







Why a URL API The CURLU handle

Set

Get CURLOPT_CURLU URLs are still messy

Why a URL API?

Lots of libcurl-using applications need to parse and handle URLs

URL parsing is messy and lacks standards Allows for different interpretations - causing security issues Unified application URL handling and libcurl URL handling Popular request in the annual user survey

The CURLU handle

Create a CURLU handle with curl_url()

Delete it with curl_url_cleanup()

Make a clone with curl_url_dup()



"Populate" the handle with URL details using curl_url_set() curl_url_set(url, CURLUPART_URL, "https://example.com", 0); curl_url_set(url, CURLUPART_SCHEME, "http", 0); curl_url_set(url, CURLUPART_PATH, "/index.html", 0);

Set - options

CURLU_URLENCODE – encode the given string

CURLU_NON_SUPPORT_SCHEME - allow a non-supported scheme to be set

CURLU_GUESS_SCHEME - guess scheme based on host name



Extract the full URL or a URL part from a CURLU handle curl_url_get(url, CURLUPART_URL, &urlptr, 0); Returns error if not enough parts have contents. curl_url_get(url, CURLUPART_SCHEME, &schemeptr, 0); Always curl_free() the returned pointer after use!

Get - options

CURLU_URLDECODE – URL decode the string before returned

CURLU_NO_DEFAULT_PORT – Don't include the port number in the returned URL if it matches the default port for the scheme.

CURLOPT_CURLU

Pass in an already parsed URL to libcurl

Alternative to CURLOPT_URL

URLs are still messy

curl speaks RFC 3986-with-tweaks

Browsers speak WHATWG URL Spec

Other tools speak either, a mix or something else

IDN doesn't make things easier

