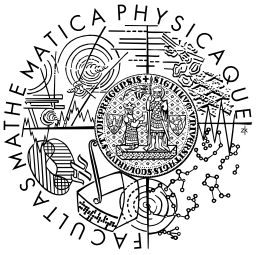

MATEMATICKO-FYZIKÁLNÍ FAKULTA
Univerzita Karlova

BAKALÁŘSKÁ PRÁCE

Michal Medvecký

Peer-to-peer síť pro sdílení novinových článků

Katedra softwarového inženýrství

Vedoucí bakalářské práce: RNDr. David Bednárek, Ph.D.
Studijní program: Informatika
Studijní obor: Programování a softwarové systémy

Praha 2022

Prohlašuji, že jsem tuto bakalářskou práci vypracoval(a) samostatně a výhradně s použitím citovaných pramenů, literatury a dalších odborných zdrojů. Tato práce nebyla využita k získání jiného nebo stejného titulu.

Beru na vědomí, že se na moji práci vztahují práva a povinnosti vyplývající ze zákona č. 121/2000 Sb., autorského zákona v platném znění, zejména skutečnost, že Univerzita Karlova má právo na uzavření licenční smlouvy o užití této práce jako školního díla podle §60 odst. 1 autorského zákona.

V ..... dne ....

Podpis autora

Poděkování.

Název práce: Peer-to-peer síť pro sdílení novinových článků

Autor: Michal Medvecký

Katedra: Katedra softwarového inženýrství

Vedoucí bakalářské práce: RNDr. David Bednárek, Ph.D., 32-KSI

Abstrakt: Abstrakt.

Klíčová slova: peer-to-peer síť internetové noviny P2P protokol

Title: Peer-to-peer network for newspaper publication

Author: Michal Medvecký

Department: Department of Software Engineering

Supervisor: RNDr. David Bednárek, Ph.D., 32-KSI

Abstract: Abstract.

Keywords: peer-to-peer network online newspaper P2P protocol

Obsah

Úvod

Následuje několik ukázkových kapitol, které doporučují, jak by se měla bakalářská práce sázet. Primárně popisují použití TEXové šablony, ale obecné rady poslouží dobře i uživatelům jiných systémů.

1. Dokumentácia protokolu

Peer-to-peer sieť na zdieľanie novinových článkov implementuje peer-to-peer protokol na zdieľanie novinových článkov, skrátene NP2PSP, z anglického „Newspaper P2P Sharing Protocol“.

1.1 Prehľad fungovania protokolu

Sekcia opisuje neformálny popis sieťového protokolu NP2PSP a jeho fungovania. Sieť používajúca protokol je vytváraná pre jedny konkrétne noviny a neslúži na komunikáciu medzi dvoma rôznymi novinami, aj keď taká komunikácia je síce možná, znamenalo by to výmenu informácií medzi dvomi oddelenými a okrem protokolu ničím nesúvisiacimi sieťami.

Sieťový protokol NP2PSP je decentralizovaný protokol postavený na princípe peer-to-peer. Základnou stavebnou jednotkou celej siete je takzvaný peer. Hoci sa jedná o peer-to-peer protokol, nie je úplná decentralizácia možná. Súvisí to s tým, že inštitúcia novín samotná má nejakú hierarchiu a teda nie je pravdou, že by dvaja peeri boli nutne na rovnakej úrovni. Tak tomu je z uhla pohľadu používateľa. Čo sa týka protokolu samotného, má každý peer rovnaké možnosti.

Peer môže byť ako čitateľom, tak aj novinárom či šéfredaktorom novín. Tieto noviny si peer vytvára sám, pričom všetky články, ktoré do novín on sám uverejní, budú k dispozícii verejne pre každého peera, ktorý si o ne požiada. Vyžadovaný je formát markdown, pričom je možné podporovať aj ďalšie spôsoby formátovania textu, ako napríklad HTML, TeX, či roff.

Na identifikáciu článkov sa používa hash článku, pričom ten by mal byť spočítaný z celého článku, z jeho ako textovej, tak aj multimediálnej časti. Podpora obrázkov tvorí minimálnu požiadavku pre podporu multimédií. Možné, a aj odporúčané rozšírenie je o podporu videa, či audia.

Peer posiela správy po sieti buď to priamo druhému peerovi, ktorý je spolu s ním na lokálnej sieti alebo má verejnú IPv4 adresu. Ak sa však jeden z peerov nachádza za smerovačom, ktorý má zapnutý NAT, je potrebné tento NAT obísť, pričom sa na tento účel používa dvojica už existujúcich protokolov STUN a TURN. Peer sťa by STUN klient naviaže spojenie prostredníctvom STUN servera aby sa pripojil k ďalšiemu STUN klientovi, teda k ďalšiemu peerovi. Tento postup však v prípade niektorých implementácií NAT nie je možný (napríklad v prípade symetrického NAT-u) a teda je potrebné, aby všetká komunikácia medzi dvoma peermi tiekla skrz jednu konkrétnu sieťovú entitu. K tomu slúži rozšírenie STUN protokolu pod názvom TURN. STUN server implementujúci TURN sa stane TURN serverom a bude sprostredkúvať komunikáciu medzi TURN klientom a TURN peerom, čo je NP2PSP peer, ku ktorému sa chceme pripojiť.

V NP2PSP sieti je TURN serverom pre dané noviny minimálne ich šéfredaktor. Bolo by možné nasadiť ďalších NP2PSP peerov, ktorí by plnili funkciu TURN serveru, ak by to pre plynulejší chod siete bolo potrebné.

Správa je do siete serializovaná pomocou Google Protocol buffers, ktoré majú k dispozícii API pre viaceré programovacie jazyky a prostredia. Takto serializovaná správa je následne zašifrovaná symetrickým kľúčom a odoslaná po sieti jej príjemcovi. Ten si ju následne rozšifruje a spracuje. Potom čo je správa zašifro-

vaná ale predtým, čo je odoslaná je potrebné pridať na začiatok správy takzvaný identifikačný bajt, ktorý identifikuje typ správy, teda či sa jedná o štandardnú správu alebo o správu obsahujúcu symetrický kľúč počas jeho výmeny pri prvej komunikácii jedného peera s tým druhým. Pridaním tohto bajtu vznikne takzvaná metaspráva. Ďalšie typy metaspráv je možné kedykoľvek pridať.

Správy sú šifrované symetricky, pričom každý peer má uložený spoločný symetrický kľúč s ďalším peerom. Ak jeden z peerov tento kľúč nejakým spôsobom stratí, je potrebné, aby sa opäť vygeneroval, podobným spôsobom, ako pri prvom „stretnutí sa“ dvoch peerov. Počas tejto výmeny symetrických kľúčov sa použije špeciálny typ metasprávy a, keďže symetrický kľúč ešte nemajú obe strany, je potrebné, aby bola táto správa šifrovaná asymetricky, pomocou verejných kľúčov oboch strán. Najprv však dôjde k podpisu správy a až tak k jej asymetrickému zašifrovaniu. Používajú sa metódy autentifikovaného šifrovania (tzv. AE, Authenticated Encryption) a teda ak by mali byť dáta nejakým spôsobom pozmenené počas prenosu, bude táto zmena zistená a ohlásená užívateľovi. Multimediálne prvky môžu byť buď to zašifrované tiež, alebo ak nie, tak sú aspoň rovnako autentifikované ako textové dáta. Na tento účel slúžia rôzne schémy typu AEAD, čo je skratka z anglického Authenticated Encryption with Additional Data, umožňujúce sledovať, či neboli nešifrované dáta počas prenosu nijak pozmenené.

2. Technická dokumentácia

2.1 title

3. Užívateľská dokumentácia

3.1 title

Závěr

Seznam použité literatury

Anděl, J. (1998). *Statistické metody*. Druhé přepracované vydání. Matfyzpress, Praha. ISBN 80-85863-27-8.

Anděl, J. (2007). *Základy matematické statistiky*. Druhé opravené vydání. Matfyzpress, Praha. ISBN 80-7378-001-1.

Cox, D. R. (1972). Regression models and life-tables (with Discussion). *Journal of the Royal Statistical Society, Series B*, **34**(2), 187–220.

Dempster, A. P., Laird, N. M. a Rubin, D. B. (1977). Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society, Series B*, **39**(1), 1–38.

Genberg, B. L., Kulich, M., Kawichai, S., Modiba, P., Chingono, A., Kilonzo, G. P., Richter, L., Pettifor, A., Sweat, M. a Celentano, D. D. (2008). HIV risk behaviors in sub-Saharan Africa and Northern Thailand: Baseline behavioral data from project Accept. *Journal of Acquired Immune Deficiency Syndrome*, **49**, 309–319.

Kaplan, E. L. a Meier, P. (1958). Nonparametric estimation from incomplete observations. *Journal of the American Statistical Association*, **53**(282), 457–481.

Lehmann, E. L. a Casella, G. (1998). *Theory of Point Estimation*. Second Edition. Springer-Verlag, New York. ISBN 0-387-98502-6.

Student (1908). On the probable error of the mean. *Biometrika*, **6**, 1–25.

Zoznam obrázkov

Zoznam tabuliek

Seznam použitých zkratek

A. Přílohy

A.1 První příloha