Uniform Resource Locator Schemes for Internet Relay Chat Servers
<draft-butcher-irc-url-01.txt>

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 person 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 uses, but also the number of
channels and servers required to support the demand.

Due to the nature of IRC as a simple 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 chat.

While IRC is progressing, the need for an appropriate Uniform
Resource Locator (URL) scheme has become apparent. Applications for
such a scheme would range quite widely, including IRC network server
lists on a website, software support contact details, or even a
meeting location with an e-mail including a specific IRC channel.

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.

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 the existing transport, as
defined in [RFC2812], and secured sessions are the same, only via a
secure transport layer such as [TLS] (or [SSL], the predecessor to
TLS).

The URL scheme for IRC follows the Generic URL Syntax, defined in
[RFC2396].

The action the URL is to instigate is to open a connection to the
specified IRC server using whatever protocol necessary. Currently
only one protocol is defined to do this, as per RFC 2812.
2.1. ABNF Syntax

The following is the definition for an IRC URL in [ABNF] grammar:

ircURL   = type "://" location "/" [ channel ] [ options ]

  type     = "irc" / "ircs"

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

authinfo = [ nickname *2( "," nickname ) ] [ ":" password ]

nickname = *( escaped / unreserved )
 ; Further restrictions may apply upon
 ; connection, depending on the server.
 ; Some common nickname characters must be
 ; encoded, as per recommendations in Section
 ; 2.4.3 in [RFC2396].

password = *( escaped / unreserved )

channel = *( escaped / unreserved / safe / "=" )

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

option = optname [ ":=" optvalue ]

optname  = *( ALPHA )

optvalue = optparam *( "," optparam )
 ; Only some options have multiple parameters.

optparam = *( escaped / unreserved / safe / "?” )
 ; Other characters (such as [Unicode]
 ; characters in [UTF-8]) must be escaped
 ; as per rules in Section 2.2.5 of [RFC2718].
 ; See also Section 4 of this document.

  safe     = ";" / "@" / ";" / ";" / ";" / ";" / ";" / ";" / ";" / ";" / ";"

The definition of "escaped" and "unreserved" is in sections 2.4.1 and
2.3 of [RFC2396] respectively.

The use of spaces, escaped (%20) or otherwise, SHOULD NOT be allowed
for channel names or nicknames, as they are currently unsupported by
existing IRC protocols.
2.2.  Authentication

Multiple nicknames MAY be specified, and pending any error or lack of availability, the IRC client software MAY request the next nickname in the list.

Clients MUST NOT accept any more than three (3) nicknames, and any more nicknames specified MUST be ignored to curtail potential abuse.

Clients may only attempt to use alternative nicknames given during the one connection. Clients MUST NOT reconnect to the server to try alternative nicknames.

Should the client run out of alternative nicknames to try, but the server is willing to accept another attempt, the client MAY either disconnect from the server and show the user an error message, or prompt the user for another nickname to try.

The use of passwords is not recommended, as they present a significant security problem. They are allowed purely for convenience. Users of the password field must be aware of the security issues discussed in Section 6 of this document.

2.3.  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 SHOULD try a number of standard ports:

- For the "irc" URI, it is RECOMMENDED that the server is attempts connection to the ports 194, 6665, 6666, 6667, 6668 and 6669.

- For the "ircs" URI, it is RECOMMENDED that 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 here are in decimal, and have been assigned by the IANA.
2.4. Channels

For compatibility with older implementations, and to allow simplification of the specification, channels MAY be specified without the use of the "channel" option detailed in Section 2.5.1.

Only one channel can be specified, and this facility has the same functionality as the "channel" option. See Section 3 of this document for examples of the equivalence between this and the "channel" option in Section 2.5.1.

2.5. Additional Options

Additional options may be added to control what action the client software MAY take following connection to the IRC server. Unsupported options should simply be ignored.

These options listed here may be expanded on at a later date by updated RFC’s.

2.5.1. The "channel" Option

This instructs the client to join the specified channel, allowing the user to participate in discussions within the channel.

The value given with the channel option is a channel name, and optionally a "key" (see Section 4.2.10 of [RFC2811]). Its value can be defined in [ABNF] as follows:

\[
\text{value} = \text{name } [ \text{","} \text{key} ]
\]

See Section 2.1 of this document for information on acceptable characters. The ",," character is not allowed in either the channel name, or key, unless it is escaped ("%x2C").

The number of channels which can be joined at once is normally restricted by the server, but no hard-limit is given by this specification as this is a matter of individual server policy. As such, multiple "channel" options may be given.

An automated message MUST NOT be sent to the channel upon joining it.

It is NOT RECOMMENDED to use the channel key feature. Please see Section 6 of this document. If a key is required to join a channel, and one is not given, the IRC client MAY wish to prompt the user for the key.
2.5.2. The "query" Option

For each "query" option, the client is requested to open some interface where by users may type a message to the given query target. For example, the client may open up a window where messages may be typed to, and received from the target.

Some clients may not have the ability to open up a specific window or dialogue box. These clients MAY prompt the user for a message to be sent to the target, or otherwise this option MAY be ignored and hence unsupported.

The option value is the same as the message target value specified in Section 3.3.1 of [RFC2812], except that the client MUST only accept one target. Multiple targets per "query" option MUST NOT be accepted, and the entire query MUST be considered invalid and ignored.

A message MUST NOT be automatically sent to the target, simply an interface created to allow the user to send a message to the target.

The IRC client software MAY wish to check the availability of the target prior to opening the interface if inclined to do so, however any method of testing the availability MUST NOT generate any automatic message being sent to the target.

Multiple targets MAY be referenced with multiple query options, however in order to reduce the potential for abuse, it is RECOMMENDED that additional query options are ignored. There are valid reasons for having multiple targets, and abuse is minimal as no messages are sent to the targets.

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://pickle@irc.efnet.org:6667/>
The above URL specifies that the IRC client should try to connect to "irc.efnet.org" on the port 6667, rather than use whatever is considered the default. It also tells the IRC client it should try to use a nickname of "pickle", if it is available.

<irc://%C4%B0dil,idil@irc.austnet.org/>

This shows a properly [UTF-8] encoded URL, specifying the nickname Idil (with the first character being a Turkish Latin capital letter "I" with a dot above it, [Unicode] code point U+0130). Failing that, the second nickname, "idil", may be used if the first one is rejected, perhaps by an older IRC server.

<irc://pickle:secret@192.0.2.1:194/>

The above URL will instruct the IRC client to connect to a server with the address 192.0.2.1, which is an IRC server that is presumably password protected. The client should request to use the nickname "pickle", with the password "secret" to authenticate the session to the remote server. This URL also enforces the standard IRC port, 194, and will stop IRC clients from hunting for ports.

<irc://irc.ircnet.net/#worldchat>
<irc://irc.ircnet.net/%23worldchat>
<irc://irc.ircnet.net/?channel=%23worldchat>

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

<irc://irc.alien.net.au/?query=pickle&channel=%2Bprivate,foo>

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’. It will also join the channel "+private" using the channel key "foo".

<irc://irc.austnet.org/%23melbourne?channel=%23sydney>

This will connect to AUSTnet and join two channels, "#melbourne" and "#sydney".

<irc://irc.undernet.org/?query=pickle%25butcher.id.au>

This will open a dialogue box prepared to send a message to "pickle%butcher.id.au". Please refer to Section 3.3.1 of [RFC2812] for more details.
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].

For example, the word for "alias" in Japanese, using Unicode code-points, is U+5225 U+540D. Correctly encoded, this would appear in the URL as: "%E5%88%A5%E5%90%8D". An example of this use in action can be found in Section 3.

5. Interoperability Considerations

Many existing implementations fail to acknowledge the correct use of the generic URL syntax defined in [RFC2396], but continue to use the format regardless. This implementation flaw is likely to be due to the first documentation of the irc: URI scheme by the W3C’s [PICS] recommendation. This implementation has never adequately considered the needs of IRC, nor even the implementation of IRC at the time.

Some implementations do not take into consideration the use of prefix characters on channels names, as defined by [RFC2811], and therefore cannot correctly reference channels appropriately. Software SHOULD NOT guess the channel type, or hence the channel’s prefix character, unless the probability of an accurate guess is 100% (for example, the server only supports one type of channel).

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.

The majority of incongruities are simply caused by the problem of developers ignoring RFC-2396.

The use of the channel name without specifying the channel option is to both maintain compatibility with the existing implementations, whilst providing an abbreviated form, similar to the design of the "mailto:" scheme defined by [RFC2368].

Some fields have been extended to allow additional characters outside of those normally needing to be encoded to allow for interoperability with existing implementations.
6. Security Considerations

Security problems arise only when the authentication portion of the URL is used, or channel keys are given. While the use of the password/key extensions is considered to be rare, they have been included for completeness.

As the passwords and channel keys are unfortunately in clear-text, any user using the IRC URL should be aware of obvious insecurities.

Furthermore, it is recommended that user 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 (also known as "spam") by way of pop-ups.

When connecting using a secure connection ("ircs://"), user-space ports (those above port 1023) should not be used automatically, as their authority is questionable. If a secure connection cannot be established, the client MUST either give up, or prompt the user before attempting an insecure ("irc://") connection.

Automated messages MUST NOT be sent to channels or other users upon connection to an IRC server as a direct action of this URL. Services MAY be contacted. The facility to send automated messages to other users has been explicitly avoided in this document to avoid abuse, common with IRC services.

Beyond this, there are security concerns with regards with associated protocols, including IRC and TLS, 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: "irc" and "ircs".

URL scheme syntax: See Section 2.1.

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.

Interoperability considerations: See Section 5.


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 IETF is to maintain change control.

8. Acknowledgments

Thanks must go to Khaled Mardam-Bey for his early implementation in his software, "mIRC", which provided the inspiration to clarify the specification.

I acknowledge the previous work of Mandar Mirashi who originally wrote an Internet-Draft to similar effect, but of which this document has no derivation.

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


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.id.au

Intellectual Property Rights Notice

The IETF takes no position regarding the validity or scope of any
intellectual property or other rights that might be claimed to
pertain to the implementation or use of the technology described in
this document or the extent to which any license under such rights
might or might not be available; neither does it represent that it
has made any effort to identify any such rights. Information on the
IETF’s procedures with respect to rights in standards-track and
standards-related documentation can be found in BCP-11. Copies of
claims of rights made available for publication and any assurances of
licenses to be made available, or the result of an attempt made to
obtain a general license or permission for the use of such
proprietary rights by implementers or users of this specification can
be obtained from the IETF Secretariat.

The IETF invites any interested party to bring to its attention any
copyrights, patents or patent applications, or other proprietary
rights which may cover technology that may be required to practice
this standard. Please address the information to the IETF Executive
Director.
Full Copyright Notice

Copyright (C) The Internet Society (2003). All Rights Reserved.

This document and translations of it may be copied and furnished to others, and derivative works that comment on or otherwise explain it or assist in its implementation may be prepared, copied, published and distributed, in whole or in part, without restriction of any kind, provided that the above copyright notice and this paragraph are included on all such copies and derivative works. However, this document itself may not be modified in any way, such as by removing the copyright notice or references to the Internet Society or other Internet organizations, except as needed for the purpose of developing Internet standards in which case the procedures for copyrights defined in the Internet Standards process must be followed, or as required to translate it into languages other than English.

The limited permissions granted above are perpetual and will not be revoked by the Internet Society or its successors or assigns.

This document and the information contained herein is provided on an "AS IS" basis and THE INTERNET SOCIETY AND THE INTERNET ENGINEERING TASK FORCE DISCLAIMS ALL WARRANTIES, EXPRESS OR IMPLIED, INCLUDING BUT NOT LIMITED TO ANY WARRANTY THAT THE USE OF THE INFORMATION HEREIN WILL NOT INFRINGE ANY RIGHTS OR ANY IMPLIED WARRANTIES OF MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE.

Document Expiration and Filename

Please note that this is a draft document and it shall expire April 2004. Its filename is draft-butcher-irc-url-01.txt