

DETAIL PŘEDMĚTU » ZADÁNÍ

< Zpět na detail předmětu (student.phtml?sn=predmet_detail&apid=231021)

< Zpět na registrace zadání (student.phtml?sn=projekty)

DETAIL ZADÁNÍ

Předmět:

ISA - Síťové aplikace a správa sítí (?gm=gm_detail_predmetu&apid=231021)

Ak. rok:

2022/2023

Název:

Čtečka novinek ve formátu Atom a RSS s podporou TLS - Projekt ISA

Vedoucí:

Ing. Libor Polčák, Ph.D. (/lide/libor-polcak-78789)

Literatura:

Vizte text zadání

Popis:

Napište program feedreader, který bude vypisovat informace uvedené ve stažených zdrojích (feed) ve formátu Atom a RSS 2.0. Pro RSS se řiďte dokumentací odkazovanou níže. Program po spuštění stáhne zadané zdroje a na standardní výstup vypíše informace požadované uživatelem (např. názvy článků).

Při vytváření programu je povoleno použít hlavičkové soubory pro práci se sokety a další obvyklé funkce používané v síťovém prostředí (jako je `netinet/*`, `sys/*`, `arpa/*` apod.), knihovnu pro práci s vlákny (`pthread`), signály, časem, stejně jako standardní knihovnu jazyka C (varianty ISO/ANSI i POSIX), C++ a STL. Pro práci s daty ve formátu XML je doporučena knihovna `libxml2`, případně jiné knihovny dostupné na serveru merlin a eva. Pro práci s TLS je doporučena knihovna `openssl`, případně jiné knihovny dostupné na serveru merlin a eva. Pro práci s XML a TLS můžete použít i jiné knihovny, musí však být jasně označeny, že jde o převzatý kód a musí být součástí odevzdávaného archivu. Jiné knihovny nejsou povoleny.

Spuštění aplikace

Použití: `feedreader <URL | -f <feedfile>> [-c <certfile>] [-C <certaddr>] [-T] [-a] [-u]`

Pořadí parametrů je libovolné. Popis parametrů:

- Povinně je uveden buď URL požadovaného zdroje (příčemž podporovaná schémata jsou `http` a `https`), nebo parametr `-f` s dodatečným parametrem určujícím umístění souboru `feedfile`. Soubor `feedfile` je textový soubor, kde je na každém řádku uvedena jedna adresa zdroje ve formátu Atom, či RSS. Prázdné řádky v souboru `feedfile` ignorujte. Řádky začínající znakem `#` v souboru `feedfile` ignorujte, jsou to komentáře. Soubor `feedfile` je obvyklý Unixový textový soubor, tzn. poslední znak na každém řádku je LF. Z toho také vyplývá, že poslední znak v souboru je LF.
- Volitelný parametr `-c` definuje soubor `<certfile>` s certifikáty, který se použije pro ověření platnosti certifikátu SSL/TLS předloženého serverem.
- Volitelný parametr `-C` určuje adresář `<certaddr>`, ve kterém se mají vyhledávat certifikáty, které se použijí pro ověření platnosti certifikátu SSL/TLS předloženého serverem.
- Pokud není uveden parametr `-c` ani `-C`, pak použijte úložiště certifikátů získané funkcí `SSL_CTX_set_default_verify_paths()`.



Při spuštění s parametrem `-T` se pro každý záznam zobrazí navíc informace o čase změny záznamu, či vytvoření záznamu (je-li ve staženém souboru obsaženo).

- Při spuštění s parametrem `-a` se pro každý záznam zobrazí jméno autora, či jeho e-mailová adresa (je-li ve staženém souboru obsaženo).
- Při spuštění s parametrem `-u` se pro každý záznam zobrazí asociované URL (je-li ve staženém souboru obsaženo).

Výstup aplikace

Pro každý zdroj vypisujte informace v následujícím tvaru:

- 1. řádek: Název zdroje uvozený znaky `"*** "` a ukončený znaky `" ***"`, např. `"*** FIT VUT v Brně ***"`.
- Na dalších řádcích budou uvedeny jednotlivé záznamy. V základním tvaru na každém řádku titulek jednotlivých záznamů, např. `"19.9-16.12.2022 Rozdělení do přednáškových skupin BIA a BIB na ZS 2022/2023"`.
- V případě použití parametrů `-T`, `-a` a `-u` budou na dalších řádcích pod zdrojem uvedeny další dodatečné informace uvozeny řetězcem `"Aktualizace: "`, `"Autor: "`, respektive `"URL: "`. Při použití dodatečných atributů budou od sebe jednotlivé záznamy odděleny jedním prázdným řádkem.
- V případě použití parametru `-f` budou od sebe jednotlivé zdroje odděleny jedním volným řádkem.

Příklady vstupů

Příklad souboru *feedfile*:

```
# FIT
http://www.fit.vutbr.cz/news/news-rss.php (http://www.fit.vutbr.cz/news/news-rss.php)
#Wikipedia
https://en.wikipedia.org/w/api.php?
hidebots=1&days=7&limit=50&hidewikidata=1&action=feedrecentchanges&feedformat=atom
(https://en.wikipedia.org/w/api.php?
hidebots=1&days=7&limit=50&hidewikidata=1&action=feedrecentchanges&feedformat=atom)
#The register
https://www.theregister.co.uk/headlines.atom (https://www.theregister.co.uk/headlines.atom)
https://www.theregister.co.uk/data_centre/headlines.atom
(https://www.theregister.co.uk/data_centre/headlines.atom)
https://www.theregister.co.uk/software/headlines.atom
(https://www.theregister.co.uk/software/headlines.atom)
https://www.theregister.co.uk/hardware/headlines.atom
(https://www.theregister.co.uk/hardware/headlines.atom)
#xkcd
https://xkcd.com/atom.xml (https://xkcd.com/atom.xml)
https://what-if.xkcd.com/feed.atom (https://what-if.xkcd.com/feed.atom)
```

Příklad spuštění programu

```
$ feedreader 'https://what-if.xkcd.com/feed.atom (https://what-if.xkcd.com/feed.atom)'
*** what if? ***

Transatlantic Car Rental
Hailstones
...

$ feedreader 'https://what-if.xkcd.com/feed.atom (https://what-if.xkcd.com/feed.atom)' -u
*** what if? ***

Transatlantic Car Rental
URL: https://what-if.xkcd.com/160/ (https://what-if.xkcd.com/160/)
```



URL: <https://what-if.xkcd.com/159/> (<https://what-if.xkcd.com/159/>)

...

```
$ feedreader 'https://what-if.xkcd.com/feed.atom (https://what-if.xkcd.com/feed.atom)' -u -T  
*** what if? ***
```

Transatlantic Car Rental

URL: <https://what-if.xkcd.com/160/> (<https://what-if.xkcd.com/160/>)

Aktualizace: 2022-09-06T00:00:00Z

Hailstones

URL: <https://what-if.xkcd.com/159/> (<https://what-if.xkcd.com/159/>)

Aktualizace: 2022-07-06T00:00:00Z

...

```
$ feedreader 'https://what-if.xkcd.com/feed.atom (https://what-if.xkcd.com/feed.atom)' -u -T-C  
/dev/null
```

Chyba: nepodařilo se ověřit platnost certifikátu serveru tools.ietf.org

Doplňující informace k zadání

- Veškeré vstupy mohou obsahovat chyby, programově tyto situace ošetřete.
- O chybách uživatele informujte na standardním chybovém výstupu.
- V rámci URL je možné volitelně specifikovat číslo portu.
- Pokud uživatel používá TLS, spojení se uskuteční pouze pokud je použitý certifikát důvěryhodný.
- V rámci komunikace HTTP nemusíte řešit přesměrování.
- Pokud je použit *feedfile* a zpracování jednoho ze zdrojů selže, program vypíše chyby a pokračuje dalšími zdroji.
- V dokumentaci popište, jakým způsobem řešíte vypisování doplňujících informací k jednotlivým zdrojům. (URL, čas, autor).
- Vytvořený kód by měl být modulární a otestovaný.
- Pište robustní (https://en.wikipedia.org/wiki/Robustness_principle) aplikace, které budou vstřícné k drobným odchylkám od specifikace (např. porušení přítomnosti povinných elementů).

Referenční prostředí pro překlad a testování

Program by měl být přenositelný. Referenční prostředí pro překlad budou servery eva.fit.vutbr.cz a merlin.fit.vutbr.cz (program musí být přeložitelný a funkční na obou systémech). Vlastní testování může probíhat na jiném počítači s nainstalovaným OS GNU/Linux, či FreeBSD, včetně jiných architektur než Intel/AMD, jiných distribucí, jiných verzí knihoven apod. Pokud vyžadujete minimální verzi knihovny (dostupné na serveru merlin a eva), jasně tuto skutečnost označte v dokumentaci a README.

Testy, které jste při řešení projektu napsali se spustí voláním " `make test` ".

Doporučená literatura

- RFC4287 (<https://tools.ietf.org/html/rfc4287>)
- Specifikace RSS 2.0 (<http://www.rssboard.org/rss-specification>)
- Základy použití knihovny OpenSSL (<http://www.ibm.com/developerworks/library/l-openssl/>)



REGISTRACE ZADÁNÍ

Registrovat od:

26.09.2022 20:10:02

Registrovat do:

09.10.2022 23:59:59

Registraci lze zrušit do:

09.10.2022 23:59:59

Zadání odevzdat do:

14.11.2022 23:59:59

Aktuálně přihlášených:

156

Maximum studentů:

174

Registrován:

ne

Přihlášení automaticky:

ne

Poslední změny provedl:[Zaregistrovat se na toto zadání](#)