New Directions for HDF5 Tools

HDF Server/HDF Compass/JSON Tools

John Readey
The HDF Group
jreadey@hdfgroup.org
HDF Server – A Web API for HDF
Why an HDF5 Web API?

Motivation to create a web api:

Anywhere reference-able data – ie URI
Network Transparency
Clients can be lighter weight
Support Multiple Writer/Multiple Reader
Enable Web UIs (JavaScript access)
Increased scope for features/performance boosters

• E.g. in memory cache of recently used data
HDF Server Highlights

Written in Python using Tornado Framework (uses h5py & PyTables)
REST-based API
HTTP request/responses in JSON
Full CRUD (create/read/update/delete) support
Most HDF5 features (Compound types, Compression, chunking, links)
Self-contained web server
Open Source
UUID identifiers for Groups/Datasets/Datatypes
Very easy to install/run
Python client lib h5pyd (h5py-distributed)
A simple diagram of the REST API...
What makes it RESTful?

Client-server model

Stateless – (no client context stored on server)

Cacheable – clients can cache responses

Resources identified by URIs (datasets, groups, attributes, etc)

Standard HTTP methods and behaviors (opposed to rpc-style):

<table>
<thead>
<tr>
<th>Method</th>
<th>Safe</th>
<th>Idempotent</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>GET</td>
<td>Y</td>
<td>Y</td>
<td>Get a description of a resource</td>
</tr>
<tr>
<td>POST</td>
<td>N</td>
<td>N</td>
<td>Create a new resource</td>
</tr>
<tr>
<td>PUT</td>
<td>N</td>
<td>Y</td>
<td>Create a new named resource</td>
</tr>
<tr>
<td>DELETE</td>
<td>N</td>
<td>Y</td>
<td>Delete a resource</td>
</tr>
</tbody>
</table>
Example URI

https://data.hdfgroup.org:7258/datasets/feef70e8-16a6-11e5-994e-06fc179af5e?host=tall.data.hdfgroup.orgvalue?select=[0:4,0:4]

- Scheme: the connection protocol: “https”
- Domain: endpoint of the server: “data.hdfgroup.org”
- Port: the port the server is running on: 7258
- Resource: identifier for the resource (UUID of dataset)
- Query param: Modify how the data will be returned
  - (e.g. hyperslab selection)

Note: no run time context!
What’s next: – Web UI

http://hdlgroup.org

Object Window

(Content changes with data type)

Display Window

Parameter 1
Parameter 2
Parameter 3
...
What’s next – access control

Security! Needed if the server is to be run on an open network

Authentication (you are who you say you are)
HTTPS (cut out the man in the middle)
Authorization (who can do what)
  • Per resource ACL’s

<table>
<thead>
<tr>
<th>UserID</th>
<th>Create</th>
<th>Read</th>
<th>Update</th>
<th>Delete</th>
<th>readACL</th>
<th>modACL</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>False</td>
<td>True</td>
<td>False</td>
<td>False</td>
<td>False</td>
<td>False</td>
</tr>
<tr>
<td>456</td>
<td>False</td>
<td>True</td>
<td>True</td>
<td>False</td>
<td>False</td>
<td>False</td>
</tr>
<tr>
<td>123</td>
<td>True</td>
<td>True</td>
<td>True</td>
<td>True</td>
<td>True</td>
<td>True</td>
</tr>
</tbody>
</table>

User 123

User 456

User X
What’s next: Scalable Server

- Support any sized repository
- Any number of users
- Any request volume
- Provide data as fast as the client can pull it in
- Targeted for public/private cloud
Try it out!

- Source Code: https://github.com/HDFGroup/h5serv
- AWS end point: https://data.hdfgroup.org:7258/
- Project Page: https://www.hdfgroup.org/projects/hdfserver/
Demo

HDF Server
HDF Compass
Simple HDF Viewer for Mac/Windows/Linux
HDF Compass – Fact Sheet

Read-only by Design
Native UI
Support largest HDF Files
Easy to develop plugin architecture
Written in Python
HDF Compass Architecture

```
HDF Compass

```

```
Plugin layer

“Frozen” Python app

```

```
Other formats

```

```
OPeNDAP

```

```
HDF

```
Try it out!

https://www.hdfgroup.org/projects/compass/
HDF5-JSON

Specification and Tools to represent HDF5 in JSON
HDF5/JSON Factsheet

“Full fidelity” representation of HDF5

Common line tools:

• H5tojson: convert HDF5 to JSON
• Jsontoh5: convert JSON to HDF5
• Jsonetcode: code generator for Python/Fortran/IDL/Matlab

Available on github: [https://github.com/HDFGroup/hdf5-json](https://github.com/HDFGroup/hdf5-json)

HDF5/JSON - Demo

Sample JSON file:

https://github.com/HDFGroup/hdf5-json/blob/master/data/json/tall.json
Questions?