```perl
#!/usr/bin/perl

# MicroWeb - ein minimalistischer Webserver zu Lehrzwecken
# (C) 2006 Veit Wahlich

package MicroWeb;
our $VERSION='1.0';


# Benoetigte Module und Pragmas importieren:
use strict;             # Wir programmieren stets strict
use warnings;           # und haetten gerne Warnmeldungen.
use IO::Socket::INET;   # Wir moechten IP-Sockets
use IO::File;           # und Dateien verwenden.
use POSIX qw(strftime); # Fuer saubere Timestamps benutzen wir strftime().


# globale Variable, speichert die Anzahl der aktuell laufenden Kindprozesse
my $children=0;

# globale Konfiguration
my $conf={
    bind_address       => '0.0.0.0',        # IP-Adresse, an die wir binden.
    bind_port          => 8080,             # TCP-Port, an den wir binden.
    root               => 'htdocs/default', # Basisverzeichnis der Dokumente.
    index_file         => 'index.html',     # Datei zu laden wenn nur ein
                                            #   Verzeichnis angefragt wurde.
    max_connections    => 10,               # Max. Anzahl paralleler
                                            #   Verbindungen (== Kindprozesse).
    max_url_length     => 4096,             # Max. Laenge einer URL in Bytes.
    max_headers_length => 4096,             # Max. Groesse der Headers in Bytes.
    client_timeout     => 5,                # Max. Wartezeit fuer Empfang von
                                            #   Anfragen.
    mime_types => {                         # Relationen Dateierweiterung zu
                                            #   MIME-Typ.
        html  => 'text/html',
        htm   => 'text/html',
        txt   => 'text/plain',
        css   => 'text/css',
        xml   => 'text/xml',
        xsl   => 'text/xml',
        jpg   => 'image/jpeg',
        jpeg  => 'image/jpeg',
        png   => 'image/png',
        gif   => 'image/gif',
        mp3   => 'audio/mpeg',
        wav   => 'audio/x-wav',
        mid   => 'audio/midi',
        mpg   => 'video/mpeg',
        mpeg  => 'video/mpeg',
        avi   => 'video/x-msvideo',
        ogg   => 'application/ogg',
        js    => 'application/x-javascript',
        swf   => 'application/x-shockwave-flash',
        gz    => 'application/x-gzip',
        tgz   => 'application/x-gzip',
        bz2   => 'application/x-bzip2',
        tbz2  => 'application/x-bzip2',
        tar   => 'application/x-tar',
        zip   => 'application/zip',
        '*'   => 'application/octet-stream'
    },
};


# Variablen wegen Verwendung in Signalhandlern in globalem Scope:
my $socket;             # Enthaelt spaeter den Listener.
my $client;             # Enthaelt spaeter das Filehandle des aktuellen Clients.


sub main(){

  my $pid;              # Enthaelt spaeter die PID des Client-Kindprozesses.

  # Ein Signalhandler faengt SIGTERM und SIGINT ab:
  $SIG{TERM}=$SIG{INT}=sub{
      # Listener sauber beenden:
      logError('Server received SIGTERM/SIGINT - exiting gracefully');
      $socket->shutdown(2);
      $socket->close();
      STDOUT->close();
      STDERR->close();
      exit(0);
    };

  # Bei SIGCHLD wurde ein Kindprozess beendet und muss geerntet werden:
  $SIG{CHLD}=\&reapPid;

  # Listener anlegen...
  $socket=new IO::Socket::INET(
      Listen    => 5,
      LocalAddr => $conf->{bind_address},
      LocalPort => $conf->{bind_port},
      Proto     => 'tcp',
      Reuse     => 1
    );

  # und auf Erfolg ueberpruefen.
  unless(defined($socket)){
    die("Unable to bind port to address: $!\n");

  }

  # Server-Endlosschleife betreten
  while(1){

    # Auf neue Client-Verbindung warten und in $client speichern:
    $client=$socket->accept();

    # Moderne Betriebssysteme geben ein $client==undef zurueck, wenn $socket
    # noch nicht wieder bereit ist - dann die Schleife von vorn beginnen:
    # *FIXME*: Hier sollte man eigentlich kurz warten, sonst 100% CPU-Last,
    # bis Socket wieder Verbindungen annimmt!
    next unless(defined($client));

    # Nur neue Verbindungen verarbeiten, wenn nicht zu viele Verbindungen
    # aktiv sind:
    if($children < $conf->{max_connections}){

      # Prozess duplizieren und PID speichern.
      $pid=fork();

      # Wenn fork() erfolgreich ist, ist $pid definiert:
      if(defined($pid)){

        # Und wenn $pid == 0 ist, sind wir gerade im Kindprozess - sonst sind
        # wir im Mutterprozess.
        if($pid == 0){
```

```
129
130             # Im Client-Kindprozess brauchen wir den Listener nicht.
131             $socket->close();
132
133             # Signalhandler im Client sind anders als die im Server:
134             $SIG{TERM}=$SIG{INT}=sub{
135                 logError('Child '.$$
136                   .' received SIGTERM/SIGINT - exiting gracefully');
137                 exitChild();
138               };
139
140             # Uebergebe Verarbeitung der Verbindung an die Zulieferer-Funktion:
141             processConnection($client);
142
143             # Ist die Verbindung verarbeitet, kann die Client-Verbindung
144             # geschlossen und der Prozess beendet werden:
145             exitChild();
146
147           }
148           # Wenn wir im Mutterprozess sind:
149           else{
150
151             # Anzahl von Kindprozessen inkrementieren:
152             $children++;
153
154             # Im Mutterprozess benoetigen wir die Client-Verbindung nicht.
155             $client->close();
156
157           }
158
159         }
160         # Wenn fork() fehlschlug, Meldung ausgeben:
161         else{
162           logError('Unable to fork: '.$!);
163         }
164
165       }
166       # Wenn die maximale Anzahl von Verbindungen ueberschritten wurde, geben
167       # wir einfach eine Meldung an den Client, ohne die Verbindung wirklich zu
168       # verarbeiten, und schliessen die Verbindung:
169       else{
170         logError('Maximum count of children reached!');
171         sendError($client,503,'Service Unavailable',
172           'Too many concurrent connections. Please try again later.');
173         $client->shutdown(2);
174         $client->close();
175       }
176
177     }
178
179   }
180
181
182   # Verarbeite eine HTTP-Client-Verbindung:
183   sub processConnection($){
184     my($client)=@_;
185
186     my $getString;        # Enthaelt spaeter den GET-Parameter-String.
187     my $file;             # Enthaelt spaeter den Pfad zur angeforderten Datei.
188     my $host;             # Enthaelt spaeter den Hostname (Host:-Header).
189
190     # Definiere einen SIGALRM-Handler:
191     $SIG{ALRM}=sub{
192         # Timeout-Meldung senden, Client-Verbindung etc. schliessen und
193         # Kindprozess beenden:
194         sendError($client,408,'Request Time-Out',
195           'The server did not receive a propper request within '
196           .$conf->{client_timeout}.' seconds.');
197         exitChild();
198       };
199
200     # Setze ein Timeout fuer den Empfang der Header vom Client:
201     alarm($conf->{client_timeout});
202
203     # Erste Zeile der Anfrage enthaelt HTTP-Methode, -URL und -Version,
204     # empfange und teile nur die ersten max_url_length+100 Bytes:
205     my($method,$url,$version)
206       =split(/ /,substr(readline($client),0,$conf->{max_url_length}+100),3);
207
208     # Sende Nachricht und beende Verbindung, falls URL zu lang:
209     if(length($url) > $conf->{max_url_length}){
210       sendError($client,414,'Request URL Too Long',
211         'The URL requested is longer than '.$conf->{max_url_length}.' bytes.');
212       exitChild();
213     }
214
215     # Extrahiere GET-Parameter aus dem URL:
216     ($url,$getString)=split(/\?/,$url,2);
217
218     # Konvertiere URI-enkodierte Hex-Werte im URL fuer Pfad zu normalen Zeichen:
219     $file=decodeUri($url);
220
221     # Fahre nur fort, wenn URL ausschliesslich aus "sichere Zeichen" besteht
222     # und mit / beginnt, ausserdem muessen double-dot-Attacken ausgeschlossen
223     # sein:
224     # *FIXME* Ansich sollte ein Server .. aufloesen koennen - das waere z.B.
225     # durch s/\/.*?\/\.\.// moeglich, wird aber hier nicht implementiert um den
226     # Code nicht zu komplex (und damit fehlertraechtig) zu machen.
227     if(not($file=~/^\/[\/a-z0-9\.\_\- ]*$/i) || $file=~/\/\.\.(\/|$)/){
228       sendError($client,400,'Bad Request',
229         'The filename ('.$file.') requested contains bad characters.');
230       exitChild();
231     }
232
233     # Ueberspringe uebermittelte HTTP-Headers (bis Leerzeile empfangen):
234     getHeaders($client);
235
236     # Setze Pfad zur angeforderten Datei im Root-Verzeichnis zusammen:
237     $file=$conf->{root}.'/'.$file;
238     # Setze Hostname auf gebundene IP-Adresse:
239     $host=$conf->{bind_address};
240
241     # Entferne alle double slashes:
242     $file=~s/\/\//\//g;
243
244     # Wenn URL nicht auf / endet und ein Verzeichnis ist, leite den Client dort
245     # hin weiter:
246     if(not($url=~/\/$/) && -d $file){
247       logAccess($client->peerhost().' 302 '.$url.' => '.$url.'/');
248       sendHeader($client,302,'Found',{Location => $url.'/'});
249       sendStatus($client,302,'Found',
250         'This document is located at <a href="'.$url.'">'.$url.'</a>');
251       exitChild();
252     }
253
254     # Akzeptiere GET-Methode:
255     if(uc($method) eq 'GET'){
256       serveFile($client,$file,$url,$host);
```

```
  }
  # Alle anderen HTTP-Methoden werden nicht unterstuetzt:
  else{
    sendError($client,405,'Method Not Allowed',
      'The method used in request is not allowed here.');
    exitChild();
  }

}


# Schicke die angeforderte Datei zum Client:
sub serveFile($$$$){
  my($client,$file,$url,$host)=@_;

  my $fh;               # Enthaelt spaeter das Dateihandle.
  my $buffer;           # Buffer fuer das Einlesen von Daten.
  my $fileExt;          # Enthaelt spaeter die Dateierweiterung.
  my $mimeType;         # Enthaelt spaeter den MIME-Type zur Dateierweiterung.

  # Index-Datei an Pfad anhaengen, falls Verzeichnis:
  if(-d $file && $file=~/\/$/){
    $file.=$conf->{index_file};
  }

  # Dateierweiterung aus Dateiname extrahieren und MIME-Type bestimmen:
  ($fileExt)=($file=~/.*\/.*?\.([a-z0-9]+)$/i);
  $mimeType=(defined($fileExt) && exists($conf->{mime_types}->{lc($fileExt)}))
    ? ($conf->{mime_types}->{lc($fileExt)}) : ($conf->{mime_types}->{'*'});

  # Wenn Datei nicht existiert, gebe Fehler aus:
  # *FIXME* Trifft auch zu, falls Datei in einem Verzeichnis ohne Zugriffsrechte
  # liegt!
  unless(-f $file){
    logAccess($client->peerhost().' 404 '.$host.' '.$file);
    sendError($client,404,'Not Found',
      'The file requested ('.$url.') does not exist on '.$host.'.');
    exitChild();
  }

  # Oeffne Datei und gebe Fehler aus, falls ohne Erfolg:
  $fh=new IO::File($file,'r')
    || do{
      logAccess($client->peerhost().' 403 '.$host.' '.$file);
      sendError($client,403,'Forbidden',
        'You are not allowed to access the file requested ('.$url.') on '
        .$host.'.');
      exitChild();
    };

  # Sende 200 OK inkl. Header der Dateigroesse:
  sendHeader($client,200,'OK',
      {
        Length         => -s $file,
        'Content-Type' => $mimeType
      }
    );

  # Schalte Dateihandle und Client-Verbindung in den Binaermodus und schiebe
  # alle Daten aus dem File zum Client durch, schliesse dann das File.
  binmode($fh);
  binmode($client);
  while(read($fh,$buffer,4096)){
    print($client $buffer);
  }
  $fh->close();

  # Erfolg protokollieren und Verbindung schliessen, Prozess beenden.
  logAccess($client->peerhost().' 200 '.$host.' '.$file.' ('.$mimeType.')');
  exitChild();

}

# Empfange alle Headers und verwerfe sie:
sub getHeaders($){
  my($client)=@_;

  my $bytes=0;          # Byte-Counter fuer maximale Header-Groesse.

  # Enthaelt die (erste) zu verarbeitende Header-Zeile.
  my $line=readline($client);

  # Bis zu ersten Leerzeile sind alles Header:
  while(defined($line) && not($line=~/^[\r\n]+$/)){

    # Erhoehe den Byte-Counter und beende Verbindung, wenn uebergelaufen:
    $bytes+=length($line);
    print($line);
    if($bytes > $conf->{max_headers_length}){
      sendError($client,413,'Request Entity Too Large',
        'Your request was bigger than '.$conf->{max_headers_length}.' bytes.');
      exitChild();
    }

    # Lese naechste Zeile:
    $line=readline($client);

  }

}


# Sende eine vollstaendige Fehlermeldung und erzeuge Log-Eintrag:
sub sendError($$$$){
  my($client,$code,$status,$message,)=@_;
  sendHeader($client,$code,$status,{});
  sendStatus($client,$code,$status,$message);
  logError($client->peerhost().' ERROR'.$code.': '.$status);
}


# Erzeuge und sende einen HTTP-Header:
sub sendHeader($$$\%){
  my($client,$code,$status,$addHeaders)=@_;

  # Default-Werte:
  my %headers=(
      'Content-Type' => 'text/html',
      'Server'       => 'MicroWeb/'.$VERSION,
    );

  # Fuege alle Header in $addHeaders %headers hinzu:
  foreach(keys(%{$addHeaders})){
    $headers{$_}=$addHeaders->{$_};
  }

  # Sende HTTP-Header an den Client:
  print($client 'HTTP/1.0 '.$code.' '.$status."\r\n");
```

```
  foreach(keys(%headers)){
    print($client $_.': '.$headers{$_}."\r\n");
  }
  print($client "\r\n");

}


# Erzeuge und sende eine HTTP-Statusmeldung in HTML:
sub sendStatus($$$){
  my($client,$code,$status,$message)=@_;
  print($client <<__EOF);
<html>
  <head>
    <title>$code $status</title>
  </head>
  <body>
    <h1>$status</h1>
    <p>
      $message
    </p>
    <hr>
    <p>
      MicroWeb $VERSION, a simple HTTP server implementation &mdash;
      &copy; 2006 Veit Wahlich
    </p>
  </body>
</html>
__EOF
}


# Erzeuge einen Timestamp mit der aktuellen lokalen Zeit fuer Log-Nachrichten
# unter Verwendung von POSIX::strftime():
sub timestamp(){
  return(strftime('[%Y-%m-%d %H:%M:%S] ',localtime(time)));
}


# Gebe eine Nachricht auf STDERR aus:
sub logError($){
  my($message)=@_;
  print(STDERR timestamp().$message."\n");
}


# Eine Nachricht auf STDOUT ausgeben:
sub logAccess($){
  my($message)=@_;
  print(STDOUT timestamp().$message."\n");
}


# Beendet den Kindprozess und schliesst zuvor sauber alle Handles:
sub exitChild(){
  $client->shutdown(2);
  $client->close();
  STDOUT->close();
  STDERR->close();
  exit(0);
}


# Der SIGCHLD-Handler erntet tote Kinder. Das nennt man wirklich so...
sub reapPid(){

  # Warte auf die PID des toten Kindprozesses...
  my $pid=wait();

  # Wenn sie richtig uebergeben wurde, dekrementiere die Anzahl laufender
  # Kindprozesse.
  if($pid > 0){
    $children--;
  }

  # Setze den SIGCHLD-Handler zur Sicherheit nochmal:
  $SIG{CHLD}=\&reapPid;

}


# Dekodiere URI-encodierten String:
sub decodeUri($){
  my($line)=@_;
  if(defined($line)){
    $line=~tr/+/ /;
    $line=~s/%([a-f0-9]{2})/pack('C',hex($1))/egi;
  }
  return($line);
}


main();
1;
```