Dr. Matthias Grabert — Abteilung Angewandte Informationsverarbeitung — 31. Mai 2002
Christian Ehrhardt — Blatt 6

Systemnahe Software (SS 2002)

Abgabetermin 11.06.2002

News (10 Punkte)

Ein Newsserver speichert Diskussionbeiträge (news) zu verschiedenen Themen geordnet in verschiedenen Gruppen. Eine Gruppe enthält dabei alle Beiträge, zu einem bestimmten Thema. Newsserver können weltweit vernetzt sein oder alleine betrieben werden. Ein solcher Server läuft auf dem Rechner miner.mathematik.uni-ulm.de auf Port 8080 (siehe unten). Auf diesem Server gibt es, wie der eine oder andere von Euch vielleicht schon weiß, eine Newsgruppe mit dem Namen miner.local.softss02, deren Thema unsere Vorlesung ist.

Die Aufgabe

In diesem Blatt soll ein Programm geschrieben werden, das den Absender, das Datum und das Subject der letzten 5 Beiträge in der Newsgruppe zur Vorlesung ausgibt.

Die Netzwerkverbindung

Normalerweise werden C-Funktionen verwendet, um eine Netzwerkverbindung zum Newsserver herzustellen. Da die Vorlesung noch nicht ganz so weit ist (und weil wir nochmal ein bißchen pipes verwenden wollen), soll das Programm nc (auf Linuxsystemen manchmal auch unter dem Namen netcat bekannt) verwendet werden um eine Netzwerkverbindung zum Newsserver herzustellen. Dazu muß folgendes Kommando "nc miner.mathematik.uni-ulm.de 8080" aufgerufen werden. Dann wird die Standardeingabe von nc an den Newsserver weitergeleitet und die Antworten des Newsservers landen auf der Standardausgabe.

Das Protokoll oder wie rede ich mit dem Newsserver

Wichtiger Hinweis: Im folgenden ist mit einer Zeile immer eine Folge von Zeichen gemeint, die durch die zwei Zeichen "\r\n" und nicht nur durch das Zeichen "\n" beendet wird.

Die Kommunikation mit dem Newsserver läuft zeilenweise ab. Der Client (also Euer Programm) schickt ein einzeiliges Kommando an den Server. Dieser antwortet darauf ebenfalls mit einer einzelnen Zeile, die aus einer dreistelligen Zahl (dem Rückgabewert) gefolgt von einem erklärenden Text besteht. Aus dem Rückgabewert kann der Client dann schließen, ob seine Anfrage erfolgreich war, oder was der Grund für den Fehlschlag gewesen ist. Außerdem kann aus dem Rückgabewert abgelesen werden, ob außer der einzeiligen Antwort noch weitere Daten folgen. Sollte das der Fall sein, dann gehören alle folgenden Zeilen ebenfalls noch zur Antwort des Servers. Das Ende der Antwort erkennt der Client an einer Zeile, die nur aus einem Punkt (gefolgt von "\r\n") besteht.

Bevor der Client irgendwelche Befehle schickt, muß er auf eine Begrüßung durch den Server warten (als Reaktion auf den Aufbau einer Netzwerkverbindung). Diese Begrüßung wird durch den "Rückgabewert" 200 oder 201 gekennzeichnet.

Die wichtigsten Kommandos

GROUP

Mit den Kommando "GROUP gruppenname\r\n" kann eine bestimmte Gruppe auf dem Server ausgewählt werden. Alle folgenden Kommandos beziehen sich dann auf diese Gruppe. Mögliche Rückgabewerte:

211 Die gewünschte Gruppe wurde erfolgreich ausgewählt. In diesem Fall folgen auf den Rückgabewert (noch in der gleichen Zeile) drei Integerzahlen. Die erste kann ignoriert werden, die zweite Zahl gibt die kleinste gültige Artikelnummer (siehe unten) in der Gruppe an, die dritte Zahl ist die größte Artikelnummer.

411 Die gewünschte Gruppe existiert auf dem Server nicht.

STAT, HEAD, BODY und ARTICLE

Die vorhandenen Beiträge (Artikel) in einer Gruppe werden auf dem Server aufsteigend numeriert, d.h. je größer die Nummer eines Beitrags, desto neuer ist er. Die Nummer des ältesten und neuesten noch vorhandenen Artikels erfährt man durch das GROUP Kommando. Nicht alle Artikelnummern zwischen diesen beiden Werten müssen vergeben sein. Ein Artikel selbst besteht aus zwei Teilen, einem Kopf und dem Inhalt. Der Kopf enthält Informationen wie den Autor und das Subject, der Inhalt enthält die eigentliche Nachricht.

STAT nr\r\n Mit diesem Kommando wird nur angefragt, ob die Artikelnummer nr tatsächlich vergeben ist.

HEAD nr\r\n Mit diesem Kommando wird der Server aufgefordert, den Kopf des Artikels mit der Nummer nr zu schicken.

BODY nr\r\n Mit diesem Kommando wird der Server aufgefordert, den Inhalt des Artikels mit der Nummer nr zu schicken.

ARTICLE nr\r\n Mit diesem Kommando wird der Server aufgefordert, sowohl den Kopf alsauch den Inhalt des Artikels mit der Nummer nr zu senden. Kopf und Inhalt werden durch eine Leerzeile voneinander getrennt.

Alle vier Kommandos lassen die gleichen Rückgabewerte zu:

220 Der Artikel wurde gefunden, Kopf und Inhalt folgen.

221 Der Artikel wurde gefunden, Kopf folgt.

222 Der Artikel wurde gefunden, Inhalt folgt.

223 Der Artikel wurde gefunden, es folgen keine weiteren Daten.

412 Es wurde noch keine Gruppe ausgewählt.

420 Dieser Fehler kann auftreten, wenn keine Artikelnummer hinter STAT, HEAD, BODY oder ARTICLE angegeben wurde.

423 und 430 Es wurde kein Artikel mit der angegebenen Nummer gefunden.

QUIT

Mit diesem Kommando wird die Kommunikation mit dem Server beendet. Die einzige gültige Antwort darauf ist der Rückgabewert 205, der besagt, daß die Verbindung beendet wird/wurde.

Was wollen wir eigentlich oder wie sieht so ein Artikel aus

Die Informationen, die über die letzten 5 Artikel ausgegeben werden sollen, befinden sich im Kopf des Artikels, der Inhalt interessiert erst mal nicht. Aus dem Kopf interessieren uns auch nur drei Dinge:

Der Absender Das ist eine Zeile im Kopf des Artikels, die mit der Zeichenfolge "From: " beginnt.

Das Datum Das ist eine Zeile im Kopf des Artikels, die mit der Zeichenfolge "Date: " beginnt.

Das Subject Das ist eine Zeile im Kopf des Artikels, die mit der Zeichenfolge "Subject: " beginnt.

Hinweise und Fallen

Carriage Return Das Zeichen "\r" (auch mit Carriage Return oder kurz CR bezeichnet) kann auf der Tastatur durch die Tastenkombination Ctrl-V gefolgt von Ctrl-M erzeugt werden, das Zeichen wird auf dem Bildschirm dann durch "^M" symbolisiert. Das ist hilfreich, um mit nc die Unterhaltung mit dem Newsserver von Hand zu testen. In einem C-Programm kann Ctrl-V Ctrl-M **nicht** verwendet werden.

fdopen Die Funktion fdopen erlaubt es einen Filedeskriptor (wie ihn open oder pipe liefern) in einen FILE-Pointer (wie von fopen) zu konvertieren. Damit ist es möglich auf einer pipe mit Funktionen wie fscanf usw. zu arbeiten. Beim Schließen muß dann allerdings fclose und nicht mehr close verwendet werden. Außerdem sollte man auch beim Lesen und Schreiben darauf achten, daß nicht gepufferte und ungepufferte Ein- und Ausgabe gemischt verwendet werden.

Pufferung Da die Ausgabe bei der Arbeit mit FILE-Pointern grundsätzlich gepuffert wird, muß der Bibliothek nach einem Kommando ausdrücklich mitgeteilt werden, daß der Puffer jetzt geleert werden soll. Geschieht dies nicht, dann wird das Kommando gepuffert bis genügend Daten beisammen sind und der Server bekommt das Kommando zunächst garnicht zu sehen. Die Funktion fflush kann dazu verwendet werden die Ausgabepuffer zu leeren.