MORETP

DAN BRUMLEVE

1. OVERVIEW

moretp is a simple text protocol that allows embedding of binary data.

The TCP byte stream is divided into packets. A packet is a line terminated by '\n' followed by an optional binary part, representing a list of strings. When decoding a packet, lines are processed to remove leading and trailing whitespace and all other whitespace sequences are normalized to a single ' '. After this processing, each line corresponds uniquely to a sequence of non-empty, non-whitespace-containing words.

If any word in a line begins with '<', the rest of the word is passed to strtoul(3) to get a byte count. The total byte count of the words in the line is the length of the binary part of the packet. To complete decoding of the packet, each word beginning with '<' is replaced by the corresponding bytes in the packet's binary part. Then the resulting word list is the decoded packet.

To encode a moretp packet, start with a list of a strings. An empty list corresponds to the packet "\n". If none of the strings contain any whitespace, then the packet is the list joined by ' ' and terminated by '\n'. If any string contains whitespace, replace it with '<' followed by its length prior to the joining, and append the original string to the final packet. Or, do this for every string whether or not it contains whitespace.

Example 1. A moretp encoder in perl.

```
sub moretpenc {
  join(' ', map /\s/ ? "<" . length : $_, @_) .
  "\n" .
  join('', grep /\s/, @_)
}
```

2. HTTP OVER MORETP

As long as the HTTP request line and request headers are not sensitive to multiple whitespace, and do not contain any '<' characters, each line can be understood as a moretp packet.

Example 2. This is both moretp and HTTP.

```
GET / HTTP/1.1
Host: localhost:9999
Connection: keep-alive
Upgrade-Insecure-Requests: 1
User-Agent: Mozilla/5.0 (X11; Linux x86_64) AppleWebKit/537.36 (KHTML, like Gecko) Chrome/72.
Accept: text/html,application/xhtml+xml,application/xml;q=0.9,image/webp,image/apng,*/*;q=0.8
Accept-Encoding: gzip, deflate, br
Accept-Language: en-US,en;q=0.9
```

Example 3. And so is this, including the request body.

```
POST /test HTTP/1.1
Host: localhost:9999
Connection: keep-alive
Content-Length: 15
Cache-Control: max-age=0
Upgrade-Insecure-Requests: 1
Origin: null
Content-Type: application/x-www-form-urlencoded
User-Agent: Mozilla/5.0 (X11; Linux x86_64) AppleWebKit/537.36 (KHTML, like Gecko) Chrome/72.
Accept: text/html,application/xhtml+xml,application/xml;q=0.9,image/webp,image/apng,*/*;q=0.8
Accept-Encoding: gzip, deflate, br
Accept-Language: en-US,en;q=0.9

foo=bar&foo=baz
```

```
To handle binary or whitespace-containing request and response bodies, we can use the trick o
```

Example 4. This is also valid moretp.

```
POST /test HTTP/1.1
Host: localhost:9999
Content-Length: 21
Content-Type: text/plain
X-Moretp-Length: <23

extra   space   between
```
In this way a moretp server can communicate with HTTP clients without any
special server-side support.