Package 'rhdf5client'

October 16, 2018

Title Access HDF5 content from h5serv

Description Provides functionality for reading data from h5serv server
from within R.

Version 1.2.3

Suggests knitr, testthat, BiocStyle, DT, reticulate

Imports S4Vectors, httr, rjson, utils

Depends R (>= 3.5), methods, DelayedArray

License Artistic-2.0

LazyLoad yes

BiocViews infrastructure

RoxygenNote 6.1.0

Collate h5serv.R indx.R H5S_Array.R HSDS_Matrix.R

VignetteBuilder knitr

biocViews DataImport, Software

git_url https://git.bioconductor.org/packages/rhdf5client

git_branch RELEASE_3_7

git_last_commit 873d846

git_last_commit_date 2018-09-28

Date/Publication 2018-10-15

Author Samuela Pollack [cre, aut],
Shweta Gopaulakrishnan [aut],
Vincent Carey [aut]

Maintainer Samuela Pollack <spollack@jimmy.harvard.edu>

R topics documented:

dim,H5S_ArraySeed-method

                    HDF Server content is assumed transposed relative to R matrix layout

Description

HDF Server content is assumed transposed relative to R matrix layout

Usage

```
## S4 method for signature 'H5S_ArraySeed'
dim(x)
```

Arguments

x   instance of H5S_ArraySeed

Value

integer(2) vector of dimensions corresponding to R's layout, assuming 2-d data

---

dimnames,H5S_ArraySeed-method

dimnames not stored with H5S_source as of Jan 2018

Description

dimnames not stored with H5S_source as of Jan 2018

Usage

```
## S4 method for signature 'H5S_ArraySeed'
dimnames(x)
```

Arguments

x   instance of H5S_ArraySeed

Value

currently returns list(NULL, NULL) as we do not store dimnames in HDF5

---

domains   *HSDS server domains accessor*

Description

HSDS server domains accessor

Usage

```
domains(object, ...)
```

Arguments

object   H5S_source instance

...   not used

Value

a data frame with domains name

Examples

```
hsdsCon = H5S_source(URL_hsds()) # hsds server connection
setPath(hsdsCon, "/shared/bioconductor/")-> hsds
domains(hsds)
```

---

dsmeta *list information about datasets available in an H5S_source*

Description

list information about datasets available in an H5S_source

Usage

```
dsmeta(src)
```

Arguments

src   H5S_source instance

Value

data frame with one row for each group and three columns. The second column has the list of datasets in the group.

Examples

```
bigec2 = H5S_source(URL_h5serv())
dsm <- dsmeta(bigec2)
dst <- unlist(dsm[1,2])    # all dataset candidates in group 1
```

---

getDatasetAttrs *getDatasetAttrs from hsds server*

Description

getDatasetAttrs from hsds server

Usage

```
getDatasetAttrs(object)
```

Arguments

object   instance of H5S_source(updated object with path to file set)

Value

list of data obtained

Examples

```
hsdsCon = H5S_source(URL_hsds()) # hsds server
setPath(hsdsCon, "/shared/bioconductor/tenx_full.h5")-> hsds
getDatasetAttrs(hsds)
```

---

getDatasetSlice   *getDatasetSlice from hsds server*

Description

getDatasetSlice from hsds server

Usage

```
getDatasetSlice(object, dsindex = 1, selectionString, ...)
```

Arguments

`object`	instance of H5S_source(updated object with path to file set)
`dsindex`	dataset index
`selectionString`	
	character with selectionString
`...`	unused

Value

list of data obtained

Examples

```
hsdsCon = H5S_source(URL_hsds()) # hsds server
setPath(hsdsCon, "/shared/bioconductor/tenx_full.h5")-> hsds
getDatasetSlice(hsds,dsindex=1,selectionString="[1:10,1:50]")
```

---

getDatasetUUIDs   *getDatasetUUIDs from hsds server*

Description

getDatasetUUIDs from hsds server

Usage

```
getDatasetUUIDs(object)
```

Arguments

`object`	instance of H5S_source(updated object with path to file set)

Value

character of dataset uuid obtained

Examples

```
hsdsCon = H5S_source(URL_hsds()) # hsds server
setPath(hsdsCon, "/shared/bioconductor/tenx_full.h5")-> hsds
getDatasetUUIDs(hsds)
```

---

getDims   *getDims from hsds server*

Description

getDims from hsds server

Usage

```
getDims(object)
```

Arguments

object   instance of H5S_source(updated object with path to file set)

Value

numeric content of dimensions

Examples

```
hsdsCon = H5S_source(URL_hsds()) # hsds server
setPath(hsdsCon, "/shared/bioconductor/tenx_full.h5")-> hsds
getDims(hsds)
```

---

getHRDF   *getHRDF from hsds server*

Description

getHRDF from hsds server

Usage

```
getHRDF(object)
```

Arguments

object   instance of H5S_source(updated object with path to file set)

Value

DataFrame of data obtained

Examples

```
hsdsCon = H5S_source(URL_hsds()) # hsds server
setPath(hsdsCon, "/shared/bioconductor/tenx_full.h5")-> hsds
getHRDF(hsds)
```

---

getReq *list information about server content available in an H5S_source hsds instance*

Description

list information about server content available in an H5S_source hsds instance

Usage

```
getReq(src)
```

Arguments

src   H5S_source instance

Value

data frame with 5 columns for one row for each user's data

---

groups *HDF5 server data groups accessor*

Description

HDF5 server data groups accessor

Usage

```
groups(object, index, ...)
```

Arguments

object   H5S_source instance

index   numeric, if present, extracts metadata about selected group (sequential ordering of groups as returned by server) access for group information for HDF5 server

...   not used

Value

a data frame with group name and number of links for each group

Examples

```
bigec2 = H5S_source(URL_h5serv())
groups(bigec2)
```

groups,H5S_source,numeric-method
selective group metadata accessor

Description

selective group metadata accessor

Usage

```
## S4 method for signature 'H5S_source,numeric'
groups(object, index, ...)
```

Arguments

object	instance of H5S_source
index	numeric
...	unused

Value

one-row data frame with group name and number of links for the group

H5S_Array *create H5S_Array instance given url (filepath) and entity (host) name*

Description

create H5S_Array instance given url (filepath) and entity (host) name

Usage

```
H5S_Array(filepath, host)
```

Arguments

filepath	a character(1) URL to port for HDF Server
host	a character(1) name of 'host' in server

Value

an instance of DelayedArray-class

Examples

```
# The true values from yriMulti data element 'banovichSE':
# > assay(banovichSE[c(1:5,329465:329469),c(1:3,63:64)])
#                    NA18498    NA18499    NA18501  |    NA18489     NA18909
# cg00000029       0.47339629  1.2943041 -0.8084735 |  0.6708168 -0.86093022
# cg00000165       1.23640861  0.2099817 -0.2683763 |  0.4446088  0.99868231
# cg00000236      -0.22258183  1.6236857 -0.8654838 |  0.1958195 -0.06090929
# cg00000289       0.65720581  0.5527470 -1.8458295 | -0.4618782  0.34934164
# cg00000363      -0.15063083  0.7498020  0.3254333 |  0.7342878  0.12940774
# #----------------------------------------------------------------------
# ch.9.98936572R -0.07954958  0.2139431 -0.4719621 |  0.6835012  0.57758798
# ch.9.98937537R  0.04254705  1.0702770  1.7356387 | -0.1531732 -1.52889773
# ch.9.98959675F -1.59253143  0.2982456 -1.1954030 | -1.3703135  0.28974909
# ch.9.98989607R -1.80646652  0.4760022  1.4771808 |  0.9479602  0.49921375
# ch.9.991104F    0.08180195 -0.2434306  1.0281002 | -0.1653721  0.55612215
#
# compare to that delivered by H5S_Array
#
H5S_Array("http://h5s.channingremotedata.org:5000", "assays")
```

H5S_Array-class   *extension of DelayedArray for HDF Server content*

Description

extension of DelayedArray for HDF Server content

H5S_ArraySeed-class   *H5S_Array for HDF Server content*

Description

H5S_Array for HDF Server content

H5S_dataset   *extract elements from H5S_dataset*

Description

extract elements from H5S_dataset

Usage

```
## S4 method for signature 'H5S_dataset,numeric,numeric,ANY'
x[i, j, ..., drop = FALSE]
```

Arguments

x	instance of H5S_dataset
i	character string usable as select option for first matrix index in HDF5 server value API
j	character string usable as select option for second matrix index in HDF5 server value API
...	unused
drop	logical defaults to FALSE

Value

matrix of data obtained

Slots

source instance of H5S_source instance

simpleName character string naming dataset

shapes list including dimension information

hrefs DataFrame of hrefs as defined in the API

allatts list of all attributes

presel string prepared for select operation in GET

transfermode default "JSON" or "binary" for binary transfer

H5S_dataset2   *H5S_dataset2 for datasets in hsds server*

Description

H5S_dataset2 for datasets in hsds server

Usage

```
H5S_dataset2(object)
```

Arguments

object	instance of H5S_source(updated object with path to file set)

Value

H5S_dataset object

Examples

```
hsdsCon = H5S_source(URL_hsds()) # hsds server
setPath(hsdsCon, "/shared/bioconductor/tenx_full.h5")-> hsds
H5S_dataset2(hsds)
```

`H5S_Matrix-class`   *extension of DelayedMatrix for HDF Server content*

Description

extension of DelayedMatrix for HDF Server content

`H5S_source`   *H5S_source identifies an HDF5/HSDS server and manages some metadata about contents*

Description

H5S_source identifies an HDF5/HSDS server and manages some metadata about contents

construct H5S_source

Usage

```
H5S_source(serverURL, domain, ...)

## S4 method for signature 'H5S_source,character'
x[[i, j]]

dataset(h5s, tag)
```

Arguments

serverURL	a URL for a port for HDF5Server
domain	character string with path to file for HSDS
...	not used
x	instance of H5S_source
i	character string intended to identify dataset on server
j	not used
h5s	instance of H5S_source
tag	character string identifying a dataset

Value

an initialized object of type H5S_source

Slots

serverURL character string with a URL

dsmeta DataFrame instance with metadata about content of h5serv server

dmains DataFrame instance with metadata about the content of hsds server

getReq DataFrame instance with metadata about hsds server

folderPath character string with path to user's folder/file on hsds server

Note

The dsmeta slot holds a DataFrame with a column dsnames that is a list with ith element a character vector of all dsnames available for the ith group. There is no effort at present to search all groups for candidate datasets.

If the domain for the HSDS server is known, pass the domain path as a character string along with ther serverURL

Examples

```
bigec2 = H5S_source(URL_h5serv()) # h5serv
bigec2
dsmeta(bigec2)[1:2,]      # two groups
dsmeta(bigec2)[1,2][[1]]  # all dataset candidates in group 1
hsdsCon = H5S_source(URL_hsds()) # hsds server connection, note : if the domain is known, pass as character st
hsdsCon
getReq(hsdsCon)
```

---

hsdsInfo   *HSDS server get request accessor*

Description

HSDS server get request accessor

Usage

```
hsdsInfo(object)
```

Arguments

object   H5S_source instance

Value

a data frame with response

Examples

```
hsdsCon = H5S_source(URL_hsds()) # hsds server connection
hsdsInfo(hsdsCon)
```

HSDS_Matrix *simplify construction of DelayedMatrix from url and path in HSDS*

Description

simplify construction of DelayedMatrix from url and path in HSDS

Usage

```
HSDS_Matrix(url, path)
```

Arguments

url   character(1) URL for HSDS object store with port

path   character(1) path from root defining HDF Cloud resource

Value

instance of DelayedArray

Examples

```
HSDS_Matrix
```

---

internalDim *acquire internal HDF5 dimension information for matrix*

Description

acquire internal HDF5 dimension information for matrix

Usage

```
internalDim(h5d)
```

Arguments

h5d   instance of H5S_dataset

Value

vector with dimensions of dataset

Examples

```
bigec2 = H5S_source(URL_h5serv())
tex <- bigec2[["tenx_100k_sorted"]]
internalDim(tex)
```

---

isplit	*isplit converts a numeric vector into a list of sequences for compact reexpression*

Description

isplit converts a numeric vector into a list of sequences for compact reexpression

sproc makes vector of type character of triplets initial:final:stride in R-conventions

Usage

```
isplit(x)

sproc(spl)
```

Arguments

x	a numeric vector (should be integers)
spl	output of isplit

Value

list of vectors of integers which can be expressed as initial/final/stride triplets

list of colon-delimited strings each with initial/final/stride triplet

Examples

```
inds = c(1:10, seq(25,50,2), seq(200,150,-2))
sproc(isplit(inds))
```

---

links	*access for link metadata for HDF5 server groups*

Description

access for link metadata for HDF5 server groups

Usage

```
links(object, index, ...)
```

Arguments

object	H5S_source instance
index	numeric group index
...	not used

Value

an object of type H5S_linkset with the linkset of the group

Examples

```
bigec2 = H5S_source(URL_h5serv())
lks <- links(bigec2, 1)    # linkset for root group
urls <- targets(lks)       # URLs of datasets in linkset
```

---

setPath   *set path for hsds server resource*

Description

set path for hsds server resource

Usage

```
setPath(object, folderPath, ...)
```

Arguments

object	H5S_source instance
folderPath	character string with path to user's folder on hsds server
...	not used

Value

an updated object with folderPath set

Examples

```
hsdsCon = H5S_source(URL_hsds()) # hsds server connection
setPath(hsdsCon, "/shared/bioconductor/tenx_full.h5")-> hsds
```

---

targets   *provide the full URLs for link members*

Description

provide the full URLs for link members

Usage

```
targets(h5linkset, index)
```

Arguments

h5linkset	instance of H5S_linkset
index	numeric index into link vector - ignored

Value

a vector of dataset tags

Examples

```
bigec2 = H5S_source(URL_h5serv())
lks <- links(bigec2, 1)   # linkset for first group (Note: first group is the root group, by construction)
urls <- targets(lks)       # URLs of datasets in linkset
```

---

transfermode<-   *replace transfer mode*

Description

replace transfer mode

Usage

```
transfermode(object) <- value
```

Arguments

object   instance of H5S_linkset

value   either "JSON" (default) or "binary"

Value

updated object of type H5S_dataset

URL_h5serv   *manage h5serv URL*

Description

manage h5serv URL

Usage

```
URL_h5serv()
```

Examples

```
URL_h5serv()
```

URL_hsds	*manage hsds URL*

Description

manage hsds URL

Usage

```
URL_hsds()
```

Examples

```
URL_hsds()
```

[,H5S_dataset,character,character,ANY-method
extract elements from H5S_dataset

Description

extract elements from H5S_dataset

Usage

```
## S4 method for signature 'H5S_dataset,character,character,ANY'
x[i, j, ...,
  drop = FALSE]
```

Arguments

x	instance of H5S_dataset
i	character string usable as select option for first matrix index in HDF5 server value API
j	character string usable as select option for second matrix index in HDF5 server value API
...	unused
drop	logical defaults to FALSE

Value

matrix of data obtained

[,H5S_dataset,character,character-method
extract elements from H5S_dataset

Description

extract elements from H5S_dataset

Usage

```
## S4 method for signature 'H5S_dataset,character,character'
x[i, j, ..., drop = FALSE]
```

Arguments

x	instance of H5S_dataset
i	character string usable as select option for first matrix index in HDF5 server value API
j	character string usable as select option for second matrix index in HDF5 server value API
...	unused
drop	logical defaults to FALSE

Value

matrix of data obtained

Index