Heliophysics Data API (HAPI)

Bob Weigel, Jon Vandegriff, Jeremy Faden, D. Aaron Roberts, Todd King, Nand Lal, Bobby Candey, Bernie Harris, Larry Brown

Overview

- HAPI is a HTTP API (Example) specification designed primarily for streaming time series data, from simple scalars to N-dimensional spectrograms.
- Main entry point http://hapi-server/
- ~7-min overview of project then demo of Python tools and capabilities

Overview

- In development over past 1.5 years. Specification is mature and development of clients and servers is ongoing.
- The specification was developed by software engineers who have developed similar services and scientists who use and/or have developed services.

Motivation for Presentation

- Project is mature enough that we need to start encouraging usage
- Lots of buy-in from server/software developers
- Need feedback and buy-in from users
- Want to hear CEDAR community perspective

Motivation for Specification

In the Heliophysics community the methods for how data providers expose data include:

A. A FTP or HTTP directory of files (usually one day of data per file);
B. A HTTP request that returns a web page with link to a file when processing is complete;
C. A HTTP request that returns a web page with link to an archive of files (zip or tgz) when processing complete; and
D. An API that returns a data stream.

Also note variations in implementation in each category, e.g., for A., providers may have different directory structures, file types, and file naming conventions.

Motivation for Specification

The data providers and the methods available include

1. CDAWeb [https://cdaweb.gsfc.nasa.gov/] - A, B, D
2. SSCWeb [https://sscweb.gsfc.nasa.gov/] - D
3. Das2 [http://das2.org/] - D
4. LiSIRD [http://lasp.colorado.edu/lisird/] - A, D
5. OMNIWeb [https://omniweb.gsfc.nasa.gov/] - A, D
6. SuperMAG [http://supermag.jhuapl.edu/] - D
7. INTERMAGNET [http://intermagnet.org] - A, C
8. CARISMA [http://www.carisma.ca] - C
9. IMAGE [http://space.fmi.fi/image] - C

A single specification could be used to describe and serve all data

Primary design considerations

1. Should be simple to write a basic HAPI server and client; and
2. Metadata should be just enough to create a plot with sensible scientific labels. Richer science-interpretation-level metadata (e.g., SPASE or provider web page) is pointed to.

Facilitating Adoption

To facilitate adoption, in parallel to the development of the specification, development of

1. Clients for Java, Python, MATLAB, IDL
2. A server validator
3. A general-use server

Endpoints

http://server/hapi/capabilities

http://server/hapi/catalog

http://server/hapi/info

http://server/hapi/data

API - Metadata

http://server/hapi/catalog - Returns a list of available datasets

https://cdaweb.gsfc.nasa.gov/hapi/catalog


```
{
    HAPI: "2.0",
  - status: {
        code: 1200,
        message: "OK"
    },
  - catalog: [
      - {
            id: "A1_K0_MPA"
        },
      - {
            id: "A2_K0_MPA"
        },
      - {
            id: "AC_AT_DEF"
```

API - Metadata

http://server/hapi/capabilities

Returns a list of file formats supported (CSV, Binary, and JSON). A HAPI server only needs to support CSV.

```
{
    HAPI: "2.0",
  - status: {
        code: 1200,
        message: "OK"
    },
  - outputFormats: [
        "csv",
        "binary",
        "json"
    ]
}
```

API - Metadata

http://server/hapi/info?**id=DATASET**

https://cdaweb.gsfc.nasa.gov/hapi/info?id=AC_H0_MFI


```
{
    HAPI: "2.0",
  - status: {
        code: 1200,
        message: "OK"
    },
  - parameters: [
    - {
          name: "Time",
          type: "isotime",
          units: "UTC",
          length: 24,
          fill: null
      },
    - {
          name: "Magnitude",
          type: "double",
          units: "nT",
          fill: "-1.0E31",
          description: "B-field magnitude"
      },
    - {
          name: "BGSEc",
          type: "double",
          units: "nT",
          fill: "-1.0E31",
          description: "Magnetic Field Vector in GSE Cartesian coordinates (16 sec)",
      - size: [
            3
        ]
```

API - Metadata

http://server/hapi/info?**id**=DATASET&parameters=P1,P2,...

Returns info for only requested parameters

https://cdaweb.gsfc.nasa.gov/hapi/info?id=AC_H0_MFI&parameters=Magnitude,BGSEc


```
{
    HAPI: "2.0",
  - status: {
        code: 1200,
        message: "OK"
    },
  - parameters: [
      - {
            name: "Time",
            type: "isotime",
            units: "UTC",
            length: 24,
            fill: null
        },
      - {
            name: "Magnitude",
            type: "double",
            units: "nT",
            fill: "-1.0E31",
            description: "B-field magnitude"
        },
      - {
            name: "BGSEc",
            type: "double",
            units: "nT",
            fill: "-1.0E31",
            description: "Magnetic Field Vector in GSE Cartesian coordinates (16 sec)",
          - size: [
                3
            ]
        }
    ],
    startDate: "1997-09-02T00:00:12Z",
    stopDate: "2018-08-10T23:59:51Z"
}
```

API - Data

hapi/data

?id=DATASET	
&parameters=P1,P2,...	*If not given, all parameters served*
&time.min=ISO8601c	*Constrained ISO8601 timestamp*
&time.max=ISO8601c	*Constrained ISO8601 timestamp*
[&format=csv, json, binary]	*Server only needs to support format=csv.*

Default result is a CSV table for parameters P1, P2, ... (parameters can be multidimensional - client uses metadata to reshape associated columns).

```
2001-01-01T00:01:33.00Z,1.1,2.1
2001-01-01T00:01:34.00Z,1.2,2.2
...
```

Development

Specification is at version 2.1 and is stable.

- Considering additions for case where frequency channels/bins change with time.
- Development of search interface.
- Need to consider using a standard for units.
- Need to consider caching specification.
- Interested in hearing about data used by CEDAR community

Known servers: http://hapi-server.org/servers

Python client demo

Python client is available via

```
pip install hapiclient --upgrade
```

See also https://github.com/hapi-server/client-python

Demo of usage: https://github.com/hapi-server/client-python-notebook