Registration Data Access Protocol Query Format
draft-ietf-weirds-rdap-query-07

Abstract

   This document describes uniform patterns to construct HTTP URLs that
   may be used to retrieve registration information from registries
   (including both Regional Internet Registries (RIRs) and Domain Name
   Registries (DNRs)) using "RESTful" web access patterns.

Status of This Memo

   This Internet-Draft is submitted in full conformance with the
   provisions of BCP 78 and BCP 79.

   Internet-Drafts are working documents of the Internet Engineering
   Task Force (IETF). Note that other groups may also distribute
   working documents as Internet-Drafts. The list of current Internet-
   Drafts is at http://datatracker.ietf.org/drafts/current/.

   Internet-Drafts are draft documents valid for a maximum of six months
   and may be updated, replaced, or obsoleted by other documents at any
   time. It is inappropriate to use Internet-Drafts as reference
   material or to cite them other than as "work in progress."

   This Internet-Draft will expire on April 04, 2014.

Table of Contents

1. Conventions Used in This Document

   The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT",
   "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this
   document are to be interpreted as described in RFC 2119 [RFC2119].

1.1. Acronyms and Abbreviations

      IDN: Internationalized Domain Name
      IDNA: Internationalized Domain Names in Applications
      DNR: Domain Name Registry
      RDAP: Registration Data Access Protocol
      REST: Representational State Transfer State Transfer. The term
      was first described in a doctoral dissertation [REST].
      RESTful: an adjective that describes a service using HTTP and the
      principles of REST.
      RIR: Regional Internet Registry

[2](#). Introduction

This document describes a specification for querying registration
data using a RESTful web service and uniform query patterns. The
service is implemented using the Hypertext Transfer Protocol (HTTP)
[RFC2616].

The protocol described in this specification is intended to address
deficiencies with the WHOIS protocol [RFC3912] that have been
identified over time, including:

o Lack of standardized command structures,
o lack of standardized output and error structures,
o lack of support for internationalization and localization, and
o lack of support for user identification, authentication, and
   access control.

The patterns described in this document purposefully do not encompass
all of the methods employed in the WHOIS and RESTful web services of
all of the RIRs and DNRs. The intent of the patterns described here
are to enable queries of:

o networks by IP address,
o autonomous system numbers by number,
o reverse DNS meta-data by domain,
o name servers by name,
o registrars by name, and
o entities (such as contacts) by identifier.

It is envisioned that each registry will continue to maintain NICNAME
/WHOIS and/or RESTful web services specific to their needs and those
of their constituencies, and the information retrieved through the
patterns described here may reference such services.

Likewise, future IETF standards may add additional patterns for
additional query types. A simple pattern namespacing scheme is
described in Section 5 to accomodate custom extensions that will not
interfere with the patterns defined in this document or patterns
defined in future IETF standards.

WHOIS services, in general, are read-only services. Therefore URL
[RFC3986] patterns specified in this document are only applicable to
the HTTP [RFC2616] GET and HEAD methods.

This document does not describe the results or entities returned from
issuing the described URLs with an HTTP GET. JSON [RFC4627] result
formatting and processing is described in
[I-D.ietf-weirds-json-response].

Additionally, resource management, provisioning and update functions
are out of scope for this document. Registries have various and
divergent methods covering these functions, and it is unlikely a
uniform approach for these functions will ever be possible.

HTTP contains mechanisms for servers to authenticate clients and for
clients to authenticate servers (from which authorization schemes may
be built) so such mechanisms are not described in this document.
Policy, provisioning, and processing of authentication and
authorization are out-of-scope for this document as deployments will
have to make choices based on local criteria. Specified
authentication mechanisms MUST use HTTP.

3. Path Segment Specification

The uniform patterns start with a base URL [RFC3986] specified by
each registry or any other service provider offering this service.
The base URL is followed by a resource-type-specific path segment.
The base URL may contain its own path segments (e.g. http://
example.com/... or http://example.com/rdap/...). The characters
used to form a path segment are limited to those that can be used to
form a URI as specified in RFC 3986 [RFC3986].

3.1. Lookup Path Segment Specification

The resource type path segments for exact match lookup are:

o 'ip': Used to identify IP networks and associated data referenced
   using either an IPv4 or IPv6 address.
o 'autnum': Used to identify autonomous system registrations and
   associated data referenced using an AS Plain autonomous system
   number.
o 'domain': Used to identify reverse DNS (RIR) or domain name (DNR)
   information and associated data referenced using a fully-qualified
   domain name.
o 'nameserver': Used to identify a name server information query
   using a host name.
o 'entity': Used to identify an entity information query using a
   string identifier.

3.1.1. IP Network Path Segment Specification

Syntax: ip/<IP address> or ip/<CIDR prefix>/<CIDR length>

Queries for information about IP networks are of the form /ip/XXX/...
or /ip/XXX/YY/... where the path segment following 'ip' is either an
IPv4 [RFC1166] or IPv6 [RFC5952] address (i.e. XXX) or an IPv4 or
IPv6 CIDR [RFC4632] notation address block (i.e. XXX/YY).

Semantically, the simpler form using the address can be thought of as a CIDR block with a bitmask length of 32 for IPv4 and a bitmask length of 128 for IPv6. A given specific address or CIDR may fall within multiple IP networks in a hierarchy of networks, therefore this query targets the "most-specific" or smallest IP network which completely encompasses it in a hierarchy of IP networks.

The IPv4 and IPv6 address formats supported in this query are described in section 3.2.2 of [RFC3986], as IPv4address and IPv6address ABNF definitions. Any valid IPv6 text address format [RFC4291] can be used, compressed or not compressed. The restricted rules to write a text representation of an IPv6 address [RFC5952] are not mandatory. However, the zone id [RFC4007] is not appropriate in this context and therefore prohibited.

This is an example URL for the most specific network containing 192.0.2.0:

/ip/192.0.2.0

This is an example of a URL the most specific network containing 192.0.2.0/24:

/ip/192.0.2.0/24

This is an example URL for the most specific network containing 2001:db8:1:1::1:

/ip/2001:db8:1:1::1

3.1.2. Autonomous System Path Segment Specification

Syntax: autnum/<autonomous system number>

Queries for information regarding autonomous system number registrations are of the form /autnum/XXX/... where XXX is an AS Plain autonomous system number [RFC5396]. In some registries, registration of autonomous system numbers is done on an individual number basis, while other registries may register blocks of autonomous system numbers. The semantics of this query are such that if a number falls within a range of registered blocks, the target of the query is the block registration, and that individual number registrations are considered a block of numbers with a size of 1.

For example, to find information on autonomous system number 65551, the following path would be used:

/autnum/65551

The following path would be used to find information on 4-byte
autonomous system number 65538:

/autnum/65538

3.1.3. Domain Path Segment Specification

Syntax: domain/<domain name>

Queries for domain information are of the form /domain/XXXX/...,
where XXXX is a fully-qualified domain name [RFC4343] in either the
in-addr.arpa or ip6.arpa zones (for RIRs) or a fully-qualified domain
name in a zone administered by the server operator (for DNRs).
Internationalized domain names represented in either A-label or
U-label format [RFC5890] are also valid domain names. IDN labels
SHOULD NOT be represented as a mixture of A-labels and U-labels.

If the client sends the server an IDN in U-label format, servers that
support IDNs MUST convert the IDN into A-label format and perform
IDNA processing as specified in RFC 5891 [RFC5891]. The server
should perform an exact match lookup using the A-label.

The following path would be used to find information describing the
zone serving the network 192.0.2/24:

/domain/2.0.192.in-addr.arpa

The following path would be used to find information describing the
zone serving the network 2001:db8:1::/48:

/domain/1.0.0.0.8.b.d.0.1.0.0.2.ip6.arpa

The following path would be used to find information for the
example.com domain name:

/domain/example.com

The following path would be used to find information for the
xn--xemple-9ua.example IDN:

/domain/xn--xemple-9ua.example

3.1.4. Name Server Path Segment Specification

Syntax: nameserver/<name server name>

The <name server name> parameter represents a fully qualified name as
specified in RFC 952 [RFC0952] and RFC 1123 [RFC1123].

Internationalized names represented in either A-label or U-label
format [RFC5890] are also valid name server names. IDN labels SHOULD
NOT be represented as a mixture of A-labels and U-labels.

If the client sends the server an IDN in U-label format, servers that
support IDNs MUST convert the IDN into A-label format and perform
IDNA processing as specified in RFC 5891 [RFC5891]. The server
should perform an exact match lookup using the A-label.

The following path would be used to find information for the
ns1.example.com name server:

/nameserver/ns1.example.com

The following path would be used to find information for the
ns1.xn--xemple-9ua.example name server:

/nameserver/ns1.xn--xemple-9ua.example

3.1.5. Entity Path Segment Specification

Syntax: entity/<handle>

The <handle> parameter represents an entity (such as a contact,
registrant, or registrar) identifier. For example, for some DNRs
contact identifiers are specified in RFC 5730 [RFC5730] and RFC 5733
[RFC5733].

The following path would be used to find information for the entity
associated with handle CID-4005:

/entity/CID-4005

3.1.6. Help Path Segment Specification

Syntax: help

The help path segment can be used to request helpful information
(command syntax, terms of service, privacy policy, rate limiting
policy, supported authentication methods, supported extensions,
technical support contact, etc.) from an RDAP server. The response
to "help" should provide basic information that a client needs to
successfully use the service. The following path would be used to
return "help" information:

/help

3.2. Search Path Segment Specification

   The resource type path segments for search are:

   o 'domains': Used to identify a domain name information search using
      a pattern to match a fully-qualified domain name.
   o 'nameservers': Used to identify a name server information search
      using a pattern to match a host name.
   o 'entities': Used to identify an entity information search using a
      pattern to match a string identifier.

   RDAP search path segments are formed using a concatenation of the
   plural form of the object being searched for, a forward slash
   character ('/', ASCII value 0x002F), and an HTTP query string. The
   HTTP query string is formed using a concatenation of the question
   mark character ('?', ASCII value 0x003F), the JSON object value
   associated with the object being searched for, the equal sign
   character ('=', ASCII value 0x003D), and the search pattern. For the
   domain and entity objects described in this document the plural
   object forms are "domains" and "entities".

3.2.1. Domain Search

   Syntax: domains?name=<domain search pattern>

   Searches for domain information are of the form

   /domains?name=XXXX

   where XXXX is a search pattern representing a domain name in
   "letters, digits, hyphen" format [RFC5890] in a zone administered by
   the server operator of a DNR. The following path would be used to
   find DNR information for domain names matching the "example*.com"
   pattern:

   /domains?name=example*.com

   Internationalized Domain Names (IDNs) in U-label format [RFC5890] can
   also be used as search patterns (see Section 4). Searches for these
   names are of the form /domains?name=XXXX, where XXXX is a search
   pattern representing a domain name in U-label format [RFC5890].

3.2.2. Name Server Search

   Syntax: nameservers?name=<nameserver search pattern>

   Searches for name server information are of the form

    /nameservers?name=XXXX

    where XXXX is a search pattern representing a host name in "letters,
    digits, hyphen" format [RFC5890] in a zone administered by the server
    operator of a DNR. The following path would be used to find DNR
    information for name server names matching the "ns1.example*.com"
    pattern:

    /nameservers?name=ns1.example*.com

    Internationalized name server names in U-label format [RFC5890] can
    also be used as search patterns (see Section 4). Searches for these
    names are of the form /nameservers?name=XXXX, where XXXX is a search
    pattern representing a name server name in U-label format [RFC5890].

3.2.3. Entity Search

    Syntax: entities?fn=<entity name search pattern>

    Syntax: entities?handle=<entity handle search pattern>

    Searches for entity information by name are of the form

    /entities?fn=XXXX

    where XXXX is a search pattern representing an entity name as
    specified in Section 6.1 of [I-D.ietf-weirds-json-response]. The
    following path would be used to find information for entity names
    matching the "Bobby Joe*" pattern.

    /entities?fn=Bobby%20Joe*

    Searches for entity information by handle are of the form

    /entities?handle=XXXX

    where XXXX is a search pattern representing an entity handle as
    specified in Section 6.1 of [I-D.ietf-weirds-json-response]. The
    following path would be used to find information for entity names
    matching the "CID-40*" pattern.

    /entities?handle=CID-40*

    URLs MUST be properly encoded according to the rules of [RFC3986].
    In the example above, "Bobby Joe*" is encoded to "Bobby%20Joe*".

4. Search Processing

Partial string searching uses the asterisk ('*', ASCII value 0x002A)
character to match zero or more trailing characters. Additional
pattern matching processing is beyond the scope of this
specification.

If a server receives a search request but cannot process the request
because it does not support a particular style of partial match
searching, it SHOULD return an HTTP 422 [RFC4918] error. When
returning a 422 error, the server MAY also return an error response
body as specified in Section 7 of [I-D.ietf-weirds-json-response] if
the requested media type is one that is specified in
[I-D.ietf-weirds-using-http].

Because Unicode characters may be combined with another Unicode
character or characters, partial matching is not feasible across
combinations of Unicode characters. Servers SHOULD NOT partially
match combinations of Unicode characters where a Unicode character
may be legally combined with another Unicode character or characters.
Clients should avoid submitting a partial match search of Unicode
characters where a Unicode character may be legally combined with
another Unicode character or characters. Partial match searches with
incomplete combinations of characters where a character must be
combined with another character or characters are invalid. Partial
match searches with characters that may be combined with another
character or characters are to be considered non-combined characters
(that is, if character x may be combined with character y but
character y is not submitted in the search string then character x is
a complete character and no combinations of character x are to be
searched).

Because Unicode characters may be combined with another Unicode
character or characters, partial matching requires that a server
maintain a list of valid character combinations to be considered a
match. When comparing DNS U-labels, servers SHOULD use the code
points specified in [RFC5892] to determine partial matches. When
comparing entity names, servers SHOULD use the normalization rules
and code points specified by [I-D.ietf-precis-nickname] to determine
partial matches.

5. Extensibility

This document describes path segment specifications for a limited
number of objects commonly registered in both RIRs and DNRs. It does
not attempt to describe path segments for all of the objects
registered in all registries. Custom path segments can be created
for objects not specified here using the process described in
Section 6 of "HTTP usage in the Registration Data Access Protocol
(RDAP)" [I-D.ietf-weirds-using-http].

Custom path segments can be created by prefixing the segment with a
unique identifier followed by an underscore character (0x5F). For
example, a custom entity path segment could be created by prefixing
"entity" with "custom_", producing "custom_entity". Servers MUST
return an appropriate failure status code for a request with an
unrecognized path segment.

6. Internationalization Considerations

There is value in supporting the ability to submit either a U-label
(Unicode form of an IDN label) or an A-label (ASCII form of an IDN
label) as a query argument to an RDAP service. Clients capable of
processing non-ASCII characters may prefer a U-label since this is
more visually recognizable and familiar than A-label strings, but
clients using programmatic interfaces might find it easier to submit
and display A-labels if they are unable to input U-labels with their
keyboard configuration. Both query forms are acceptable.

Internationalized domain and name server names can contain character
variants and variant labels as described in RFC 4290 [RFC4290].
Clients that support queries for internationalized domain and name
server names MUST accept service provider responses that describe
variants as specified in "JSON Responses for the Registration Data
Access Protocol" [I-D.ietf-weirds-json-response].

7. IANA Considerations

This document does not specify any IANA actions.

8. Security Considerations

Security services for the operations specified in this document are
described in "Security Services for the Registration Data Access
Protocol" [I-D.ietf-weirds-rdap-sec].

Search functionality typically requires more server resources (such
as memory, CPU cycles, and network bandwidth) when compared to basic
lookup functionality. This increases the risk of server resource
exhaustion and subsequent denial of service due to abuse. This risk
can be mitigated by developing and implementing controls to restrict
search functionality to identified and authorized clients. If those
clients behave badly, their search privileges can be suspended or
revoked. Rate limiting as described in Section 5.5 of "HTTP usage in
the Registration Data Access Protocol (RDAP)"
[I-D.ietf-weirds-using-http] can also be used to control the rate of
received search requests. Server operators can also reduce their
risk by restricting the amount of information returned in response to
a search request.

9. Acknowledgements

This document is derived from original work on RIR query formats developed by Byron J. Ellacott of APNIC, Arturo L. Servin of LACNIC, Kaveh Ranjbar of the RIPE NCC, and Andrew L. Newton of ARIN. Additionally, this document incorporates DNR query formats originally described by Francisco Arias and Steve Sheng of ICANN and Scott Hollenbeck of Verisign.

The authors would like to acknowledge the following individuals for their contributions to this document: Francisco Arias, Marc Blanchet, Ernie Dainow, Jean-Philippe Dionne, Behnam Esfahbod, Edward Lewis, and John Levine.

10. References

10.1. Normative References

[I-D.ietf-precis-nickname]
          Saint-Andre, P., "Preparation and Comparison of
          Nicknames", draft-ietf-precis-nickname-06 (work in
          progress), July 2013.

[I-D.ietf-weirds-json-response]
          Newton, A. and S. Hollenbeck, "JSON Responses for the
          Registration Data Access Protocol (RDAP)", draft-ietf-
          weirds-json-response-05 (work in progress), August 2013.

[I-D.ietf-weirds-rdap-sec]
          Hollenbeck, S. and N. Kong, "Security Services for the
          Registration Data Access Protocol", draft-ietf-weirds-
          rdap-sec-05 (work in progress), August 2013.

[I-D.ietf-weirds-using-http]
          Newton, A., Ellacott, B., and N. Kong, "HTTP usage in the
          Registration Data Access Protocol (RDAP)", draft-ietf-
          weirds-using-http-07 (work in progress), July 2013.

[RFC0952] Harrenstien, K., Stahl, M., and E. Feinler, "DoD Internet
          host table specification", RFC 952, October 1985.

[RFC1123] Braden, R., "Requirements for Internet Hosts - Application
          and Support", STD 3, RFC 1123, October 1989.

[RFC1166] Kirkpatrick, S., Stahl, M., and M. Recker, "Internet
          numbers", RFC 1166, July 1990.

[RFC2119]   Bradner, S., "Key words for use in RFCs to Indicate
            Requirement Levels", BCP 14, RFC 2119, March 1997.

[RFC2616]   Fielding, R., Gettys, J., Mogul, J., Frystyk, H.,
            Masinter, L., Leach, P., and T. Berners-Lee, "Hypertext
            Transfer Protocol -- HTTP/1.1", RFC 2616, June 1999.

[RFC3986]   Berners-Lee, T., Fielding, R., and L. Masinter, "Uniform
            Resource Identifier (URI): Generic Syntax", STD 66, RFC
            3986, January 2005.

[RFC4290]   Klensin, J., "Suggested Practices for Registration of
            Internationalized Domain Names (IDN)", RFC 4290, December
            2005.

[RFC4291]   Hinden, R. and S. Deering, "IP Version 6 Addressing
            Architecture", RFC 4291, February 2006.

[RFC4343]   Eastlake, D., "Domain Name System (DNS) Case Insensitivity
            Clarification", RFC 4343, January 2006.

[RFC4632]   Fuller, V. and T. Li, "Classless Inter-domain Routing
            (CIDR): The Internet Address Assignment and Aggregation
            Plan", BCP 122, RFC 4632, August 2006.

[RFC4918]   Dusseault, L., "HTTP Extensions for Web Distributed
            Authoring and Versioning (WebDAV)", RFC 4918, June 2007.

[RFC5396]   Huston, G. and G. Michaelson, "Textual Representation of
            Autonomous System (AS) Numbers", RFC 5396, December 2008.

[RFC5730]   Hollenbeck, S., "Extensible Provisioning Protocol (EPP)",
            STD 69, RFC 5730, August 2009.

[RFC5733]   Hollenbeck, S., "Extensible Provisioning Protocol (EPP)
            Contact Mapping", STD 69, RFC 5733, August 2009.

[RFC5890]   Klensin, J., "Internationalized Domain Names for
            Applications (IDNA): Definitions and Document Framework",
            RFC 5890, August 2010.

[RFC5891]   Klensin, J., "Internationalized Domain Names in
            Applications (IDNA): Protocol", RFC 5891, August 2010.

[RFC5892]   Faltstrom, P., "The Unicode Code Points and
            Internationalized Domain Names for Applications (IDNA)",
            RFC 5892, August 2010.

   [RFC5952]   Kawamura, S. and M. Kawashima, "A Recommendation for IPv6
               Address Text Representation", RFC 5952, August 2010.

10.2. Informative References

   [REST]   Fielding, R. and R. Taylor, "Principled Design of the
               Modern Web Architecture", ACM Transactions on Internet
               Technology Vol. 2, No. 2, May 2002.

   [RFC3912]   Daigle, L., "WHOIS Protocol Specification", RFC 3912,
               September 2004.

   [RFC4007]   Deering, S., Haberman, B., Jinmei, T., Nordmark, E., and
               B. Zill, "IPv6 Scoped Address Architecture", RFC 4007,
               March 2005.

   [RFC4627]   Crockford, D., "The application/json Media Type for
               JavaScript Object Notation (JSON)", RFC 4627, July 2006.

Appendix A. Change Log

   Initial -00: Adopted as working group document.
   -01: Added "Conventions Used in This Document" section. Added
      normative reference to draft-ietf-weirds-rdap-sec and some
      wrapping text in the Security Considerations section.
   -02: Removed "unified" from the title. Rewrote the last paragraph
      of section 2. Edited the first paragraph of section 3 to more
      clearly note that only one path segment is provided. Added
      "bitmask" to "length" in section 3.1. Changed "lowest IP network"
      to "smallest IP network" in section 3.1. Added "asplain" to the
      description of autonomous system numbers in section 3.2. Minor
      change from "semantics is" to "semantics are" in section 3.2.
      Changed the last sentence in section 4 to more clearly specify
      error response behavior. Added acknowledgements. Added a
      paragraph in the introduction regarding future IETF standards and
      extensibility.
   -03: Changed 'query' to 'lookup' in document title to better
      describe the 'exact match lookup' purpose of this document.
      Included a multitude of minor additions and clarifications
      provided by Marc Blanchet and Jean-Philippe Dionne. Modified the
      domain and name server sections to include support for IDN
      U-labels.
   -04: Updated the domain and name server sections to use .example IDN
      U-labels. Added text to note that mixed IDN labels SHOULD NOT be
      used. Fixed broken sentences in Section 6.
   -05: Added "help" path segment.
   -06: Added search text and removed or edited old search text.

   -07: Fixed query parameter typo by replacing "/?" with "?". Changed
      "asplain" to "AS Plain". Added entity search by handle.
      Corrected section references. Updated IDN search text.

Authors' Addresses

   Andrew Lee Newton
   American Registry for Internet Numbers
   3635 Concorde Parkway
   Chantilly, VA 20151
   US

   Email: andy@arin.net
   URI: http://www.arin.net

   Scott Hollenbeck
   Verisign Labs
   12061 Bluemont Way
   Reston, VA 20190
   US

   Email: shollenbeck@verisign.com
   URI: http://www.verisignlabs.com/