Uniform Resource Locator Schemes for
Internet Relay Chat Entities
<draft-butcher-irc-url-04.txt>

Status of this Memo

Copyright Notice

Abstract

   This document specifies two URL (Uniform Resource Locator) schemes,
   using the URI (Uniform Resource Indicator) names "irc" and "ircs",
   for the location of IRC (Internet Relay Chat) servers. These URLs
   allow for easy location of an IRC server, optionally also specifying
   an IRC channel to join, or a person's nickname to contact upon
   connection.

1. Introduction

   Since its introduction, Internet Relay Chat (IRC) has become widely
   known and used within the Internet Community as a real-time chat
   medium. IRC networks are steadily growing larger, not only with
   regards to the number of regular users, but also the number of
   channels and servers required to support the diverse demand.

   Due to the nature of IRC as a real-time chat service, it has been
   known to be used for a wide variety of uses such as software support,
   job interviews, and of course just for a casual conversation.

   For years now, the need for an appropriate Uniform Resource Locator
   (URL) scheme has been apparent. Applications for such a scheme range
   quite widely, including IRC network's server lists on their website,
   technical support contact details, or even a meeting location within
   an e-mail, giving a specific IRC channel or nickname to contact.

   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 [RFC2119].

   In this document, the term "client" is defined as the IRC client
   software, and the term "user" is the end-user of that software. The
   term "entity" refers to an addressable IRC entity, such as a user,
   service, or channel.

2. URL Definition

   An IRC URL begins with either the Uniform Resource Identifiers (URIs)
   "irc", or "ircs", denoting normal and secured connections
   respectively. Normal sessions are via existing transport (such as
   that in [RFC2812]) and is to be considered insecure. Secured
   sessions are much the same, only secured via a "blanket security"
   method such as [SSL], or negotiating a [TLS] session.

   The URL scheme for IRC follows the Generic URL Syntax ([RFC2396]).

   The action the URL instigates is to open a connection to the
   specified IRC server using whatever protocol necessary, and make
   contact with the given user, service or channel, if also requested.

   There is no requirement for IRC client software implementing the
   "irc" scheme to also implement the "ircs" scheme, or vice-versa. It
   is, however, RECOMMENDED that clients implementing the "ircs" scheme
   are also capable of handling a normal connection via the "irc"
   scheme.

2.1. ABNF Syntax

   Below is the definition for the IRC URL scheme in [ABNF] grammar:

   ircURL   = ircURI "://" location "/" [entity] [flags] [options]

   ircURI   = "irc" / "ircs"
                ; See Section 2, above, for details.

   location = [authinfo "@"] hostport
                ; See Section 3.2.2 of [RFC2396] for the definition
                ; of 'hostport'.

   authinfo = [username] [":" password]
                ; See Section 2.2 of this document for details.

   username = *(escaped / unreserved)

   password = *(escaped / unreserved) [";" passtype]

   passtype = *(escaped / unreserved)

   entity   = ["#"] *(escaped / unreserved)
                ; Note the prefix, "#", may be used for channel names
                ; without escapes. Please see Section 2.5.1.

   flags   = (["," enttype] ["," hosttype])
          /= (["," hosttype] ["," enttype])

   enttype = "," ("isuser" / "ischannel")

   hosttype = "," ("isserver" / "isnetwork")

   options = "?" option *("&" option)

   option   = optname ["=" optvalue]

   optname = *(ALPHA / "-")
                ; Option names are case-insensitive.

   optvalue = optparam *("," optparam)

   optparam = *(escaped / unreserved)

           The definition of "escaped" and "unreserved" is in sections
           2.4.1 and 2.3 of [RFC2396] respectively. Clients MUST be
           aware of protocol limitations. For example, using "IRC-2"
           (RFC2812), it's impossible to use codepoint U+0020 in names.

2.2. Authentication

To allow for complete authentication of a session, a username MAY be
provided with the password. The username MUST NOT be passed to the
server as a nickname. For example, while registering a connection
using the "IRC-2" protocol, the username would be passed as the first
parameter of the "USER" command (See Section 3.1.3 of [RFC2812]).

The characters available for use in a username may be restricted by
the protocol used, and the IRC server software used.

The use of the password field is not recommended, as it presents a
significant security problem. Authors of IRC URLs using the
authentication field, including a password, should make themselves
aware of the security issues discussed in Section 6 of this document.

See Section 3 for examples of username/password pair authentication,
and traditional server password only authentication.

2.3. Server Names

Servers can be named with either their hostname, or address, like
other URL schemes, but also with an IRC Network's name. The
difference can be explicitly specified using the "isserver" and
"isnetwork" keywords in the "hosttype" section (see Section 3 for
examples).

As many modern IRC clients maintain lists of major IRC networks and
their respective servers, determination of a server to connect to
from a given network name should be a trivial task.

If the host name used is not a raw address (such as an IPv4, IPv6, or
other network address), the name cannot be resolved (through DNS or
other means), and does not contain a period character (U+002E), the
client MAY consider the given host name as a network name to find an
appropriate IRC server.

If the IRC client does not contain IRC Network name lists, and
"isnetwork" has been specified, the client MUST NOT attempt to
resolve the network name as a hostname.

2.4. Server Ports

Special consideration must be given to URLs without ports specified.
Almost all IRC servers are contactable on a variety of standard ports
as allocated by the IANA. Should an IRC URL be specified without a

port, a client MAY try a number of standard ports:

- For the "irc" URI, the client SHOULD attempt connection to the
  port 6667, and MAY attempt connection to the ports 194, 6665,
  6666, 6668 and 6669, in that order. Port 194 is likely to be a
  more "authentic" server, however at this time the majority of
  IRC servers are available on port 6667, at least.

- For the "ircs" URI, the default port used is 994. User-space
  ports (those above port 1023) may have questionable
  authenticity, and SHOULD NOT be used unless explicitly
  specified.

Port numbers shown are in decimal, and have been assigned by the
IANA. Section 3.2.2 of [RFC2396] suggests only one port may be used
as a default port, and does not state a preference for or against
port hunting. The act of port hunting for the "irc" scheme when no
port is specified is therefore left up to the discretion of client
authors.

For URL equivalency, clients SHOULD consider default ports without
considering port-hunting. For example, <irc://some.server/> and
<irc://some.server:6667/> should be considered equivalent, as should
<ircs://some.server/> and <ircs://some.server:994/>.

Note that the port 194 is officially the "standard" port for IRC
servers, the current practise is to use port 6667. This document may
be updated in the future if and when port 194 obtains an increased
prevailance.

2.5. Entity Names

Only one entity can be named per URL. The named entity SHOULD be
presumed to be a channel name, unless the "enttype" section (see
Section 2.1) of the URL is provided to determine the entity type.

An automated message MUST NOT be sent to the addressed entity.

2.5.1. Channel Names

When "enttype" contains "ischannel", or "enttype" is omitted
completely, the entity name provided is a channel name.

While it is discouraged, channel names prefixed with the "#" (U+0023)
character may be specified without encoding the character (as "%23")
in the URL. Implementers MAY accept this, despite it being an

exception to Section 2.4.3 of [RFC2396], because channels of this
type are currently very common, and will remain so in the foreseeable
future.

Clients SHOULD attempt to determine valid channel name prefix
characters from the server it has connected to, such as via an
"RPL_ISUPPORT" reply. If the client is unable to determine valid
prefix characters for the server it is connected to, the client
SHOULD attempt to join the channel without modifying its name. If
joining the channel failed, the prefix character "#" may be used.

If the client discovers the channel name given is considered to be
invalid because it is missing a valid prefix character, the client
SHOULD prepend a default prefix character to the name.

Since default prefix characters for channels may differ between IRC
servers, the client SHOULD try to determine the default channel
prefix for the server it is connected to, such as the first prefix
character given by "CHANTYPES" in "RPL_ISUPPORT". If the client is
still unable to determine a prefix character, a prefix character of
'#' (U+0023) MAY be presumed.

2.5.2. Nicknames

When "enttype" contains "isuser", the entity given refers to a user.
The given entity name may simply be a nickname, or it may contain
more specific information such as the user's hostname, username, or a
server they use.

A user entity is referred to using the following syntax (in [ABNF]
grammar):

    userent = nickname ["%21" username] ["%40" hostname]

The definitions of "nickname", "username", and "hostname" are all
identical to the definition of "entname", as defined in Section 2.1
of this document.

It's RECOMMENDED that the client parse this name, as most servers
will not accept this syntax directly. For example, the client may
wish to make use of the IRC-2 "WHO" command to discover if the entity
is valid and available.

2.6. Additional Options

   Additional options may be used to provide additional information
   about the entity you're addressing.

   These options listed here may be expanded on at a later date by
   future documents. Unsupported options MUST be ignored by the client.
   The client author is not obligated to utilise the "options" section
   (see Section 2.1) of the URL, but it is RECOMMENDED to do so.

2.6.1. The "key" Option

   This option is only valid if the entity name given is a channel name.
   If the entity name is not a channel name, then this option MUST
   simply be ignored.

   The option's value provides a "key" to be given to the server when
   joining the given channel name, and is used for channels which
   require a "key" to join them. If a channel key is found to be
   required and one is not provided with this option, the IRC client may
   wish to prompt the user for the key.

   Please see Section 6 of this document.

3. Examples

   While examples of every situation cannot be shown here because of
   space considerations, the following examples provide a rough overview
   of how the IRC URL can be used.

      <irc://irc.undernet.org/>

   In its simplest form, the above complete URL can be used to direct a
   client to a specific IRC server, which in this case is
   "irc.undernet.org". The client should presume to use default port
   settings.

      <irc://irc.ircnet.net/#worldchat,ischannel>
      <irc://irc.ircnet.net/%23worldchat,ischannel>
      <irc://irc.ircnet.net/#worldchat>
      <irc://irc.ircnet.net:6667/%23worldchat>

   All four of these URLs connects to the IRCnet network, and will join
   the client to the channel "#worldchat" upon connection. All of these
   URLs are considered identical.

   <irc://irc.alien.net.au/pickle,isuser>

This will connect to the server "irc.alien.net.au" and will provoke
the client to open up a window (or similar) associated with sending
messages to the nickname 'pickle'.

   <irc://irc.austnet.org/%23foobar?key=bazqux>

This will connect to AUSTnet and join the channel "#foobar", using
the channel key "bazqux".

   <irc://undernet/pickle%25butcher.id.au,isuser>

This will open a dialogue box prepared to send a message to "pickle"
with the server name "butcher.id.au". This URL will connect to the
network named as "undernet". For this to work correctly, the client
must be configured appropriately to know of at least one server's
address associated with this name.

   <irc://:pass@irc.efnet.org:194/>

The above URL specifies that the IRC client should try to connect to
"irc.efnet.org" on the port 194, rather than use the default port(s).
It also tells the IRC client it should try to connect to the server
using the server password "pass".

   <irc://%C4%B0dil:g%C3%BCzel@irc.austnet.org/>

This shows a properly [UTF-8] encoded URL, specifying the username
"Idil" (with the first character being a Turkish Latin capital letter
"I" with a dot above it, [Unicode] codepoint U+0130) and the password
"guzel" (with a diaeresis on the u, codepoint U+00FC).

4. **Internationalisation Considerations**

   With the inevitable adoption of [Unicode] on IRC, and indeed the
   Internet as a whole, URLs MUST be encoded using the [UTF-8] character
   set, with (potentially) unsafe octets encoded using %HH notation
   (where HH is a hexadecimal value), as per Section 2.2.5 of
   [RFC-2718]. An example of this in action can be found in Section 3.

   Some IRC servers use such character sets as US-ASCII and KOI-8. It
   is left up to the client and the server to negotiate an appropriate
   character set for communication between the two, as more servers are
   now implementing specific character-set preferences. It is also left
   up to the client to convert entity names from UTF-8 into the
   appropriate character set.

   At the time of writing, [UTF-8] is set to become the popular choice
   (announced via RPL_ISUPPORT) as it's easy to implement with very
   minimal changes to existing server software. Other IRC servers are
   opting to announce a preferred character set, but allow the client to
   switch character sets on the fly, using CAP/CAPAB negotiation, oft
   implemented using the UNIX98 iconv() function (or something similar).

5. Interoperability Considerations

   Many existing implementations fail to acknowledge the correct use of
   the generic URL syntax defined in [RFC2396], but act like they use
   the format.

   Some current implementations will need slight modification to accept
   the extended format defined in this specification, however most
   implementations which parse the URL in a standard form will continue
   to work for most IRC URLs.

   The presumption of a channel name without explicitly specifying the
   entity type is designed to maintain compatibility with the existing
   implementations. The practise of omitting the channel prefix
   character, or not encoding it, is also for compatibility, but is
   STRONGLY DISCOURAGED.

   There are interoperability issues with existing IRC servers as a
   result of the restricted characters available for channel names and
   nicknames. The restriction of acceptable characters has been left up
   to the IRC server authors and not the URL scheme, as not to hinder
   advances in IRC protocols and servers.

   Some existing IRC servers will accept nickname/password pairs,
   however at the time of writing these servers do not use this for
   actually authenticating the session, but instead identifying
   nicknames to nickname registration services. The use of
   username/password pairs is used for actual authentication, and has
   been included.

[6](#). **Security Considerations**

   Security problems naturally arise when a server password and/or a
   channel key is specified (using the "key" option). While the use of
   the password and channel key sections is considered to be rare, and
   they have been included for uses such as for shortcut/bookmark lists,
   or to be used as a user command.

   As the passwords and channel keys are unfortunately passed as clear
   text, any user using the IRC URL should be aware of obvious
   insecurities. It is strongly discouraged to use these fields in a
   public sense, such as on a website.

   Furthermore, it is recommended that client software does not
   automatically initiate the connection specified by the URL without
   the knowledge and consent of the user. To do so would open the
   implementation up to a variety of malicious activities including, but
   not limited to, the purposes of direct advertising or channel
   advertising (known as "spam") via "pop-ups" or other means.

   When connecting using a secure connection ("ircs://"), user-space
   ports (those above port 1023) should be treated with suspicion, as
   their authority could be questionable. If a secure connection cannot
   be established, the client MUST NOT automatically default to an
   insecure ("irc://") connection. To do so would denigrate the "ircs"
   scheme and restrict its usefulness.

   Automated messages MUST NOT be sent to any entity upon connection to
   an IRC server as a direct result of execution of an IRC URL. Sending
   messages to channels and other users should be left up to the user,
   not the URL author or the client software. The facility to send
   automated messages to other users has been explicitly avoided in this
   document to avoid abuse, common with IRC.

   Clients MUST be aware of protocol limitations, especially when
   dealing with entity names, as the probability for exploitation is
   high. For example, a URL with a nickname including "%0D%0A" could be
   used to exploit a client using using the "IRC-2" protocol,
   potentially allowing a malicious URL author to execute any command
   they wish.

   Beyond this, there are security concerns with regards with associated
   protocols, including the IRC server-to-user protocols themselves,
   [TLS] and [UTF-8], which must be taken into consideration, but are
   beyond the scope of this document.

7. **IANA Considerations**

   The following is registration for the URL schemes as per [RFC2717]:

      URL scheme name: Two URI's are described herein: "irc" and "ircs".

      URL scheme syntax: See Section 2.1, and indeed Section 2 as a
      whole.

      Character encoding considerations: Characters must be encoded in
      UTF-8 and escaped. See Section 4.

      Intended usage: The scheme initiates connection to an IRC server,
      normally through the execution of IRC Client software. Further-
      more, the scheme may then initiate further commands, such as
      joining channels, as outlined above.

      Interoperability considerations: See Section 5.

      Security considerations: See Section 6.

      Relevant publications: The IRC protocol is defined by [RFC2812].
      Either [SSL] or [TLS] may be used for the "ircs" scheme, depending
      on client and server configuration.

      Person & email address to contact for further information: The
      Author; See Section 10 for details.

      Author/Change controller: The Author's details are contained
      within Section 10. The IETF is to maintain change control.

8. **Acknowledgments**

   I acknowledge the previous work of Mandar Mirashi who originally
   wrote an Internet-Draft to similar effect.

   The input of Petr Baudis, Robert Ginda, Piotr Kucharski, Perry
   Lorier, Khaled Mardam-Bey, Dominick Meglio, James Ross, and Samuel
   Sieb, was greatly appreciated, and this draft would not exist without
   their valued participation. I also thank them for their patience
   while I was travelling overseas.

   I would also like to acknowledge those members of the IRC development
   community who encouraged me to publish this document, after more than
   18 months of pretermission.

9.   References

    [ABNF]   Crocker, D., and Overell, P., "Augmented BNF for Syntax
               Specifications: ABNF", RFC 2234, November 1997.

    [PICS]   Miller, J., Resnick, P., Singer, D., "Rating Services and
               Rating Systems (and Their Machine Readable Descriptions)",
               Version 1.1, <http://www.w3.org/TR/REC-PICS-services>,
               October 1996.

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

    [RFC2368] Hoffman, P., Masinter, L., Zawinski, J., "The mailto URL
               scheme", RFC 2368, July 1998.

    [RFC2396] Berners-Lee, T, Fielding, T., Masinter, L., "Uniform
               Resource Identifiers (URI): Generic Syntax", RFC 2396,
               August 1998.

    [RFC2717] Petke, R., King, I., "Registration Procedures for URL
               Scheme Names", RFC 2717, November 1999.

    [RFC2718] Masinter, L., Alvestrand, H., Zigmond, D., Petke, R.,
               "Guidelines for new URL Schemes", RFC 2718, November 1999.

    [RFC2811] Kalt, C., "Internet Relay Chat: Channel Management", RFC
               2811, April 2000.

    [RFC2812] Kalt, C., "Internet Relay Chat: Client Protocol", RFC 2812,
               April 2000.

    [SSL]   Hickman, K., "The SSL Protocol", Netscape Communications
               Corp., February 9, 1995.

    [TLS]   Dierks, T. and Allen, C., "The TLS Protocol Version 1.0",
               RFC 2246, January 1999.

    [Unicode] The Unicode Consortium. The Unicode Standard, Version
               4.0.0, (Reading, MA, Addison-Wesley, 2003. ISBN
               0-321-18578-1).

    [UTF-8]   Yergeau, F., "UTF-8, a transformation format of ISO 10646",
               RFC 2279, January 1998.

10. **Author's Address**

Simon Butcher
Alien Internet Services

PO Box 7041
Croydon South
VIC 3136
Australia

Phone: +61-3-9879-8052
Fax: +61-3-9893-2793
Email: simonb@alien.net.au
       simon@butcher.name
       simon@butcher.id.au

11. **Intellectual Property Rights Notice**

12. **Full Copyright Notice**

Document Discussion Forum

   Discussion for this document is currently taking place using the
   "irc-url@alien.net.au" mailing list in order to keep discussion open
   and archived appropriately. Those interested in this document are
   strongly encouraged to join in on the discussion.

   To subscribe to the mailing list, see:
      <http://alien.net.au/mailman/listinfo/irc-url_alien.net.au>

   Archives of the mailing list are available at:
      <http://alien.net.au/pipermail/irc-url_alien.net.au>

Document Expiration and Filename

   Please note that this is a draft document and it shall expire July
   2004. Its filename is <draft-butcher-irc-url-04.txt>.