FDSN Web Service Specification
fdsnws-availability

PROPOSAL DRAFT Version 1.0 (2019-04-03)

Purpose

To specify a web service interface for the exchange of time series data availability within the context of the International Federation of Digital Seismograph Networks (FDSN). The intention is to provide a specification that, when implemented at different FDSN data centers, can be used interchangeably by the same client software. Combined with common FDSN web service specifications, this document fully defines the request parameters and expected results.

Common service characteristics

This document must be combined with the common service characteristics defined in version 1.1, or later version 1.x, of the "FDSN-WS-Specification" for a full specification. The common service characteristics include versioning scheme, general calling pattern, common service methods, common error responses, and more.

Service methods

The following methods shall be supported by the service:

> **query** – to submit a data request
>
> **queryauth** – to authenticate and submit a request, optional
>
> **extent** - to submit a data request
>
> **extentauth** – to authenticate and submit a request, optional
>
> **version** – to request the full service version number
>
> **application.wadl** – to request a WADL for the interface

Purpose of the *query* and *extent* methods:

The fdsnws-availability service is designed to return a description of available time series data. To support a broad range of use cases there are two major result styles. The first, returned by the *query* and *queryauth* methods, includes full resolution time series listings with an option to merge overlapping time spans. The second, returned by *extent* and *extentauth*, includes time series listings with only the earliest and latest data available.

Purpose of the *queryauth* and *extentauth* methods:

The optional *queryauth* and *extentauth* methods behave identically to their *query* and *extent* counterparts except that submission requires authentication. The data center may then include restricted data that would be available using the credentials in the results.

Common parameters for the *query* and *extent* methods:

The service shall accept requests formulated using the parameters identified in Table 1. The **alias** values are acceptable synonyms for the given **parameter** name.

Table 1. Common parameters for the *query* and *extent* methods:

Parameter	Alias	Support	Default	Minimum	Maximum	Type	Unit
starttime	**start**	**required**	**[Any]**	**[Any valid time]**		**time**	**UTC**
Limit results to time series samples on or after the specified start time.							
endtime	**end**	**required**	**[Any]**	**[Any valid time]**		**time**	**UTC**
Limit results to time series samples on or before the specified end time.							
network	**net**	**required**	**[Any]**	**[Valid ASCII or * or ?]**		**string**	
Select one or more network codes. Can be FDSN network codes or data center defined codes. Multiple codes are comma-separated.							
station	**sta**	**required**	**[Any]**	**[Valid ASCII or * or ?]**		**string**	
Select one or more FDSN station codes. Multiple codes are comma-separated.							
location	**loc**	**required**	**[Any]**	**[Valid ASCII or * or ?]**		**string**	
Select one or more FDSN location identifiers. Multiple identifiers are comma-separated. As a special case "--" (two dashes) will be translated to a string of two space characters to match blank location IDs.							
channel	**cha**	**required**	**[Any]**	**[Valid ASCII or * or ?]**		**string**	
Select one or more FDSN channel codes. Multiple codes are comma-separated.							
quality		**optional**	**[Any]**	**[Valid ASCII or * or ?]**		**string**	
Select a specific FDSN quality indicator, handling is data center dependent.							
merge		**optional**	**[None]**			**string**	
If set to one or more of the following values, time spans are merged as described. Multiple values may be specified as a comma-delimited list, e.g. *merge=samplerate,quality*.							
samplerate: time spans from data with differing sample rates will be grouped together. If specified this field will be omitted from the result.							
quality: time spans from data with differing quality codes will be grouped together. If specified this field will be omitted from the result.							
overlap: time spans from data that overlap will be merged together. This option does not apply to the *extent* method.							
orderby		**optional**	**nslc_time_quality_samplerate**			**string**	
Sort results by one of the following values in the order specified:							
nslc_time_quality_samplerate: network, station, location, channel, time-range, quality, sample-rate (**default**)							
latestupdate: update-date (past to present), network, station, location, channel, time-range, quality, sample-rate							
latestupdate_desc: update-date (present to past), network, station, location, channel, time-range, quality, sample-rate							
timespancount: number of timespans (small to large), network, station, location, channel, time-range, quality, sample-rate							
timespancount_desc: number of timespans (large to small), network, station, location, channel, time-range, quality, sample-rate							
limit		**optional**	**[Unlimited]**			**integer**	
Limit results to the specified number of timespans.							
includerestricted		**optional**	**false**			**boolean**	
If true, all data are reported. If false, only data that can be openly accessed …							
format		**optional**	**text**			**string**	
Specify format of result, either 'text' (the default) or 'geocsv' or 'json' or 'request'.							
nodata		**optional**	**204**	***204 or 404***		**string**	
Select HTTP status code for "no data", either '204' (default) or '404'.							

Parameters specific to the *query* method:

Additional parameters for the *query* method are in Table 2.

Table 2. Additional *query* method parameters:

Parameter	Alias	Support	Default	Minimum	Maximum	Type	Unit
mergegaps		**optional**	**0.0**			**float**	
If set, merge time spans that are separated by the specified tolerance in seconds.							
show		**optional**	**[None]**	**"latestupdate"**		**string**	
If set to ***latestupdate***, the latest times at which data contributing to the returned time spans were loaded into the repository are included in the result. This option applies to all formats except 'request'.							

Requests using the *query* and *extent* methods:

These parameters may be submitted using either of the HTTP GET or POST methods. The POST method is useful to support the submission of a large number of data selections.

For the GET method, the parameters should be submitted as key=value pairs and may not be specified more than once; if a parameter is submitted multiple times the result is undefined.

For the POST method, all parameters must be submitted as part of the POST body. The non-repeating parameters should be specified as key=value pairs on separate lines, while the data identifier (*network*, *station*, *location*, *chanel*) and time (*starttime*, *endtime*) parameters are repeated as many times as necessary following this pattern:

```
parameter1=value
parameter2=value
NET STA LOC CHA STARTTIME ENDTIME
NET STA LOC CHA STARTTIME ENDTIME
NET STA LOC CHA STARTTIME ENDTIME
```

This would be POSTed to the URI "<site>/fdsnws/availability/1/query" or "<site>/fdsnws/availability/1/extent".

All rules for parameters apply equally whether submitted using the GET or POST methods with the exception of blank location IDs, which must be specified as "--" in the POST body due to spaces being used as the field separator.

Responses for the ***query*** and ***extent*** methods

The results of a successful request shall be returned as text by default, using the MIME type **text/plain**. If the *format* parameter is 'geocsv' the results shall be returned in GeoCSV formatting using MIME type **text/csv**. If the format parameter is 'request' the results shall be returned in the FDSN web service POST-style request format using MIME type **text/plain**.

The earliest and latest times included in a 'request' format response should be truncated to the start and end times if they are submitted in the request. This allows the result to be used as a request without selecting more than originally desired. At a data center's option the earliest and latest times for all other formats may be trimmed as well.

3

Behavior for the ***queryauth*** and ***extentauth*** methods

HTTP Digest Authentication (RFC 2617) should be requested from the client. After successful authentication, a request should be accepted and handled as per the *query* or *extent* method respectively. Authentication credentials (user database and associated licensing schemes) are data center specific.

Behavior for the ***version*** method

The service shall return the return the implementation version as a simple text string using the MIME type **text/plain**. Any parameters submitted with the method will be ignored.

Behavior for the ***application.wadl*** method

The service shall return a WADL conformant description of the interface using the MIME type **application/xml**. Any parameters submitted with the method will be ignored. The WADL shall describe all parameters supported by the interface and is primarily used to document which optional parameters are supported.

Behavior for the **error** conditions

All errors shall use the type and pattern described in the common FDSN web service specifications.

Text output format

This 'text' output format contains one timespan per line with fields separated by spaces (ASCII 32) as needed for column alignment. This output is aligned for human readability. Lines beginning with a hash character ("#": ASCII decimal 35) should be considered comment lines. A single header line describing the columns is strongly recommended.

The text format for the *query* method follows this pattern:

```
#Network Station Location Channel Quality SampleRate Earliest Latest
NET   STA   LOC   CHAN   QUAL   SAMPRATE   EARLIEST   LATEST
NET   STA   LOC   CHAN   QUAL   SAMPRATE   EARLIEST   LATEST
…
```

For the *query* method, if the *show* parameter is set to *latestupdate*, the text format includes an additional column with the latest update time:

```
#Network Station Location Channel Quality SampleRate Earliest Latest Updated
NET   STA   LOC   CHAN   QUAL   SAMPRATE   EARLIEST   LATEST   UPDATED
NET   STA   LOC   CHAN   QUAL   SAMPRATE   EARLIEST   LATEST   UPDATED
…
```

For the *query* method, if the *merge* parameter is set to *samplerate,quality* these respective fields are omitted from the result:

```
#Network Station Location Channel Earliest Latest Updated
NET   STA   LOC   CHAN   EARLIEST   LATEST   UPDATED
NET   STA   LOC   CHAN   EARLIEST   LATEST   UPDATED
…
```

The text format for the *extent* method follows this pattern:

```
#Network Station Location Channel Quality SampleRate Earliest Latest Updated TimeSpans Restriction
NET   STA   LOC   CHAN   QUAL   SAMPRATE   EARLIEST   LATEST   UPDATED   SPANCOUNT   STATUS
NET   STA   LOC   CHAN   QUAL   SAMPRATE   EARLIEST   LATEST   UPDATED   SPANCOUNT   STATUS
…
```

In these text listings:

- NET, STA, LOC, CHAN and QUAL are FDSN data source identifiers
- SAMPRATE is sample rate in Hertz
- EARLIEST, LATEST and UPDATED are date-times in the form specified in the common FDSN web service specifications
- SPANCOUNT is the number of timespans contributing to an extent
- STATUS is the restricted status, one of: OPEN, RESTRICTED, or PARTIAL

GeoCSV output format

The 'geocsv' version of results is a text format designed to contain enhanced metadata and be consistently machine readable. The output format contains the exact same data at the 'text' format within a [GeoCSV](#) container, which provides annotations to describe the data structure. For this output, the column delimiter should be vertical bar characters ("|": ASCII decimal 124). Field entries cannot contain vertical bar characters.

To illustrate, the GeoCSV results for the *extent* method would follow this pattern:

```
#dataset: GeoCSV 2.0
#delimiter: |
#field_unit: unitless|unitless|unitless|unitless|unitless|hertz|ISO_8601|ISO_8601|ISO_8601|unitless|
unitless
#field_type: string|string|string|string|string|float|datetime|datetime|datetime|integer|string
Network|Station|Location|Channel|Quality|SampleRate|Earliest|Latest|Updated|TimeSpans|Restriction
NET|STA|LOC|CHAN|QUAL|SR|EARLIEST|LATEST|UPDATED|SPANCOUNT|STATUS
NET|STA|LOC|CHAN|QUAL|SR|EARLIEST|LATEST|UPDATED|SPANCOUNT|STATUS
…
```

JSON output format

The 'json' output format contains the same information available in the text-based formats. The schema is defined in JSON Schema in the associated file named 'fdsnws-availability-schema-1.0.json'.

[NOTE: if approved, the schema document can be on the fdsn.org site and referenced]

Request output format

The 'request' output format is a text listing of the selected time spans or extents in the FDSN web service POST-style request format. This response is designed for easy submission to other FDSN web services, e.g. fdsnws-dataselect, fdsnws-station. The pattern of this format is:

```
NET STA LOC CHA STARTTIME ENDTIME
NET STA LOC CHA STARTTIME ENDTIME
NET STA LOC CHA STARTTIME ENDTIME
```

Time span definition

The *query* method of this service returns timespans representing the coverage of continuous time series data or the range of non-time series data. The earliest and latest times reported are, respectively, the times of the first and last sample times in the span. When determining continuous time spans from otherwise fragmented data, such as data records, it is recommended to use a time tear tolerance equal to or less than ½ the sampling period.

Time span representations must not be generated that combine spans across independent networks, specifically the case where network codes are re-used for temporary networks.

Examples

Requesting extents in default 'text' format with all extra columns:

https://DataCenter/fdsnws/availability/1/extent?network=IU&station=ANMO&channel=BHZ

```
#Network Station Location Channel Quality SampleRate Earliest Latest Updated TimeSpans Restriction
IU    ANMO --  BHZ  M    20.0 1989-08-29T22:07:20.482000Z 1998-10-26T17:38:43.640000Z 2018-05-08T21:47:09Z  611      OPEN
IU    ANMO 00  BHZ  M     0.0 2002-08-28T18:17:51.000000Z 2008-05-23T23:09:24.000000Z 2017-12-06T03:42:35Z  9051     OPEN
IU    ANMO 00  BHZ  M    20.0 1998-10-26T20:35:58.310000Z 2018-07-09T20:45:47.369000Z 2018-07-10T08:29:54Z  2238     OPEN
IU    ANMO 00  BHZ  M    40.0 2018-07-09T20:46:40.594000Z 2019-02-04T23:59:59.994000Z 2019-02-05T09:27:24Z  1        OPEN
IU    ANMO 10  BHZ  M     0.0 2002-11-28T04:00:00.000000Z 2007-06-01T19:29:00.000000Z 2017-12-06T03:09:51Z  47       OPEN
IU    ANMO 10  BHZ  M    40.0 1998-10-26T20:35:59.072000Z 2019-02-04T23:59:59.994000Z 2019-02-05T09:27:24Z  176      OPEN
```

Requesting extents in GeoCSV format:

https://DataCenter/fdsnws/availability/1/extent?format=geocsv&network=IU&station=ANMO&channel=BHZ

```
#dataset: GeoCSV 2.0
#delimiter: |
#field_unit: unitless|unitless|unitless|unitless|unitless|hertz|ISO_8601|ISO_8601|ISO_8601|unitless|unitless
#field_type: string|string|string|string|string|float|datetime|datetime|datetime|integer|string
network|station|location|channel|quality|sample_rate|earliest|latest|updated|timespans|restriction
IU|ANMO||BHZ|M|20.0|1989-08-29T22:07:20.482000Z|1998-10-26T17:38:43.640000Z|2018-05-08T21:47:09Z|611|OPEN
IU|ANMO|00|BHZ|M|0.0|2002-08-28T18:17:51.000000Z|2008-05-23T23:09:24.000000Z|2017-12-06T03:42:35Z|9051|OPEN
IU|ANMO|00|BHZ|M|20.0|1998-10-26T20:35:58.310000Z|2018-07-09T20:45:47.369000Z|2018-07-10T08:29:54Z|2238|OPEN
IU|ANMO|00|BHZ|M|40.0|2018-07-09T20:46:40.594000Z|2019-04-02T23:59:59.994000Z|2019-04-03T08:27:06Z|4|OPEN
IU|ANMO|10|BHZ|M|0.0|2002-11-28T04:00:00.000000Z|2007-06-01T19:29:00.000000Z|2017-12-06T03:09:51Z|47|OPEN
IU|ANMO|10|BHZ|M|40.0|1998-10-26T20:35:59.072000Z|2019-04-02T23:59:59.994000Z|2019-04-03T08:27:06Z|180|OPEN
```

Requesting extents in 'request' format for a 36 hour time window:

https://DataCenter/fdsnws/availability/1/extent?format=request&network=IU&station=ANMO&channel=BHZ&starttime=2018-01-31T00:00:00&endtime=2018-02-01T12:00:00

```
IU ANMO 00 BHZ 2018-01-31T00:00:00.000000 2018-02-01T12:00:00.000000
IU ANMO 10 BHZ 2018-01-31T00:00:00.000000 2018-02-01T12:00:00.000000
```

Requesting extents in 'json' format for a 36 hour time window:

```
{
    "created": "2019-02-13T22:29:09Z",
    "schemaVersion": "1.0",
    "datasources": [
        {
            "network": "XX",
            "station": "STA",
            "location": "",
            "channel": "BHZ",
            "quality": "D",
            "samplerate": 20.0,
            "earliest": "2018-01-31T00:00:00.000000Z",
            "latest": "2018-02-01T12:00:00.000000Z",
            "timespanCount": 2,
            "updated": "2019-02-03T14:01:00Z",
            "restriction": "OPEN"
        },
        {
            "network": "XX",
            "station": "STA",
            "location": "",
            "channel": "HHZ",
            "quality": "D",
            "samplerate": 100.0,
            "earliest": "2018-01-31T00:00:00.000000Z",
            "latest": "2018-02-01T12:00:00.000000Z",
            "timespanCount": 1,
            "updated": "2019-02-03T14:01:00Z",
            "restriction": "RESTRICTED"
        }
    ]
}
```

Requesting timespans in 'json' format for a 36 hour time window, including the additional *latestupdate* values:

https://DataCenter/fdsnws/availability/1/query?format=json&show=latestupdate&network=IU&station=ANMO&channel=BHZ&starttime=2018-01-31T00:00:00&endtime=2018-02-01T12:00:00

```
{
    "created": "2019-02-13T22:29:09Z",
    "schemaVersion": "1.0",
    "datasources": [
        {
            "network": "XX",
            "station": "STA",
            "location": "",
            "channel": "BHZ",
            "quality": "D",
            "samplerate": 20.0,
            "updated": "2019-02-03T14:01:00Z",
            "timespans": [
                [
                    "2018-01-31T00:00:00.000000Z",
                    "2018-02-01T05:01:14.250000Z"
                ],
                [
                    "2018-02-01T05:02:32.050000Z",
                    "2018-02-01T12:00:00.000000Z"
                ]

            ]
        },
        {
            "network": "XX",
            "station": "STA",
            "location": "",
            "channel": "HHZ",
            "quality": "D",
            "samplerate": 100.0,
            "updated": "2019-02-03T14:01:00Z",
            "timespans": [
                [
                    "2018-01-31T00:00:00.000000Z",
                    "2018-02-01T12:00:00.000000Z"
                ]
            ]
        }
    ]
}
```

References

Web Application Description Language (WADL) - http://www.w3.org/Submission/wadl/

GeoCSV, tabular text formatting for geoscience data - http://geows.ds.iris.edu/documents/GeoCSV.pdf

JSON Schema vocabulary format - https://json-schema.org/

Changes

2019-02-13 - Initial specification proposal

2019-04-03 - Redrafted proposal based on Working Group feedback

- Change *timespan* method to *query*.
- Remove 'show' parameter for *extent* method and make latest update, time span count and restriction status fixed fields in the output format.
- Add new 'mergegaps' parameter to *query* method.
- Remove functionally duplicated 'mergetimespans' method in favor of existing 'merge=timespans'.
- Clarify when merging sample rates or qualities the respective field is omitted in the result.
- Clarify that timespan representations should not be created that cross individual network designations, specifically for network codes that are re-used for temporary deployments.
- Change text format headers to fully express data identifiers for consistency with other FDSN web service formats.