The GSAC[1] Retailer Service CGI API

July 11, 2003
Michael Scharber

Cecil H. and Ida M. Green Institute of Geophysics and Planetary Physics, Scripps Institution of Oceanography, La Jolla, California

Contents

Introduction

The GSAC Retailer Service API provides a simple, Common Gateway Interface (CGI) to centralized GSAC content stored in Retailer databases. This basic interface is comprised of a set of request tokens specifically for the GSAC project. Through this interface, which is accessible through either HTTP POST or GET requests, a user or application can create custom GSAC queries matching the immediate and specific needs of the user/application. This interface, in turn, forms the communication linkage between GSAC applications and the set of existing GSAC Retailers involved in the project.

The requests themselves, which very much resemble standard relational database SQL, are received by the GSAC Retailer server (the host to which the HTTP request is made), parsed by that server, interpreted and then executed (as SQL) on the local database schema of the Retailer. The results are then packaged in a format of the users choosing, prepended by a small number of Retailer "directive" lines, and returned to the client.

Each request made through this API is treated as a wholly autonomous, sessionless transaction whereby there is only one request made and one response provided. Subsequent requests are not related to one another (currently).

It is important to note that ALL GSAC client applications (CGI, command-line executables, HTML forms, Java applets, etc) use this API to service client requests. The transparency of the API was intended to serve a broad spectrum of retrieval methods, allowing for the development of a variety of client platforms by anyone with an expressed interest in utilizing the GSAC for their own needs and/or the needs of their user community.

All GSAC Retailer Services are accessed in the same fashion, via a single Universal Resource Identifier (URI) (using GET) or HTML form (using POST). GET requests utilize a set of name=value pairs, separated by ampersands (&), following a question mark (?) after the hostname and "/GSAC" of the GSAC Retailer Service being accessed. HTTP POST request work in a similar fashion, the primary difference being the the browser or application intercepts an HTML form and assembles the request for HTTP transmission for the user.

At any rate, one can access a GSAC Retailer Service directly with only limited web development experience (HTML at minimum). Outlined below are the set of available (required and optional) tokens offered by all GSAC Retailer Services, followed by some samples to illustrate the translation of popular natural language queries into actual GSAC Retailer Service requests.

---

1 GSAC stands for GPS Seamless Archive Centers and is a UNAVCO (University NAVSTAR Consortium) project.

GSAC Retailer Service Tokens

S=[help | status | info | count]

About Stands for the type of query requested of the GSAC Retailer.
Required No, however, will default to "help" unless otherwise specified.
Format Value must be one of the four supported enumerations (see below).
Enumeration The following types of queries are currently supported:

help	Returns this document.
status	Brief information about the current status of the requested GSAC Retailer Service.
info	GSAC Monument/Data Holdings records matching constraints applied by other request tokens.
count	Number of records matching constraints applied by other request tokens.

O=[ascii | html | csv | dhr::mc | dhr::dhf]

About Stands for output format.
Required No, will default to ascii unless otherwise specified.
Format Value must be one of five supported enumerations (see below).
Enumeration The following formats are currently supported:

ascii	basic (space-separated) ASCII text
html	HTML format, with hyperlinks where applicable
csv	Comma Separated Value (CSV) format
dhr::mc	CSV records converted from Perl GSAC::DHR::MC object structures.
dhr::dhf	CSV records converted from Perl GSAC::DHR::DHF object structures.

C=<string>

About Stands for the name of the client calling the Retailer Service. In place only to provide future flexibility in GSAC Retailer Services to allow for the transmission and receipt of informative messages and/or directives from a GSAC Retailer Service to the client in instances where a client upgrade is suggested or required.
Required No. This token is completely optional and, currently, without function.
Format No particular format required.
(optional, stands for client name)

F=<field clause>

About Stands for “query return fields”.

Required No, though without it you will only receive a total “count” of matching records.

Format A comma-separated list of valid GSAC Retailer Schema return fields formatted like:

f1,f2,f3...fn

where the set of valid field names (represented by the pattern fn above, are listed below.

Enumeration The set of possible field names for use in field clause are listed below. Further information on the actual GSAC Wholesaler publication field counterparts to the schema names listed below can be found at:

http://www.unavco.ucar.edu/data_support/data/gsac/GSAC-1.html

The format is:

gsac-schema-table-alias.field-name

> dh represents the data_holding table, a relation where the entities modeled are individual pieces of data published by a given GSAC Wholesaler to the GSAC. Typically these records are associated with one or more URLs representing one or more locations where a particular set of data (usually a single file) can be found.
> mc represents the monument_catalog table, a relation where the entities modeled are individual monuments (associated with a single point in three-dimensional Cartesian coordinates) as published by a particular GSAC Wholesaler. Any given monument may be published by more than one Wholesaler, using different identifiers, if so desired.

dh.unique_info_id	mc.unique_site_id
dh.wholesaler	mc.wholesaler
dh.data_type	mc.char_id
dh.unique_site_id	mc.descriptive_id
dh.start_time	mc.dhr_create_time
dh.end_time	mc.x
dh.dhr_create_time	mc.y
dh.info_url	mc.z
dh.file_size	mc.coord_accuracy
dh.file_create_time	mc.latitude
dh.file_checksum	mc.longitude
dh.provider	mc.height
dh.file_grouping	mc.data_record_count
dh.file_compression	mc.data_record_min_time
dh.source_unique_info_id	mc.data_record_max_time
dh.source_wholesaler	

R=<restriction clause>

About

Stands for “query restrictions”. Embodies the constraints to place on a request of the following types:

info
count

This component is used to act as the limiting aspect of a query and may include use of sets of parenthases to delimit logical precedence.

Required

No. Though most queries will fail (as deemed “too expensive”) if no restrictions are applied.

Format

Like the fields used in the **F** component of a request, fully-prefixed GSAC Retailer Schema field identifiers (like "dh.info_url", "mc.unique_site_id", etc) must be used, along with a combination of relational operators (listed below), without spaces. Together with special functions, a restrictions clause can be built.

Relational Operators

AND	Just like a logical "and" in SQL
OR	Just like a logical "or" in SQL
NEQ	In place of "!=", means "not equal to"
EQ	In place of "=", means "equal to"
IN	In place of " in ", used when comparing a value to a set of possible values. For example (" where name in ('James','Paul','Tim')".
ISNULL	In place of "is null" means "no value"
LIKE	For use with wildcards such as "%". For example (“where char_idLIKE'sio%')
ORDERBY	For use in specifying ordering of return columns.
GT	In place of ">", means "greater than"
GTE	In place of ">=", means "greater than or equal to"
LT	In place of "<", means "less than"
LTE	In place of "<=", means "less than or equal to"
MINUS	In place of "-" (in mathematical sense)
PLUS	In place of "+" (in mathematical sense)

Special Functions

----RETAILER-DEPENDENT----

These functions are provided at the discretion of the Retailer. Use the **status** *GSAC Retailer Service request to see if a particular Retailer actually implements one or more of these functions.*

bounding_box(xmin ymin,xmax ymax)
Use longitude and latitude {decimal degrees} to narrow your query to a rectangular geographical region.

proximity(x,y,r)
Use longitude and latitude {decimal degrees} and a radius (r) in kilometers to constrain a query to a circular geographical region.

Notes

- Like SQL, all string identifiers in an **R** request component used in direct comparitive operations must be enclosed by single quotation marks. For example, “mc.char_idEQ'sio3'”.
- When using datetime values in comparitive operations you must use the 13 digit datetime format used by GSAC Retailer schemas : YYYYDDDHHMMSS.
- Depending on the client accessing a GSAC Retailer Service, various problems associated with the inclusion of whitespace or other special characters in a GSAC Retailer Service URI may result. Accordingly, below is a list of some of the characters in the ISO 8859 Latin 1 character set which may cause such problems, along with their associated hexidecimal escape identifiers:

%20 : " "	%2a : "*"	%3e : ">"	%7c : "|"
%21 : "!"	%2b : "+"	%3f : "?"	%7d "}"
%22 : """	%2c : ","	%40 : "@"	%7e : "~"
%23 : "#"	%2d : "-"	%5b : "["	
%24 : "$"	%2e : "."	%5c : "\"	
%25 : "%"	%2f : "/"	%5d : "]"	
%26 : "&"	%3a : ":"	%5e : "^"	
%27 : "'"	%3b : ";"	%5f : "_"	
%28 : "("	%3c : "<"	%60 : "`"	
%29 : ")"	%3d : "="	%7b : "{"	

CV=<string>

About	Stands for the version of the client calling the Retailer Service. In place only to provide future flexibility in GSAC Retailer Services to allow for the transmission and receipt of informative messages and/or directives from a GSAC Retailer Service to the client in instances where a client upgrade is suggested or required.
Required	No. This token is completely optional and, currently, without function.
Format	No particular format required. (optional, stands for client name)

U=<string>

About	Stands for the name of the user calling the Retailer Service. In place only to provide future flexibility in GSAC Retailer Services to allow for user-specific functionality such as preferences, session-tracking, computational quotas and stored queries.
Required	No. This token is completely optional and, currently, without function.
Format	No particular format required. (optional, stands for client name)

DT=[on | off]

About	Flag indicating whether or not particular return fields from the GSAC Retailer Schema which happen to represent a date and time value should be translated from their native database representation into the standard GSAC datetime format or not. By default the status of this token is “on”, unless otherwise specified. Specifying “off” can significantly reduce the time required to return certain information from queries which include one or more datetime return fields. In “on” mode datetime return fields are formatted like : YYYY-DDDT:HH:MM:SSZ In “off” mode datetime return fields are formatted like : YYYYDDDHHMMSS
Required	No. This token is completely optional, and will default to “on” unless otherwise specified.
Format	Value must be either “on” or “off” (no quotation marks).

Example GSAC Retailer Service URI Requests

a. "Return a list of monuments from the GSAC that are within 1000 kilometers of the point 120° West, 50° North".

http://gsac.ucsd.edu/GSAC?C=GSAC%20Wizard&CV=1.30&O=csv&S=info&F=mc.wholesaler,mc.char_id,mc.longitude,mc.latitude,mc.height&R=proximity(-120,50,1000)

b. "Return a list of all data records available in the GSAC which pertain to data for the timespan of 2003 day 137 through 2003 day 139, AND whose associated monument is within 1000 kilometers of the point 120° West, 50° North".

http://gsac.ucsd.edu/GSAC?C=GSAC%20Wizard&CV=1.30&O=csv&S=info&F=dh.wholesaler,dh.info_url,dh.file_size,dh.start_time&R=(proximity(-120,50,1000))AND(((mc.data_record_min_timeLT2003139235959ANDmc.data_record_min_timeGT2003137000000)OR(mc.data_record_max_timeGT2003137000000ANDmc.data_record_max_timeLT2003139235959)OR(mc.data_record_min_timeLT2003137000000ANDmc.data_record_max_timeGT2003139235959)))AND(dh.start_timeGTE2003137000000ANDdh.start_timeLTE2003139235959)

c. "Return help from a particular GSAC Retailer Service".

http://gsac.ucsd.edu/GSAC?S=help

d. "Return the number of monuments within a specific geographical region of interest."

http://gsac.ucsd.edu/GSAC?S=count&R=bounding_box(-120,50,-110,30)

e. "Return the number of data files published to the GSAC from all participating Wholesalers which are associated with the common multi-character identifier 'sio3'".

http://gsac.ucsd.edu/GSAC?S=count&R=mc.char_idEQ'sio3'ANDdh.info_urlNEQ'bogus'

GSAC Contact Information

GSAC development is an ongoing body of work conducted primarily by staff at the Scripps Orbit and Permanent Array Center (SOPAC) with peer supervision provided by UNAVCO Inc. Organizations or individuals interested in learning more about what the GSAC can offer in terms of a specific application or computing problem are invited to contact any of the GSAC peer community members listed below:

Fran Boler	fboler@unavco.ucar.edu
Michael Scharber	mscharber@ucsd.edu
Yehuda Bock	ybock@ucsd.edu
Chuck Meertens	chuckm@unavco.ucar.edu

In addition to the contacts listed above, there are two GSAC websites available to the public:

http://gsac.ucsd.edu
http://www.unavco.ucar.edu/data_support/data/gsac/gsac.html