Heliophysics Data API (HAPI)

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

Overview

- HAPI is a HTTP API specification designed primarily for streaming time series data, from simple scalars to N-dimensional spectrograms with time-varying bins.
- Main entry point https://hapi-server/

Overview

- Active development since 2017. Specification is mature as are clients and servers and other software tools.
- The specification was developed by software engineers who have developed similar services and scientists who use and/or have developed services.

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

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
C. A HTTP request that returns a web page with link to zip or tgz
D. An API that returns a data stream.

Some data providers and the methods available include

1. **AMDA** [http://amda.irap.omp.eu/service/hapi] - B, D
2. CARISMA [http://www.carisma.ca] - C
3. **CCMC/iSWA** [https://iswa.gsfc.nasa.gov/] - D
4. **CDAWeb** [https://cdaweb.gsfc.nasa.gov/] - A, B, D
5. *CAIO* [https://csa.esac.esa.int/csa/aio/] - C, D
6. **Das2** [http://das2.org/] - D
7. IMAGE [http://space.fmi.fi/image] - C
8. **INTERMAGNET** [http://intermagnet.org] - A, C
9. **LiSIRD** [http://lasp.colorado.edu/lisird/] - A, D
10. *MADRIGAL* [https://openmadrigal.org/] - A, D
11. **OMNIWeb** [https://omniweb.gsfc.nasa.gov/] - A, D
12. *PDS* [https://pds-ppi.igpp.ucla.edu/hapi]
13. **SSCWeb** [https://sscweb.gsfc.nasa.gov/] - D
14. *SuperMAG* [http://supermag.jhuapl.edu/] - D
15. **VirES-for-Swarm** [https://vires.services/hapi] - D

Bold = HAPI available
Italic = HAPI under development

A single specification for method (streaming API) can be used to describe and serve data from all of these providers

Primary design considerations

1. Should be simple to write a basic HAPI server and client
2. Metadata should be just enough to create a plot with sensible scientific labels.

   More detailed science-interpretation-level metadata (e.g., SPASE or provider web page) is pointed to in HAPI metadata.

Facilitating Adoption

To facilitate adoption, in parallel to the development of the specification, we have developed

1. Clients for Autoplot, IDL, Java, Javascript, MATLAB, Python, R;
2. A web interface for selecting data from any HAPI server + plotting + script builder (https://hapi-server.org/servers/;
3. A server validator; and
4. A general-use servers (In Java and Javascript; only need to provide metadata + command line program that returns data); less generalized code for Python server exists.

Spec and full list of software at https://github.com/hapi-server

Development

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

Specification is at version 3.1 and is stable. For 3.1:

- Additions for case where frequency channels/bins change with time
- Allow for key that indicates the standard used for unit strings
- Allow for indicating coordinate system when appropriate
- Allow for embedding of external metadata

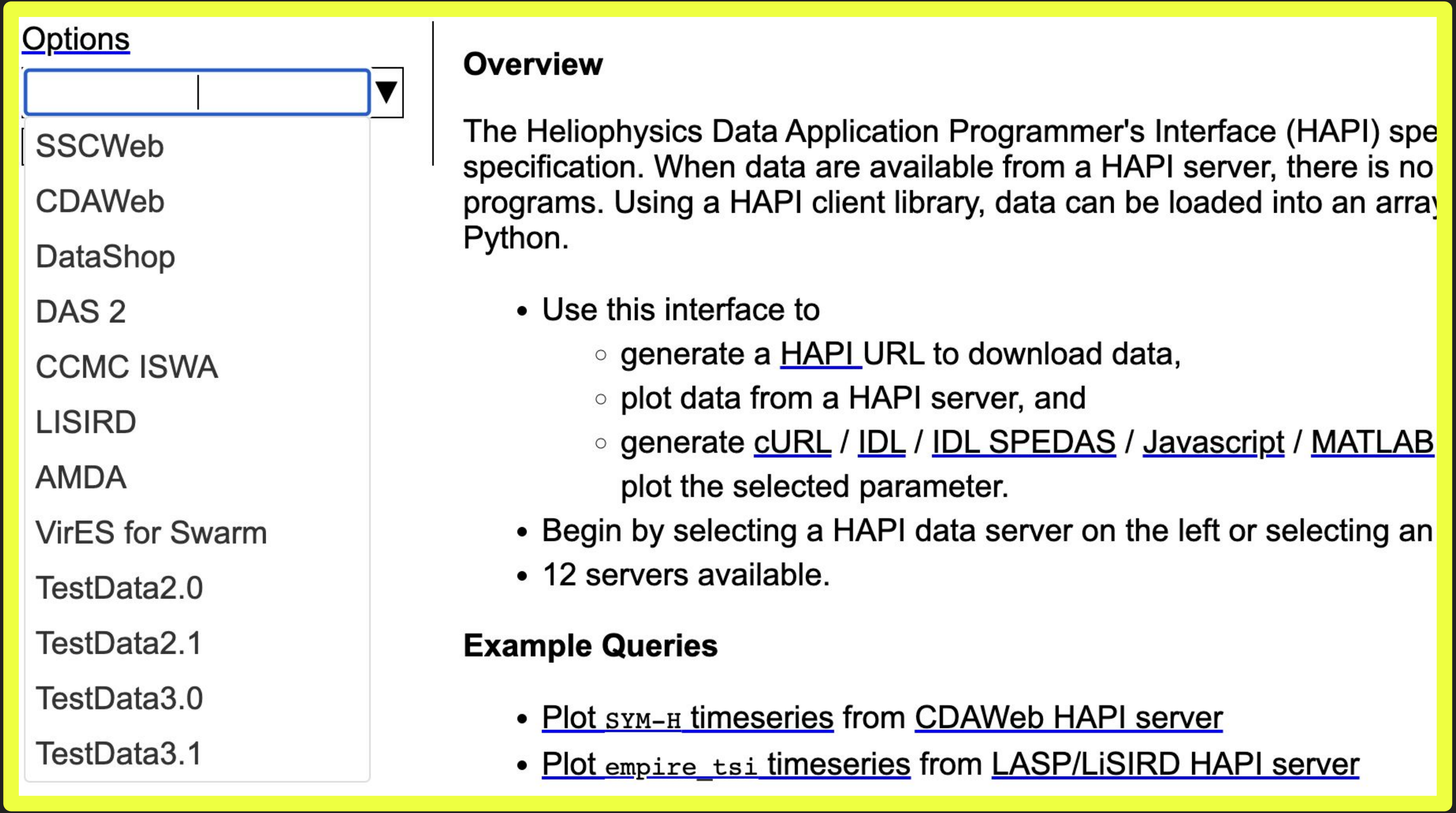
- Several groups are developing servers for simulation output; ongoing discussions on how to handle custom request parameters
- Active development of general servers

Summary

- HAPI is a HTTP API specification designed primarily for streaming time series data, from simple scalars to N-dimensional spectrograms with time-varying bins.
- Data from many providers is available through HAPI software clients, and we continue to add to the list.

Back-up slides

Web Interface to all Servers



Overview

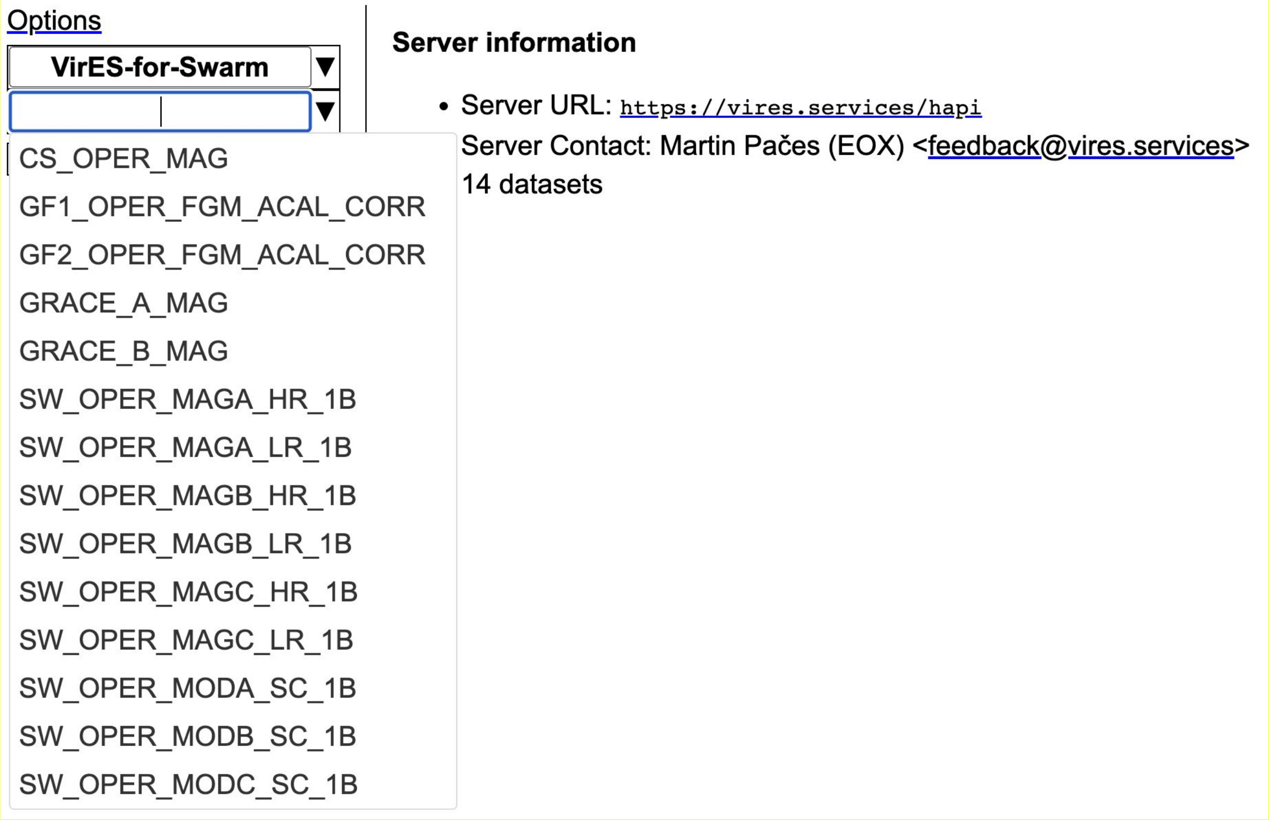
The Heliophysics Data Application Programmer's Interface (HAPI) spe
specification. When data are available from a HAPI server, there is no
programs. Using a HAPI client library, data can be loaded into an array
Python.

- Use this interface to
  - generate a HAPI URL to download data,
  - plot data from a HAPI server, and
  - generate cURL / IDL / IDL SPEDAS / Javascript / MATLAB
    plot the selected parameter.
- Begin by selecting a HAPI data server on the left or selecting an
- 12 servers available.

Example Queries

- Plot `SYM-H` timeseries from CDAWeb HAPI server
- Plot `empire_tsi` timeseries from LASP/LiSIRD HAPI server

Web Interface to all Servers



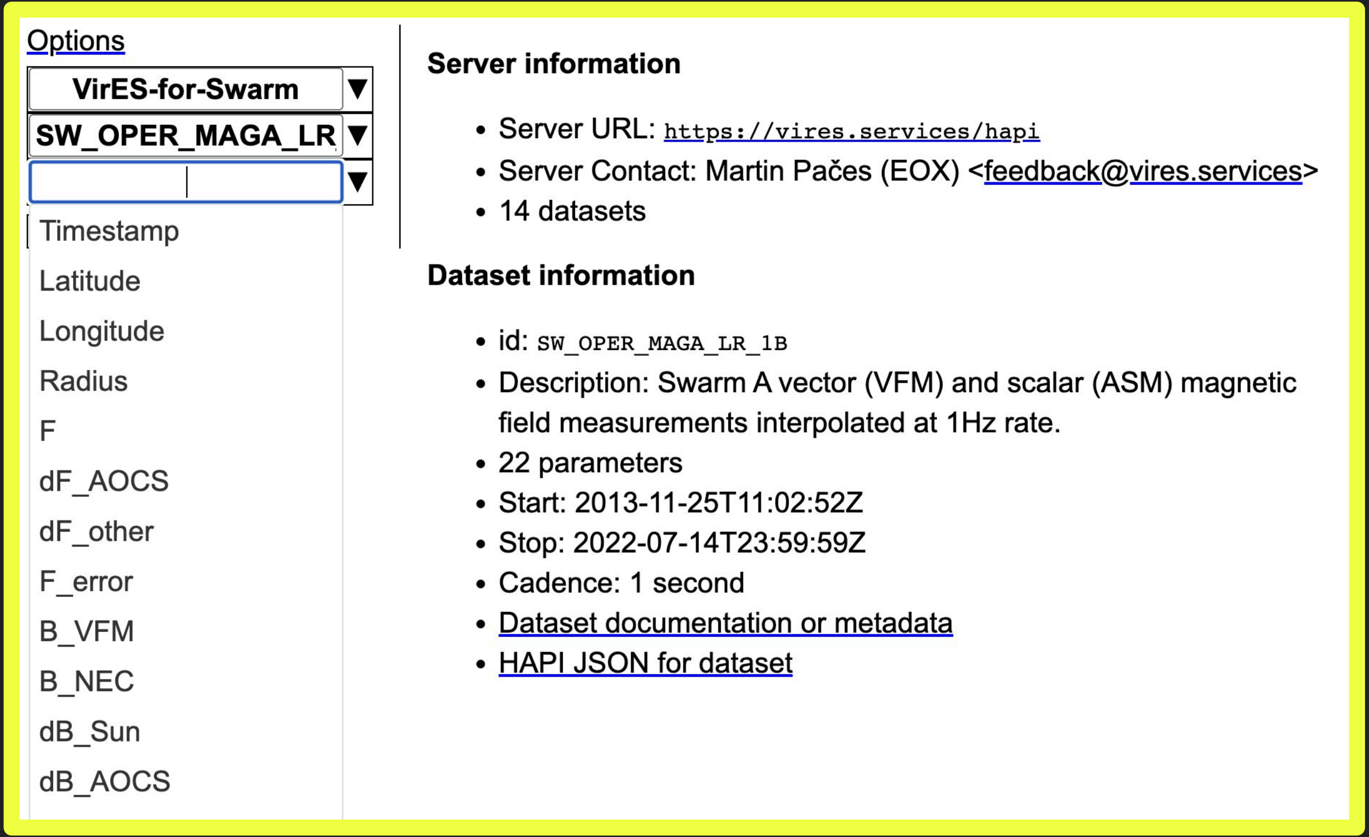
Server information

- Server URL: `https://vires.services/hapi`

Server Contact: Martin Pačes (EOX) <feedback@vires.services>

14 datasets

Web Interface to all Servers



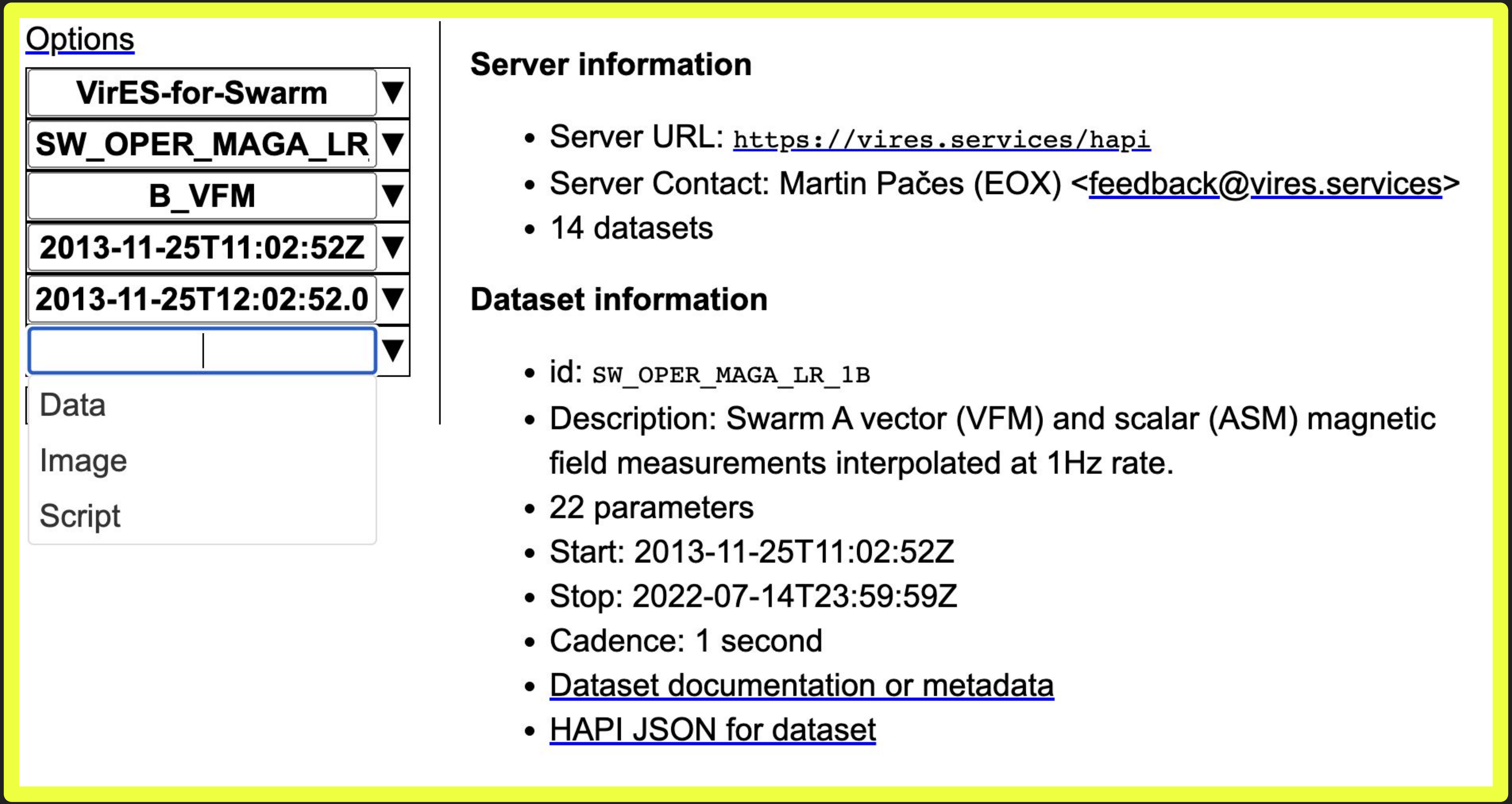
Server information

- Server URL: `https://vires.services/hapi`
- Server Contact: Martin Pačes (EOX) <feedback@vires.services>
- 14 datasets

Dataset information

- id: `SW_OPER_MAGA_LR_1B`
- Description: Swarm A vector (VFM) and scalar (ASM) magnetic field measurements interpolated at 1Hz rate.
- 22 parameters
- Start: 2013-11-25T11:02:52Z
- Stop: 2022-07-14T23:59:59Z
- Cadence: 1 second
- Dataset documentation or metadata
- HAPI JSON for dataset

Web Interface to all Servers



Options

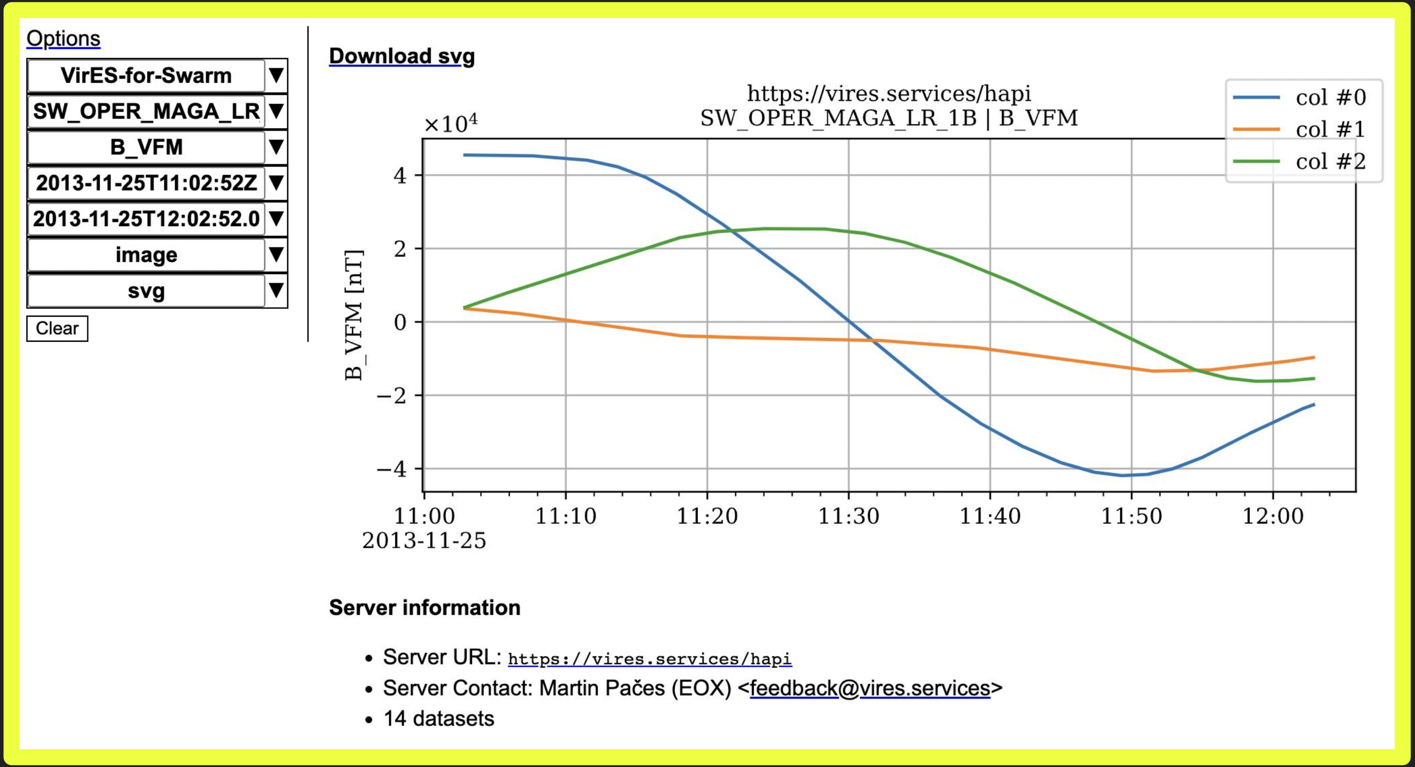
Server information

- Server URL: `https://vires.services/hapi`
- Server Contact: Martin Pačes (EOX) <feedback@vires.services>
- 14 datasets

Dataset information

- id: `SW_OPER_MAGA_LR_1B`
- Description: Swarm A vector (VFM) and scalar (ASM) magnetic field measurements interpolated at 1Hz rate.
- 22 parameters
- Start: 2013-11-25T11:02:52Z
- Stop: 2022-07-14T23:59:59Z
- Cadence: 1 second
- Dataset documentation or metadata
- HAPI JSON for dataset

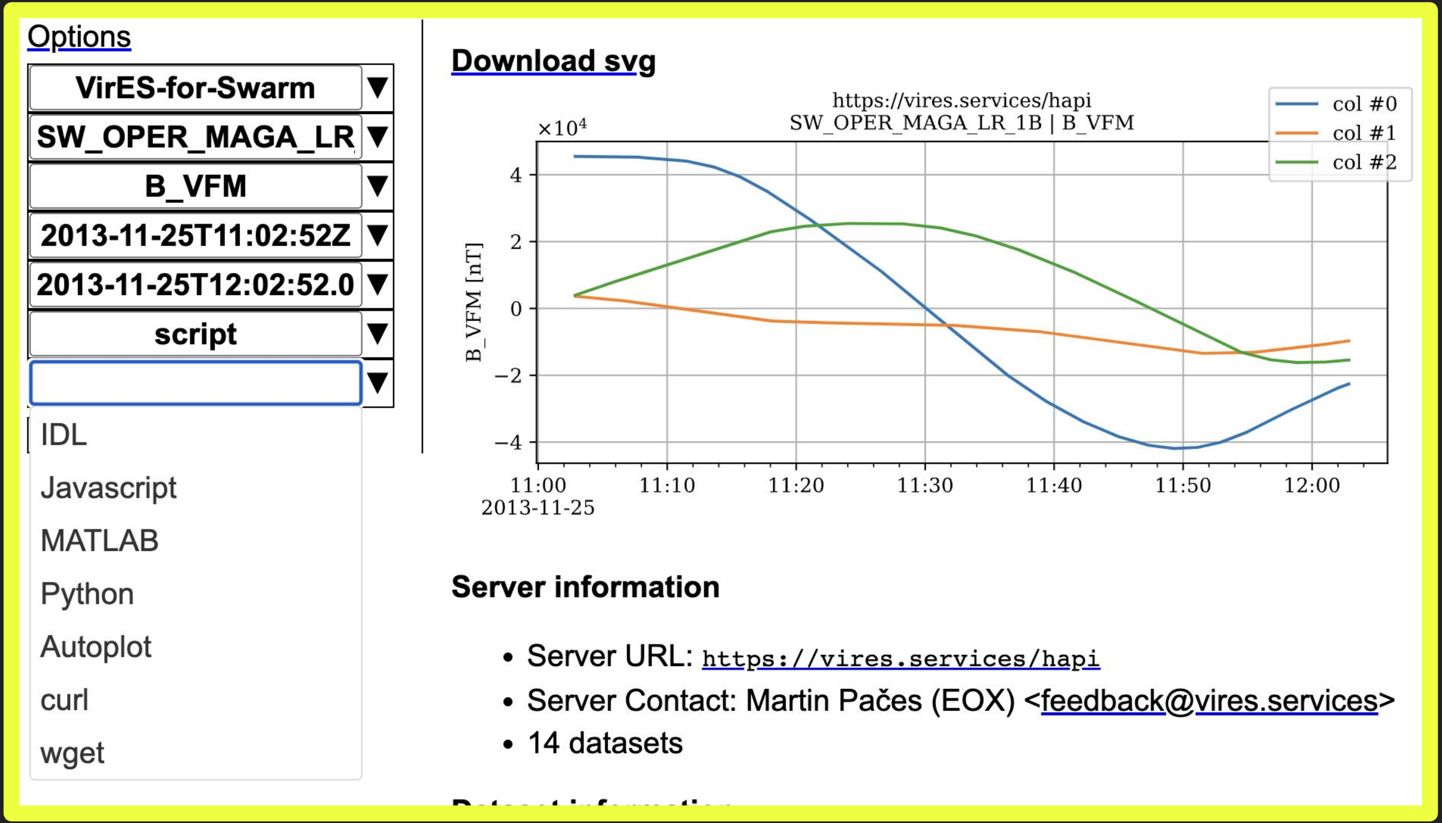
Web Interface to all Servers



Server information

- Server URL: `https://vires.services/hapi`
- Server Contact: Martin Pačes (EOX) <feedback@vires.services>
- 14 datasets

Web Interface to all Servers



Options

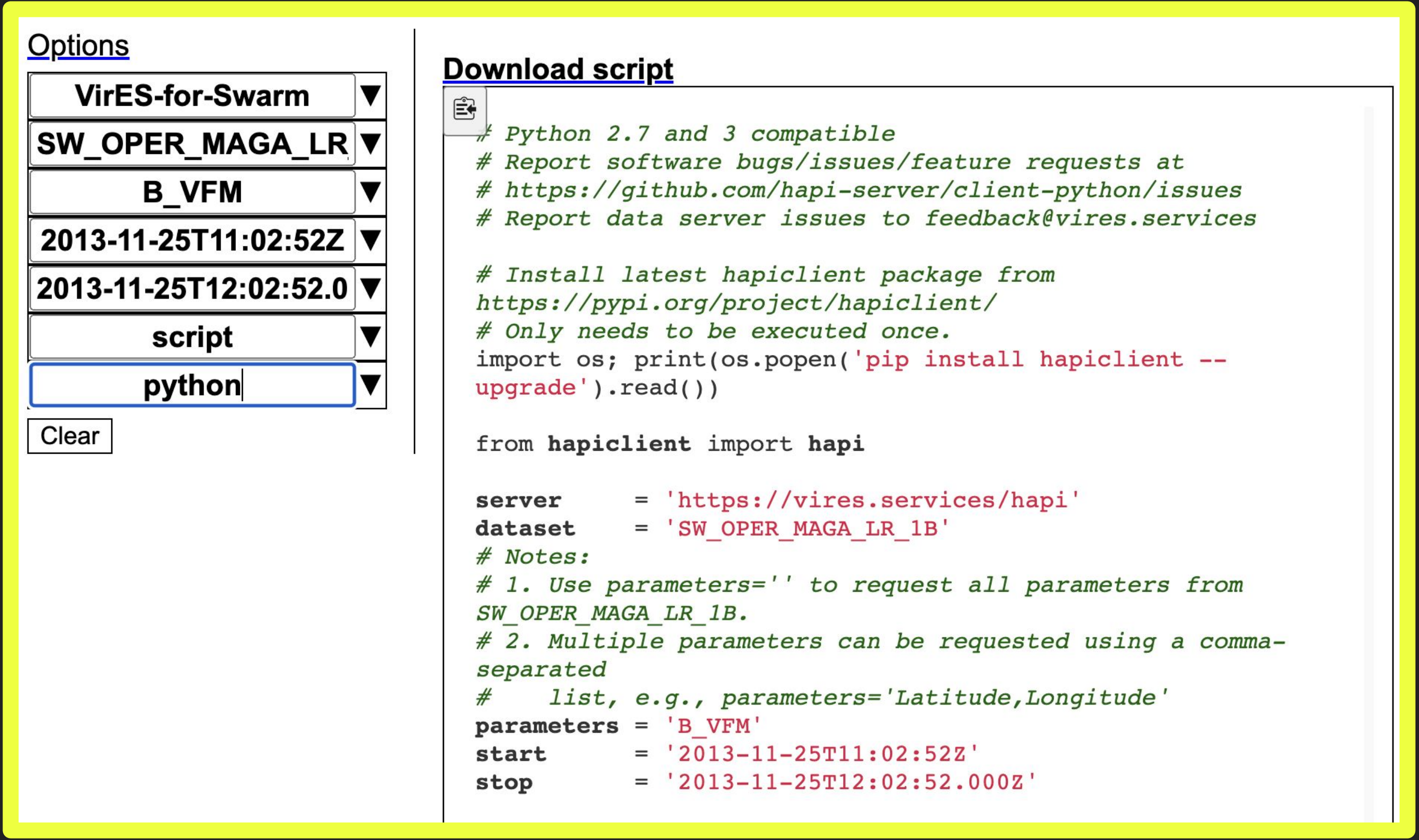
- VirES-for-Swarm ▼
- SW_OPER_MAGA_LR ▼
- B_VFM ▼
- 2013-11-25T11:02:52Z ▼
- 2013-11-25T12:02:52.0 ▼
- script ▼
- ▼

Server information

- Server URL: `https://vires.services/hapi`
- Server Contact: Martin Pačes (EOX) <feedback@vires.services>
- 14 datasets

Dataset information

Web Interface to all Servers



Options

VirES-for-Swarm ▼

SW_OPER_MAGA_LR ▼

B_VFM ▼

2013-11-25T11:02:52Z ▼

2013-11-25T12:02:52.0 ▼

script ▼

python ▼

Clear

Download script

```
# Python 2.7 and 3 compatible
# Report software bugs/issues/feature requests at
# https://github.com/hapi-server/client-python/issues
# Report data server issues to feedback@vires.services

# Install latest hapiclient package from
https://pypi.org/project/hapiclient/
# Only needs to be executed once.
import os; print(os.popen('pip install hapiclient --
upgrade').read())

from hapiclient import hapi

server     = 'https://vires.services/hapi'
dataset    = 'SW_OPER_MAGA_LR_1B'
# Notes:
# 1. Use parameters='' to request all parameters from
SW_OPER_MAGA_LR_1B.
# 2. Multiple parameters can be requested using a comma-
separated
#    list, e.g., parameters='Latitude,Longitude'
parameters = 'B_VFM'
start      = '2013-11-25T11:02:52Z'
stop       = '2013-11-25T12:02:52.000Z'
```

API - Endpoints

http://server/hapi/capabilities

http://server/hapi/catalog

http://server/hapi/info

http://server/hapi/data

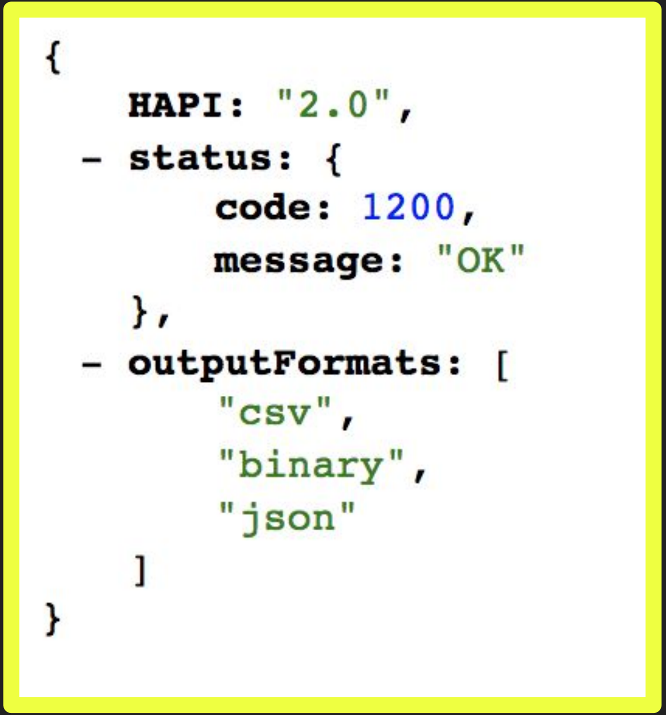
API - Metadata

http://server/hapi/capabilities

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

Note that the data output is not meant to be a file format for exchange of data. It is a simple streaming format intended for use primarily by HAPI clients.

The optional binary format is simple and fast - essentially CSV with commas and newlines removed and numbers replaced with binary representation.

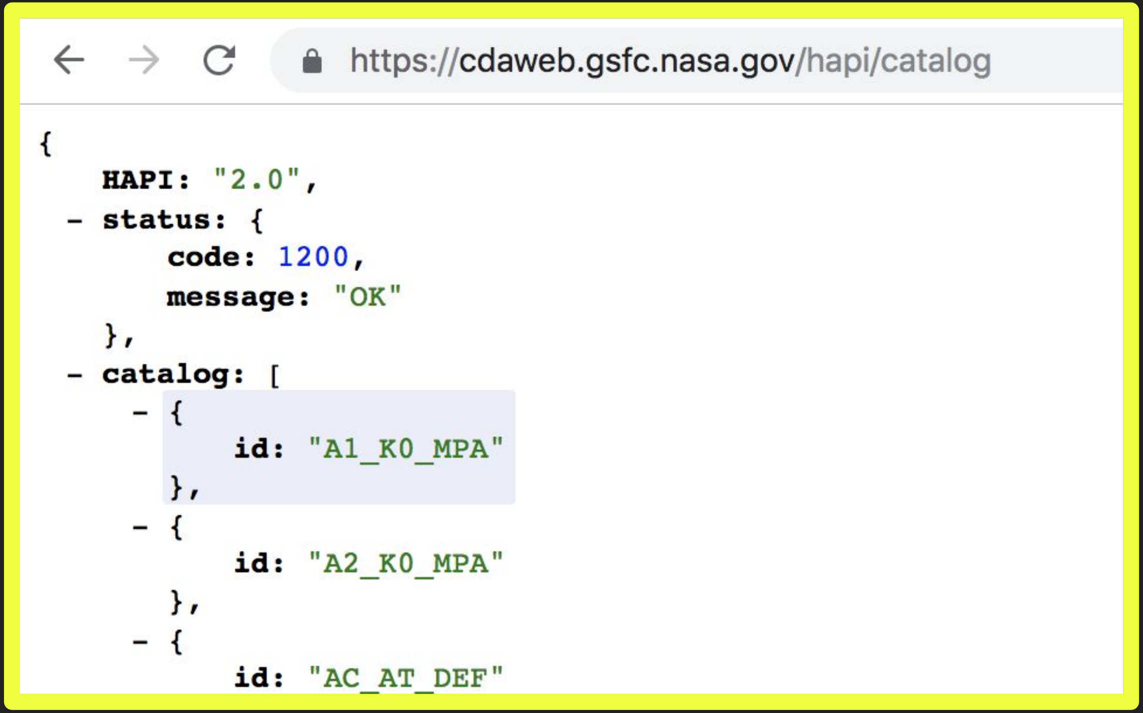


```
{
    HAPI: "2.0",
  - status: {
        code: 1200,
        message: "OK"
    },
  - outputFormats: [
        "csv",
        "binary",
        "json"
    ]
}
```

API - Metadata

http://server/hapi/catalog

Returns a list of available datasets

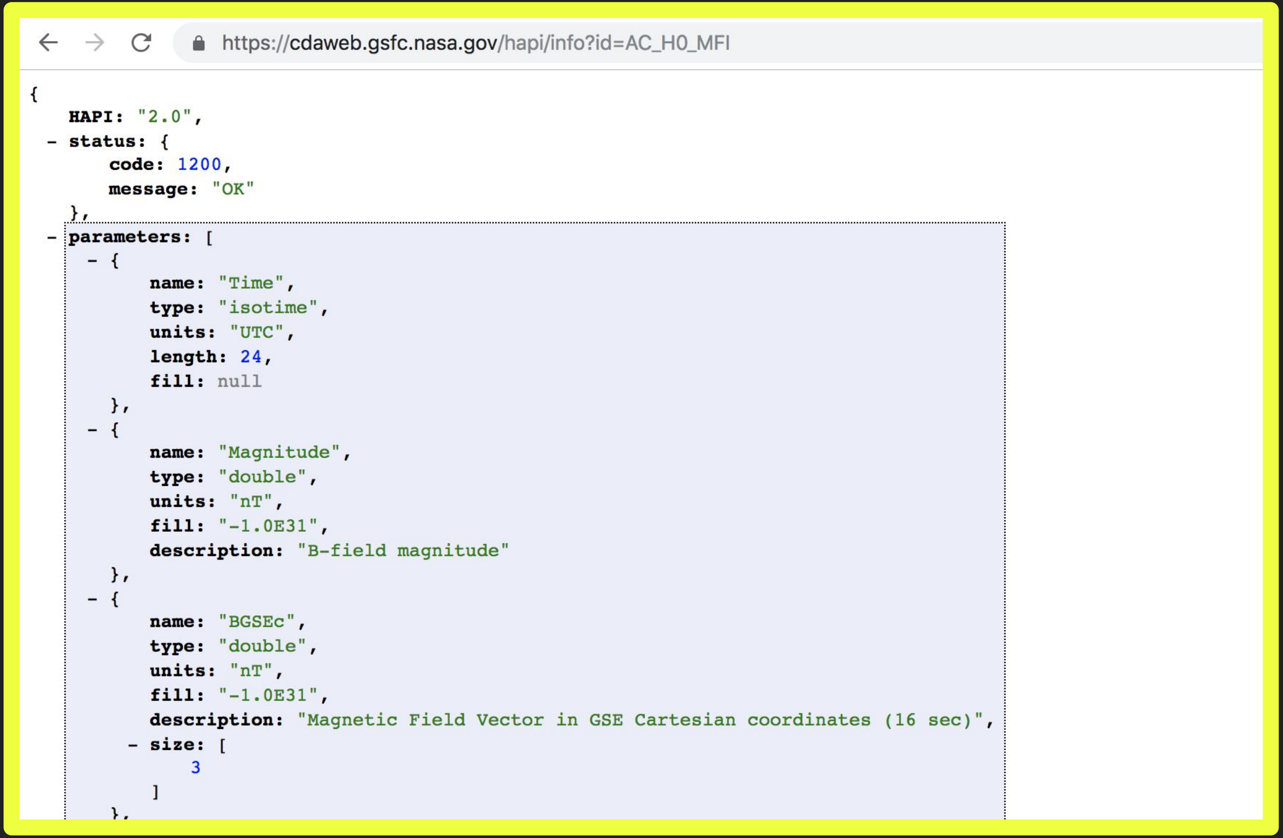


```
{
   HAPI: "2.0",
 - status: {
      code: 1200,
      message: "OK"
   },
 - catalog: [
    - {
         id: "A1_K0_MPA"
      },
    - {
         id: "A2_K0_MPA"
      },
    - {
         id: "AC_AT_DEF"
```

API - Metadata

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

Returns info for all parameters in **DATASET**



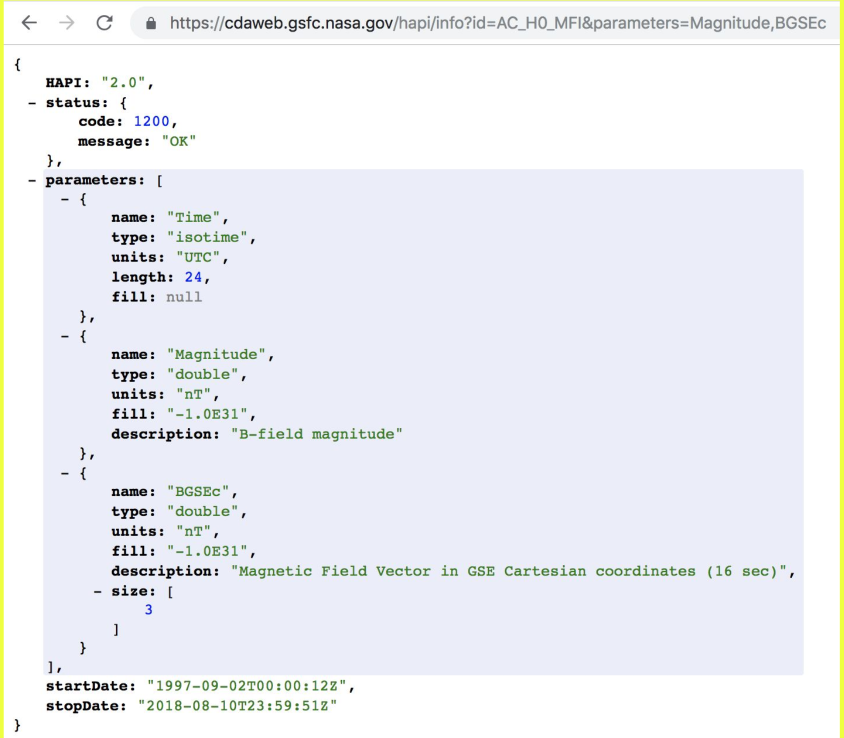
https://cdaweb.gsfc.nasa.gov/hapi/info?id=AC_H0_MFI

```
{
   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_
&start=ISO8601*	_*Constrained ISO8601 timestamp_
&stop=ISO8601*	_*Constrained ISO8601 timestamp_
[&format={csv,json,binary}]	_Server only needs to support format=csv._

Default output 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
...
```