

Java调用R语言



Java如何实现与R的通信

1. 嵌入式模式JRI

rJava或者使用Java的exec方法在JVM中执行Rscript命令。这种方案会把R引擎加入到JVM的内存环境。优点是高效，缺点是紧耦合，扩展受限，并且如果在R脚本语言中出现致命的错误，有可能导致JVM的崩溃。

2. 远程通信模式Rserve

Rserve抽象了R语言网络接口，基于TCP/IP协议实现与多语言之间的通道。

Rserve介绍

Rserve是一个基于TCP/IP协议的，允许R语言与其他语言通信的C/S结构程序，支持C/C++,Java,PHP,Python,Ruby,Nodejs等。Rserve提供了远程连接、认证、文件传输等功能。可以设计R作为后台服务，处理统计建模、数据分析、绘图等任务。

Rserver安装

系统环境 macOS Sierra Verison 10.12.4

```
→ wangc ~ R
```

```
R version 3.4.0 (2017-04-21) -- "You Stupid Darkness"  
Copyright (C) 2017 The R Foundation for Statistical Computing  
Platform: x86_64-apple-darwin15.6.0 (64-bit)
```

```
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.

```
> install.packages('Rserve')
--- Please select a CRAN mirror for use in this session ---
trying URL 'https://mirror.lzu.edu.cn/CRAN/bin/macosx/el-
capitan/contrib/3.4/Rserve_1.7-3.tgz'
Content type 'application/octet-stream' length 337626 bytes (329
KB)
=====
downloaded 329 KB
```

The downloaded binary packages are **in**
/var/folders/83/g6vnxfjs3qs_xtqd3q6f28f00000gn/T//RtmpqzJsL
e/downloaded_packages
>

启动Rserve

```
→ wanc ~ R CMD Rserve
/Library/Frameworks/R.framework/Resources/bin/Rcmd: line 62: exec:
Rserve: not found
→ wanc ~
```

报错可以以下面的命令执行

```
→ wanc ~ R CMD
/Library/Frameworks/R.framework/Resources/library/Rserve/libs/Rserv
e

R version 3.4.0 (2017-04-21) -- "You Stupid Darkness"
Copyright (C) 2017 The R Foundation for Statistical Computing
Platform: x86_64-apple-darwin15.6.0 (64-bit)

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.

Rserv started **in** daemon mode.

#查看进程

```
→ wangc ~ ps -ef | grep Rserve
  501  3029      1   0 12:51PM ??          0:00.00
/Library/Frameworks/R.framework/Resources/library/Rserve/libs/Rserv
e
  501  3060  1031   0 12:52PM ttys000    0:00.00 grep --color=auto
--exclude-dir=.bzip --exclude-dir=CVS --exclude-dir=.git --exclude-
dir=.hg --exclude-dir=.svn Rserve
```

#查看端口

```
→ wangc ~ lsof -iTCP -sTCP:LISTEN | grep Rserve
Rserve    3029 wangc    3u  IPv4 0x3f8f0a39cc47e719      0t0  TCP
localhost:6311 (LISTEN)
```

Java远程连接Rserve

- 远程连接Rserve

刚刚启动的时候，使用的是本地模式，如果想远程连接需要增加参数 `--RS-enable-remote`

#杀掉刚才的Rserve进程

```
→ wangc ~ kill -9 3029
```

#打开远程模式重新启动

```
→ wangc ~ R CMD
/Library/Frameworks/R.framework/Resources/library/Rserve/libs/Rserv
e --RS-enable-remote
```

```
R version 3.4.0 (2017-04-21) -- "You Stupid Darkness"
Copyright (C) 2017 The R Foundation for Statistical Computing
Platform: x86_64-apple-darwin15.6.0 (64-bit)
```

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.

Rserv started in daemon mode.

#查看端口

```
→ wangc ~ lsof -iTCP -sTCP:LISTEN | grep Rserve
Rserve    4911 wangc    3u  IPv4  0x3f8f0a39cc424e21      0t0  TCP
*:6311    (LISTEN)
```

*: 6311表示不限制IP访问了。

- 新建Java Maven工程

引入REngine和RserveEngine

```
<dependencies>
  <!--
  https://mvnrepository.com/artifact/org.rosuda.REngine/REngine -->
  <dependency>
    <groupId>org.rosuda.REngine</groupId>
    <artifactId>REngine</artifactId>
    <version>2.1.0</version>
  </dependency>
  <!--
  https://mvnrepository.com/artifact/org.rosuda.REngine/Rserve -->
  <dependency>
    <groupId>org.rosuda.REngine</groupId>
    <artifactId>Rserve</artifactId>
```

```
<version>1.8.1</version>
</dependency>
</dependencies>
```

创建测试的Java类

```
public class Demo {

    public static void main(String[] args) throws Exception {
        Demo demo = new Demo();
        demo.callRserve();
        demo.callScript();
    }

    public void callRserve() throws RserveException,
    REXPMismatchException {
        RConnection c = new RConnection("127.0.0.1");
        REXP x = c.eval("R.version.string");
        System.out.println(x.asString()); // 打印变量x

        double[] arr = c.eval("rnorm(10)").asDoubles();
        for (double a : arr) { // 循环打印变量arr
            System.out.println(a + ",");
        }
    }

    public void callScript() throws RserveException,
    REXPMismatchException {
        RConnection rc = new RConnection("127.0.0.1");
        rc.eval("source(\"/Users/wangc/area.R\")");
        REXP rexp = rc.eval("area(10)");
        System.out.println("Area of 10 is " + rexp.asDouble());
    }
}
```

用于测试的area.R文件内容

```
area<-function(r){
    pi * r^2
}
```

运行结果

```
R version 3.4.0 (2017-04-21)
2.009860220520962,
0.5071132946452399,
-0.7096564671732806,
1.1366166668030737,
0.9782241007249238,
0.19057420975790718,
-0.10839852088552279,
-0.12713574382540224,
-0.10723228542731107,
-1.067523513605585,
Area of 10 is 314.1592653589793
```

Rserve高级使用：Rserve配置管理

通过配置文件，管理Rserve服务器

- Rserv.conf 配置文件
- source.R 初始化脚本文件

查看当前Rserve服务器的默认配置信息

```
→ wanc ~ R CMD
/Library/Frameworks/R.framework/Resources/library/Rserve/libs/Rserve
--RS-settings
Rserve v1.7-3

config file: /etc/Rserv.conf
working root: /tmp/Rserv
port: 6311
local socket: [none, TCP/IP used]
authorization required: no
plain text password: not allowed
passwords file: [none]
allow I/O: yes
allow remote access: no
control commands: no
interactive: yes
max.input buffer size: 262144 kB
```

- config file: 我本地无此文件/etc/Rserv.conf, 系统会默认跳过这项
- working root: R运行时工作目录 /tmp/Rserv

- port: 通信端口6311
- local socket: TCP/IP协议
- authorization: 认证未开启
- plain text password: 不允许明文密码
- passwords file: 密码文件，未指定
- allow I/O: 允许IO操作
- allow remote access: 远程访问未开启
- control commands: 命令控制未开启
- interactive: 允许通信
- max.input buffer size: 文件上传限制262mb

修改默认配置，新建文件：/etc/Rserv.conf

```
→ wangc R cat /etc/Rserv.conf

workdir /tmp/Rserv
remote enable
fileio enable
interactive yes
port 6311
maxinbuf 262144
encoding utf8
control enable
auth required
pwdfile /Users/wangc/R/RserveAuth.txt
plaintext enable
source /Users/wangc/R/Rserve/source.R
eval xx=1
```

RserveAuth.txt的内容

```
→ wangc R cat RserveAuth.txt
admin 123456
```

使用RScclient访问，认证报错。

```
> library(RScclient)
> conn<-RS.connect()
> RS.eval(conn,rnorm(10))
```

```
Error in RS.eval(conn, rnorm(10)) :  
  command failed with status code 0x41: authentication failed  
>
```

用RClient登陆

```
> conn<-RS.connect()  
> RS.login(conn,"admin","123456",authkey=RS.authkey(conn))  
[1] TRUE  
> RS.eval(conn,rnorm(5))  
[1] 0.05446922 0.81716847 -0.94353182 -0.78270845 -0.17954282  
>
```

性能测试

[Rserve与Java的跨平台通信](#) 我记得用RClient可以连多个，不过Rserve本身是单线程呢，每次也只能处理一个请求。

Rserve同时只允许一个客户端连接。因此，如果第二个线程试图连接时，它就会一直处在等待状态。

RConnectionPool针对的是多个Rserve服务器的情况

