clusters are independent.
- Monitoring the execution of parallel R programs with detection of problems.

sfCluster workflow



sfCluster execution modes

Execution modes for running sfCluster

- **batch** (-b) like “R CMD BATCH”. *Default*
- **interactive** (-i) interactive R shell
- **monitor** (-m) batch + visual debugging

- **sequential** (-s): sequential execution without cluster

Optionally, these modes can be installed as R additions:

```
R CMD PARBATCH
R CMD PARINT
R CMD PARMON
R CMD SEQ
```

Example interactive mode

```
jo@biom9:~$ sfCluster -i --cpus=16 --mem=200
Session-ID   : bjrrj9v2_R
biom8.imbi.uni-freiburg.de: 1 CPUs assigned (1 possible).
biom9.imbi.uni-freiburg.de: 1 CPUs assigned (1 possible).
biom10.imbi.uni-freiburg.de: 1 CPUs assigned (1 possible).
knecht5.fdm.uni-freiburg.de: 8 CPUs assigned (8 possible).
knecht4.fdm.uni-freiburg.de: 5 CPUs assigned (8 possible).
ASSIGNED 16 cpus on 5 machines (16 requested).

-- sfCluster: START R-interactive session --

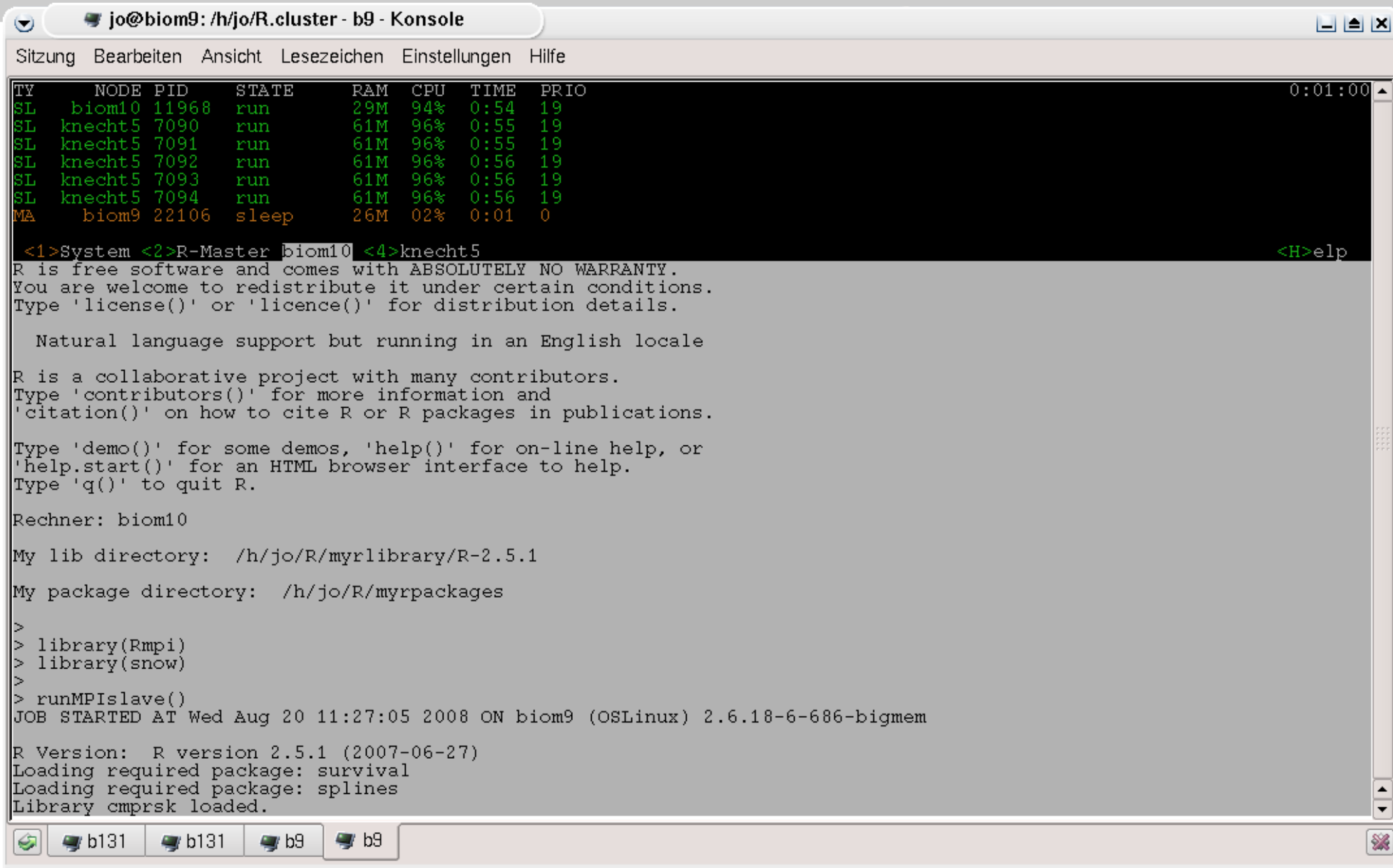
> library(snowfall)
> sfInit()
    16 slaves are spawned successfully. 0 failed.
Startup Lockfile removed: /h/jo/.sfCluster/SFINIT_jo_bjrrj9v2_R_1113_080820
JOB STARTED AT Wed Aug 20 11:14:08 2008 ON biom9 (OSLinux) 2.6.18-6-686-bigmem

R Version:   R version 2.5.1 (2007-06-27)
snowfall 1.43 initialized (parallel=TRUE, CPUs=16)

> q()
Save workspace image? [y/n/c]: n

-- sfCluster: INTERACTIVE session finished. --
LAM/MPI cluster successfully halted
```

Example screenshot of the monitoring mode



The screenshot shows a terminal window titled "jo@biom9: /h/jo/R.cluster - b9 - Konsole". The window contains a system monitoring table and an R console session.

TY	NODE	PID	STATE	RAM	CPU	TIME	PRIO
SL	biom10	11968	run	29M	94%	0:54	19
SL	knecht5	7090	run	61M	96%	0:55	19
SL	knecht5	7091	run	61M	96%	0:55	19
SL	knecht5	7092	run	61M	96%	0:56	19
SL	knecht5	7093	run	61M	96%	0:56	19
SL	knecht5	7094	run	61M	96%	0:56	19
MA	biom9	22106	sleep	26M	02%	0:01	0

<1>System <2>R-Master biom10 <4>knecht5 <H>elp

R is free software and comes with ABSOLUTELY NO WARRANTY.
You are welcome to redistribute it under certain conditions.
Type 'license()' or 'licence()' for distribution details.

Natural language support but running in an English locale

R is a collaborative project with many contributors.
Type 'contributors()' for more information and
'citation()' on how to cite R or R packages in publications.

Type 'demo()' for some demos, 'help()' for on-line help, or
'help.start()' for an HTML browser interface to help.
Type 'q()' to quit R.

Rechner: biom10

My lib directory: /h/jo/R/myrlibrary/R-2.5.1

My package directory: /h/jo/R/myrpackages

>
> library(Rmpi)
> library(snow)
>
> runMPIslave()
JOB STARTED AT Wed Aug 20 11:27:05 2008 ON biom9 (OSLinux) 2.6.18-6-686-bigmem

R Version: R version 2.5.1 (2007-06-27)
Loading required package: survival
Loading required package: splines
Library cmprsk loaded.

sfCluster execution options

- Request specific number of CPUs.
- Request specific R version for execution.
- Send mail at success or failure (one or many addresses).
- Set nice level of all slaves.
- Restore results from `sfClusterApplySR ...`
- *... and many more* (just call: `sfCluster --help`)

Examples using sfCluster

Run R programs (example: file.R):

Sequential:

```
$ R file.R
```

```
$ sfCluster -s file.R
```

Interactive mode (parallel execution):

```
$ sfCluster -i --cpus=10
```

Monitoring mode

```
$ sfCluster -m --cpus=20 --memory=1.5G --rversion=2.7.2 file.R
```

Batch mode

```
$ sfCluster -b --cpus=30 --nice=9 --mail file.R
```

Optionally, R commands are installable, e.g.

```
$ R-2.7.2 CMD PARMON file.R --args --cpus=20 --memory=1.5G
```

sfCluster administration options

- Show current usage of resources in cluster universe (with determination of **free resources**).
- Show current **running sessions** (per user or all users).
- Convenient **session shutdown** (kill). Can be used by (administration user) *root*.
- sfCluster allows the definition of “**subuniverses**” in the whole cluster universe, which are accessible to specific user groups.
- Installation via **tarball**, specific options for autoconf (e.g. R CMD additions). Debian packages on the way.

Example of administration

```
jo@biom9:~$ sfCluster -o --all
```

SESSION	STATE	USR	M	MASTER	#N	RUNTIME	R-FILE / R-OUT
MWhCBAj6_R	run	jo	M0	biom9.imbi	6	0:00:09	boot.R / boot.Rout
4DTqQJWF_R-2.7.1	run	arthur	BA	biom9.imbi	20	1:24:54	simul_pcsh.R / [...]

```
jo@biom9:~$ sfCluster --universe --mem=0.5G
```

```
Assumed memuse: 512M (use '--mem' to change).
```

Node	Max-Load	CPUs	RAM	Free-Load	Free-RAM	FREE-TOTAL
biom8.imbi.uni-freiburg.de	5	8	15.9G	1	13.6G	1
biom9.imbi.uni-freiburg.de	7	8	15.9G	1	12.4G	1
biom10.imbi.uni-freiburg.de	8	8	15.9G	1	12.4G	1
biom11.imbi.uni-freiburg.de	2	4	7.9G	0	4.6G	0
knecht5.fdm.uni-freiburg.de	8	8	15.7G	8	0.7G	1
knecht4.fdm.uni-freiburg.de	8	8	15.7G	8	3.0G	6
knecht3.fdm.uni-freiburg.de	8	8	15.7G	7	4.3G	7
knecht1.fdm.uni-freiburg.de	4	4	7.8G	4	7.5G	4
biom6.imbi.uni-freiburg.de	no-sched	4	7.9G	-	-	-

```
Potential usable CPUs: 21
```

```
jo@biom9:~$ sfCluster --kill MWhCBAj6_R
```

```
Try to "smart" shutdown remote sfCluster (biom9.imbi.uni-freiburg.de, pid 15491)
```

```
Waiting for sfCluster to halt: ..... succeeded. Force wipeout remains.
```

```
[...]
```

More informations on our website

Parallel computing with R using sfCluster/snowfall (IMBI Freiburg) - Iceweasel <2>

File Edit View History Bookmarks Tools Help

http://www.imbi.uni-freiburg.de/paralle/ Wikipedia (EN)

- Overview
- Clusters
- Get started
- Download
- Documentation
- Talks
- FAQ
- Packages
- News / History
- Contact

Parallel computing in R with sfCluster/snowfall

Overview

Although there are many working cluster solutions for R, all of them need to have the user to setup a cluster, connect to servers or perform any other, non-R specific, task. **sfCluster/snowfall** is a solution to run parallel R programs easier through MPI clusters, as users can concentrate on their R-code and are not forced into managing environments for parallel computing.

snowfall is an R package based on [snow](#). It features no additional *technical* abstraction layer, but enriched functionality increasing usability. It offers all snow cluster techniques, sequential mode for "run anywhere", extended error checkings and some functions for comfort. It is also the connection to sfCluster, but can be used without it as well.

sfCluster is a Unix tool for managing and observing clusters. It automatically set up clusters for the user and shutdown them after finish. If something nasty is going on during execution, sfCluster will notice and react. sfCluster is based on LAM/MPI (with a port to OpenMPI coming in the near future).

MPI, PVM, Socket, what?

There are several techniques to bind computers together for parallel computing, where snow and so snowfall can use four:

- **Socket-connection**: the easiest, where everything is managed inside R. Socket connections run over direct TCP/IP connection and so can be used on virtual any machines. If you just want to use parallelization on one computer (laptop or workstation) or on very few machines, you are fine. Biggest advantage is that you do not have to install additional software to use this kind of clusters.
- **MPI**: Message Passing Interface. Basically an definition of a networking protocoll. There are several different implementations today, where openMPI is the most common and widely used. sfCluster uses the bit more older LAM, but will feature openMPI in the future, as well. [Open MPI home](#).
- **PVM**: Parallel Virtual Machine. Most Unix distributions offers packages for PVM. [PVM home](#).
- **NetWorkSpaces/NWS** is a framework for coordinating programs written in scripting languages. It has support for parallel computing with it's Sleight mode. [NetWorkSpaces for R](#).

If you are not using Socket-clusters, you have to install and configure the chosen cluster solution first. Please consult your local administrator first. If you do not know what this is all about, you most likely are fine using Socket-clusters first.

Get started

SSH access without password

cluster techniques secure shell (SSH) connections needed. As these require the input of passwords, you should have access without password on these machines you want to use (even if it is only your local machine).

Done

Summary

Current state

- snowfall works very well as extended “usability” snow wrapper and is **widely used** now.
- From a technological view, sfCluster is not on the cutting edge. But: **it works**. And it works well.

Outlook

- snowfall will be enhanced on demand (next release: better Windows support, many fixes).
- sfCluster: support of **OpenMPI**, also: **easier installation**.

References

R packages: **snow**, **Rmpi**

Ananth Grama, Anshul Gupta, Vipin Kumar, and George Karypis. *Introduction to Parallel Computing*. Pearson Education, second edition, 2003.

G. Burns, R. Daoud, and J. Vaigl. *LAM: An Open Cluster Environment for MPI*. Technical report, 1994.
<http://www.lam-mpi.org/download/files/lam-papers.tar.gz>

A. Rossini, L. Tierney, and N. Li. *Simple parallel statistical computing in R*. *Journal of Computational and Graphical Statistics*, 16(2): 399-420, 2007.

Appendix

Typical problems and pitfalls concerning parallel computing in R

- Verfügbarkeit und Konsistenz von Paketen, Bibliotheken und Daten auf allen Maschinen des Clusters.
- Bibliothekspfade (benutzerinstallierte Pakete, verschiedene R Versionen)
- R Versionen identisch auf allen Rechnern.
- 32/64 Bit Problem: Bibliotheken mit kompiliertem Code (Stichwort: Netzlaufwerke).
- Speichernutzung potenziert sich.
- Reproduzierbarkeit von Zufallszahlen.