Zadanie programistyczne nr 3
z Sieci komputerowych

1 Opis zadania

Napisz program `transport`, który będzie łączyć się z określonym serwerem i wysyłając — zgodnie z opisanym niżej protokołem — pakiety UDP, pobierze od serwera plik o określonej wielkości.

Program powinien przyjmować trzy argumenty: *port*, *nazwa_pliku* i *rozmiar* w bajtach. Celem jest pobranie od specjalnego serwera UDP nasłuchującego na porcie *port* komputera o adresie `156.17.4.30` pierwszych *rozmiar* bajtów pliku i zapisanie ich w pliku *nazwa_pliku_wynikowego*. W celu komunikacji z serwerem UDP należy wysyłać mu datagramy UDP zawierające następujący napis:

`GET` *start długość*`\n`

Wartość *start* powinna być liczbą całkowitą z zakresu $[0, 10\,000\,000]$, zaś *długość* z zakresu $[1, 1\,000]$. Znak `\n` jest uniksowym końcem wiersza, zaś odstępy są pojedynczymi spacjami. Jedyną zawartością datagramu musi być powyższy napis: serwer zignoruje datagramy, które nie spełniają tej specyfikacji. W odpowiedzi serwer wyśle datagram, na którego początku będzie znajdować się napis:

`DATA` *start długość*`\n`

Wartości *start* i *długość* są takie, jakich zażądał klient. Po tym napisie znajduje się *długość* bajtów pliku: od bajtu o numerze *start* do bajtu o numerze *start*+*długość*-1. Uwaga: bajty pliku numerowane są od zera.

1.1 Przykład

Jeśli program zostanie uruchomiony w następujący sposób:

```
$> ./transport 40001 output 1100
```

to może wysłać w do portu 40001 serwera `156.17.4.30` dwa datagramy o treściach:

```
GET 0 700\n
```

oraz

```
GET 700 400\n
```

następnie zaczekać na odpowiedzi serwera i zapisać odpowiednie 1100 bajtów do pliku `output`.

1.2 Dodatkowe wymagania

Pamiętaj, że pobierane dane są danymi binarnymi i mogą zawierać bajt równy zero. Należy je zatem zapisywać do pliku funkcją `write()`, `fwrite()` lub podobną. Twój program powinien zapisywać dane wyłącznie na końcu pliku wyjściowego. W szczególności niedopuszczalne jest uzyskiwanie dostępu do różnych miejsc pliku wyjściowego za pomocą funkcji `lseek()` lub podobnej. Twój program powinien zużywać co najwyżej 5 MB pamięci operacyjnej (łącznie z binarnym kodem programu; patrz część *Sposób oceniania programów*).

Program powinien obsługiwać błędne dane wejściowe, zgłaszając odpowiedni komunikat. Program nie powinien wypisywać niepotrzebnych komunikatów diagnostycznych, ale może wypisywać postęp w pobieraniu kolejnych części pliku.

Serwer UDP oczekuje na portach od 40001 do 40003; na takich portach można testować swój program. Nie należy ograniczać działania programu do powyższych portów: podczas oddawania programu będzie on testowany na osobnej instancji serwera UDP działającej na innym porcie.

1.3 Zawodna komunikacja

Należy pamiętać, że datagramy UDP mogą zostać zagubione, zduplikowane lub nadejść w innej kolejności niż były wysyłane. Wykorzystywany serwer UDP sprzyja nauczeniu się tych reguł, działając w następujący sposób:

1. Odpowiedź na dane żądanie zostaje wysłana z prawdopodobieństwem ok. 1/2.
2. Każda wysyłana odpowiedź zostanie zduplikowana z prawdopodobieństwem ok. 1/5. W tej definicji duplikat też jest traktowany jako odpowiedź, więc może pojawić się też duplikat duplikatu (z prawdopodobieństwem ok. 1/25) itd.
3. Każdy wysyłany datagram jest wysyłany z losowym opóźnieniem wynoszącym od 0,5 do 1,5 sekundy.
4. Serwer utrzymuje kolejkę co najwyżej 1000 datagramów, które ma wysłać.

1.4 Uwagi implementacyjne

Konieczne jest sprawdzanie, czy adres źródłowy i port źródłowy datagramu jest prawidłowy. Możesz założyć, że jeśli Twój program otrzymuje dane od adresu IP i portu, do którego dane uprzednio wysłał, to są one zgodne ze specyfikacją i nie trzeba tego sprawdzać.[1]

Twój program będzie testowany pod kątem poprawności i wydajności. W najprostszym wariancie można zaprogramować go jako algorytm typu *stop-and-wait*. Do odczekiwania i sprawdzania, czy gniazdo zawiera datagram, można wykorzystać funkcję `select()`. Tak zaimplementowany został program `transport-slower` (statycznie skonsolidowaną wersję 64-bitową można pobrać ze strony wykładu). Za poprawną implementację takiego podejścia można dostać ok. 4 punktów.

Podejście typu *stop-and-wait* jest bardzo nieefektywne. Aby je poprawić, możesz zaimplementować algorytm przesuwnego okna, utrzymujący odpowiednie liczniki czasu dla wszystkich otrzymanych segmentów. (Ograniczenie na pamięć ma na celu wymuszenie, żeby okno nie mieściło wszystkich datagramów: prefiks danych, który udało nam się pobrać należy zapisać do pliku). Taki algorytm wykorzystuje program `transport-faster` (również do pobrania ze strony wykładu). Za taką implementację można dostać maksymalną liczbę punktów.

[1] W prawdziwych zastosowaniach byłby to bardzo zły pomysł.

Twój program nie musi pobierać danych zgodnie z opisanymi wyżej schematami. Program będzie testowany w pracowni 109, czyli parametry wpływające na jego efektywność (czas odczekiwania i rozmiar okna) należy dobrać eksperymentalnie w sieci instytutowej. Napisanie programu, który będzie dynamicznie dostosowywał się do parametrów łącza nie jest wymagane.

2 Uwagi techniczne

Pliki Sposób utworzenia napisu oznaczanego dalej jako *imie_nazwisko*: Swoje (pierwsze) imię oraz nazwisko proszę zapisać wyłącznie małymi literami zastępując litery ze znakami diakrytycznymi przez ich łacińskie odpowiedniki. Pomiędzy imię i nazwisko należy wstawić znak podkreślenia.

Swojemu ćwiczeniowcowi należy dostarczyć jeden plik o nazwie *imie_nazwisko*`.tar.bz2` z archiwum (w formacie `tar`, spakowane programem `bzip2`) zawierającym pojedynczy katalog o nazwie *imie_nazwisko* z następującymi plikami.

- Kod źródłowy w C lub C++, czyli pliki `*.c` i `*.h` lub pliki `*.cpp` i `*.h`. Każdy plik `*.c` i `*.cpp` na początku powinien zawierać w komentarzu imię, nazwisko i numer indeksu autora.
- Plik `Makefile` pozwalający na kompilację programu po uruchomieniu `make`.
- Ewentualnie plik `README`.

W katalogu tym **nie** powinno być żadnych innych plików, w szczególności skompilowanego programu, obiektów `*.o`, czy plików źródłowych nie należących do projektu.

Kompilacja Kompilacja i uruchamianie przeprowadzane zostaną w 64-bitowym środowisku Linux. Kompilacja w przypadku C ma wykorzystywać standard C99 z ewentualnymi rozszerzeniami GNU (opcja kompilatora `-std=c99` lub `-std=gnu99`), zaś w przypadku C++ — standard C++11 z ewentualnymi rozszerzeniami GNU (opcja kompilatora `-std=c++11` lub `-std=gnu++11`). Kompilacja powinna wykorzystywać opcje `-Wall` i `-Wextra`. Podczas kompilacji nie powinny pojawiać się ostrzeżenia.

3 Sposób oceniania programów

Poniższe uwagi służą ujednoliceniu oceniania w poszczególnych grupach. Napisane są jako polecenia dla ćwiczeniowców, ale studenci powinni **koniecznie się** z nimi zapoznać, gdyż będziemy się ściśle trzymać poniższych wytycznych. Programy będą testowane na zajęciach w obecności autora programu. Na początku program uruchamiany jest w różnych warunkach i otrzymuje za te uruchomienia od 0 do 10 punktów. Następnie obliczane są ewentualne punkty karne. Oceniamy z dokładnością do 0,5 punktu. Jeśli ostateczna liczba punktów wyjdzie ujemna wstawiamy zero. (Ostatnia uwaga nie dotyczy przypadków plagiatów lub niesamodzielnych programów).

Testowanie: punkty dodatnie Rozpocząć od kompilacji programu. W przypadku programu niekompilującego się stawiamy 0 punktów, nawet jeśli program będzie ładnie wyglądał.

Do testów zostanie uruchomiona osobna instancja serwera, działająca na innych portach; porty te zostaną podane ćwiczeniowcom. Chodzi o to, żeby podczas oceniania zadań zminimalizować efekty związane z jednoczesnym dostępem do serwera przez wiele osób.

Przy uruchamianiu programu należy wykorzystywać polecenie `/usr/bin/time -v`. Umożliwi to pomiar czasu i zajętość pamięci (pola `Elapsed time` i `Maximum resident set size`).

3 pkt. Uruchomić program do pobrania ok. 15 000 bajtów, gdzie liczba bajtów nie jest wielokrotnością 1 000. Na takich samych danych uruchomić program `transport-slower`; niech t będzie czasem jego działania. Program studenta otrzymuje punkty, jeśli jego czas działania jest nie większy niż $4 \cdot t + 5$ sek, zajętość pamięci nie większa niż 5 MB, a pliki generowane przez oba programy są identyczne.

1 pkt. Uruchomić program do pobrania ok. 15 000 bajtów. Zatrzymać go w trakcie wykonywania; sprawdzić Wiresharkiem jaki jest jego port źródłowy. Następnie poleceniem `nc` wysłać do tego portu źródłowego datagram zawierający śmieci. Wznowić działanie programu i sprawdzić, czy generowany jest poprawny plik.

3 pkt. Jak w pierwszym punkcie, ale pobieramy ok. 1 000 000 bajtów i porównujemy czas z czasem działania programu `transport-faster`

3 pkt. Jak w pierwszym punkcie, ale pobieramy ok. 9 000 000 bajtów i porównujemy czas z czasem działania programu `transport-faster`

Punkty karne Punkty karne przewidziane są za następujące usterki.

-2 pkt. Za każdy rozpoczęty megabajt ponad limit 5 MB na zajętość pamięciową programu.

-5 pkt. Zapis do pliku wyjściowego w innym miejscu niż na jego końcu.

-1 pkt. Brak sprawdzania poprawności danych na wejściu.

do -3 pkt. Zła / nieczytelna struktura programu: wszystko w jednym pliku, brak modularności i podziału na funkcjonalne części, niekonsekwentne wcięcia, powtórzenia kodu.

-2 pkt. Aktywne czekanie zamiast zasypiania do momentu otrzymania pakietu.

-1 pkt. Brak sprawdzania poprawności wywołania funkcji systemowych, takich jak `recvfrom()`, `write()` czy `bind()`.

-1 pkt. Nietrzymanie się specyfikacji wejścia i wyjścia. Przykładowo: wyświetlanie nadmiarowych informacji diagnostycznych, inna niż w specyfikacji obsługa parametrów.

-1 pkt. Zły plik `Makefile` lub jego brak: program powinien się kompilować poleceniem `make`, polecenie `make clean` powinno czyścić katalog z tymczasowych obiektów (plików `*.o`), polecenie `make distclean` powinno usuwać skompilowane programy i zostawiać tylko pliki źródłowe.

-1 pkt. Niewłaściwa kompilacja: nietrzymanie się opcji podanych w zadaniu, ostrzeżenia wypisywane przy kompilacji, kompilacja bezpośrednio do pliku wykonywalnego bez tworzenia obiektów tymczasowych `*.o`.

(-3/-6 pkt) Kara za wysłanie programu z opóźnieniem: -3 pkt. za opóźnienie do 1 tygodnia, -6 pkt. za opóźnienie do 2 tygodni. Programy wysyłane z większym opóźnieniem nie będą sprawdzane.

Marcin Bieńkowski