Katedra informatiky
Přírodovědecká fakulta
Univerzita Palackého v Olomouci

BAKALÁŘSKÁ PRÁCE

Bezpečná emailová služba s filtrací spamu

2022

Matěj Cukr

Vedoucí práce:
doc. Mgr. Jan Outrata, Ph.D.

Studijní program: Aplikovaná informatika,
prezenční forma

Bibliografické údaje

Autor:	Matěj Cukr
Název práce:	Bezpečná emailová služba s filtrací spamu
Typ práce:	bakalářská práce
Pracoviště:	Katedra informatiky, Přírodovědecká fakulta, Univerzita Palackého v Olomouci
Rok obhajoby:	2022
Studijní program:	Aplikovaná informatika, prezenční forma
Vedoucí práce:	doc. Mgr. Jan Outrata, Ph.D.
Počet stran:	40
Přílohy:	1 flash disk
Jazyk práce:	český

Bibliographic info

Author:	Matěj Cukr
Title:	Secure email service with spam filtering
Thesis type:	bachelor thesis
Department:	Department of Computer Science, Faculty of Science, Palacký University Olomouc
Year of defense:	2022
Study program:	Applied Computer Science, full-time form
Supervisor:	doc. Mgr. Jan Outrata, Ph.D.
Page count:	40
Supplements:	1 flash drive
Thesis language:	Czech

Anotace

Práce popisuje princip fungování a implementaci emailové služby, která je předmětem praktické části. Je kladen důraz na bezpečnost serveru i komunikace a také na filtraci spamu s postupným zdokonalováním. Na začátku práce je popsán postup vytvoření a zprovozenění samotného emailového serveru. Následně je rozebráno zajištění bezpečnosti jednotlivých částí a způsob, jakým je možné spam filtrovat. Jak službu využívat je uvedeno v uživatelské příručce. V poslední části je uvedena konfigurace vybraných prvků, použité programy a pomocné skripty.

Synopsis

The thesis describes the principle of functioning and implementation of the email service, which is the subject of the practical part. The emphasis is on the security of the server and communication and also on spam filtering with gradual improvement. At the beginning of the thesis, the procedure of creating and commissioning the email server itself is described. Subsequently, the security of individual parts and the way in which spam can be filtered are discussed. How to use the service is described in the user manual. In the last part, the configuration of selected elements, used programs and auxiliary scripts are presented.

Klíčová slova: email; spam; filtrace spamu; emailová komunikace

Keywords: email; spam; spam filtering; email communication

Děkuji vedoucímu své bakalářské práce panu doc. Mgr. Janu Outratovi, Ph.D. za jeho vedení, podporu a odborné rady. Děkuji také kolegyni Kataríně Olejkové za překlad spam filter pluginu do slovenštiny.

Místopřísežně prohlašuji, že jsem celou práci včetně příloh vypracoval samostatně a za použití pouze zdrojů citovaných v textu práce a uvedených v seznamu literatury.

datum odevzdání práce podpis autora

Obsah

Seznam obrázků

Seznam tabulek

Seznam zdrojových kódů

1 Úvod

Většina čtenářů jistě ví, jak email vypadá a jak jej poslat, jsou však odstíněni od detailů, co se děje v pozadí a jak přesně celá emailová komunikace funguje. Pro plné pochopení této práce je potřeba rozšířených znalostí toho, co je to email, spam, doménové jméno, emailová komunikace, emailový server a jak emailová komunikace skutečně probíhá.

1.1 Základní pojmy

Vysvětlení základních pojmů v emailové komunikaci.

Email

Email je ve své podstatě dopis, jen v elektronické podobě a s více možnostmi. Každý email se skládá z hlavičky a těla. Hlavička obsahuje pouze ASCII znaky a její obsah tvoří:

- odesíltel,
- příjemce nebo příjemci,
- předmět,
- datum a čas odeslání emailu,
- kopie,
- skrytá kopie,
- unikátní ID emailu,
- content-type (jak má být zpráva zobrazena),
- záznamy o všech MTA, přes které se email poslal.

Díky standardu MIME (Multipurpose Internet Mail Extensions) je možné v těle emailu použít i jiná kódování než ASCII a také lze připojit přílohy.

Spam

Základní rozdělení emailů je na vyžádané a nevyžádané (nebo-li spam). Jak už název napovídá, vyžádané emaily jsou ty, které si uživatel vyžádal, o které má zájem a mají pro něj nějaký význam.

Oproti tomu nevyžádané emaily nemají pro uživatele žádný význam, jen zabírají místo a znepřehledňují emailovou schránku. Často nabízejí koupi produktů či služeb. Některé nevyžádané emaily se mohou snažit získat od uživatele přihlašovací údaje nebo číslo kreditní karty, to se označuje jako phishing.

Emailová adresa a doménové jméno

Každý uživatel má svou vlastní unikátní emailovou adresu. První část emailové adresy je uživatelské jméno, druhou částí je doménové jméno (dále jen doména). Mezi těmito částmi se nachází typický znak emailových adres, zavináč (@). Příkladem může být emailová adresa *jnovak@seznam.cz*, kde *jnovak* je uživatelské jméno a *seznam.cz* je doména. Příklady dalších domén: gmail.com, yahoo.com, hotmail.com, z českých pak email.cz, posta.cz, upol.cz.

Emailová komunikace

Systém odesílání, doručování a přijímání emailů se nazývá emailová komunikace. Jedná se o nejstarší a nejdéle používaný elektronický komunikační prostředek na světě. Téměř každý člověk s připojením k internetu používá emailovou komunikaci. Denně jsou po celém světě odeslány stovky miliard emailů. [1]

Většina systémů emailové komunikace je postavena na principu klient-server. Emailový klient slouží uživatelům ke správě elektronické pošty - odesílání, přijímání a mazání emailů. Emailový server zajišťuje faktický přenos emailů.

Emailový server

Jedná se o server, který zajišťuje odesílání, přijímání nebo ukládání emailů. Jeden server zpravidla zajišťuje více, i všechny, zmíněné procesy. V praxi se nejčastěji používá jeden server pro jednu doménu, tak je tomu i v této práci. Více serverů pro jednu doménu se zpravidla používá při velkém množství uživatelů kvůli rozložení zátěže a také jako záloha pro případ výpadku jednoho z nich. V některých případech se dá provozovat na jednom serveru i více domén. O tom, co se s emailem na daném serveru stane, rozhodují tzv. agenti.

1.2 Emailový agenti

Každý email při své cestě od odesílatele až ke svému příjemci projde přes tři různé agenty.

Mail Transfer Agent (MTA)

Program MTA zajišťuje přenos emailu od odesílatele až na server příjemce. MTA běží na serveru a pro přenos emailů používá protokol SMTP, SMTPS nebo Submission. Při přenosu projde email typicky přes několik MTA serverů. Mezi nejrozšířenější MTA patří Postfix, Sendmail a Microsoft Exchange Server.

Mail Delivery Agent (MDA)

Program MDA zajišťuje doručení a uložení emailu do konkrétní schránky zvoleného příjemce. MDA běží na serveru, většinou společně s MTA. Pokud MTA rozhodne, že email patří místnímu uživateli (uživateli na lokálním serveru), předá

email MDA a ten jej doručí do schránky uživatele. K zobrazení emailu se používají protokoly IMAP, IMAPS nebo POP3. Mezi nejrozšířenější MDA patří Dovecot, Procmail a Maildrop.

Mail User Agent (MUA)

Jedná se o aplikaci koncového uživatele, která mu umožňuje přistupovat k emailové schránce. Existují dva druhy MUA:

- samostatně instalovatelné aplikace (např. Outlook, Thunderbird, Spark),
- webové služby (např. Gmail, Yahoo!, Seznam).

Hlavní výhody emailové komunikace jsou bezpochyby rychlost, jednoduchost a dostupnost. Na zřízení vlastní emailové schránky stačí pár minut a je zpravidla zcela zdarma. K posílání a zobrazování emailů pak postačí pouze připojení k internetu a chytré zařízení (např. počítač, mobilní telefon, tablet). Ovšem právě kvůli těmto výhodám se dnes potýkáme s jednou zásadní nevýhodou - obrovské množství spamu. Právě spam tvoří většinu poslaných emailů.

2 Emailový server

Pro účely této práce byl na katedře informatiky vyčleněn virtualizovaný server s operačním systémem Linux a platformou Debian 11.

2.1 Použitý software

Seznam použitého softwaru pro zprovoznění a chod emailového serveru. Konkrétní nastavení je v samostatné kapitole 4.1 Konfigurace serveru.

Apache

Webový open-source HTTP server. Apache je nejrozšířenějším webovým serverem a běží na něm většina webových stránek. V tomto případě zajišťuje chod pro RoundCube a společně tvoří MUA, uživatel tak může přistupovat ke své poště skrze webového klienta přímo v prohlížeči.

Certbot

Automatizovaný nástroj spravující certifikáty Let's Encrypt. Tyto certifikáty se používají pro zabezpečenou komunikaci prostřednictvím protokolu HTTPS. Díky tomu je veškerá komunikace mezi uživatelem a webovým serverem šifrovaná. Certbot se stará o vystavování, aktualizování i odebírání certifikátů.

Fail2ban

Software chránící server před brute-force attacks, nebo-li útoky hrubou silou. Více je popsáno v kapitole 2.3.2 Bezpečnost serveru.

MariaDB server

Databázový server uchovávající data všech emailových adres, hesel, domén a aliasů.

Adminer

Webový nástroj pro správu MariaDB.

Postfix

Open-source MTA pro správu odesílání emailů.

Dovecot

Open-source IMAP a POP3 server starající se o ukládání a zobrazování emailů. Zastává roli MDA.

Dovecot ManageSieve

Rozšíření Dovecot o uživatelské filtry příchozí pošty.

RoundCube

Open-source webový klient na bázi IMAP. Umožňuje uživateli přistupovat do jeho emailového schránky skrze webový prohlížeč.

Redis

Open-source databáze, která díky ukládání dat na principu klíč-hodnota funguje velmi rychle a efektivně. Na serveru je využíván programem Rspamd pro rychlé nalezení a porovnání emailů.

Rspamd

Rychlý open-source software pro filtraci spamu. Nabízí také webový nástroj pro zobrazení a nastavení filtrace spamu.

PHP

Podpora skriptovacího jazyka PHP, který využívá Apache pro zobrazení RoundCube.

PHP-zip

Rozšíření PHP podporující vytvoření kompresních archivů zip. Používá se v případě, že chce uživatel stáhnout více příloh jednoho emailu najednou jako jeden soubor.

Mono-complete

Software umožňující kompilaci a spuštění C# kódů přímo v Linuxu.

2.2 Postup řešení

V této kapitole je uveden postup řešení při vytváření emailového serveru a princip jeho fungování. Postup instalace a základní konfigurace vychází z internetového návodu [2], který používá stejný software.

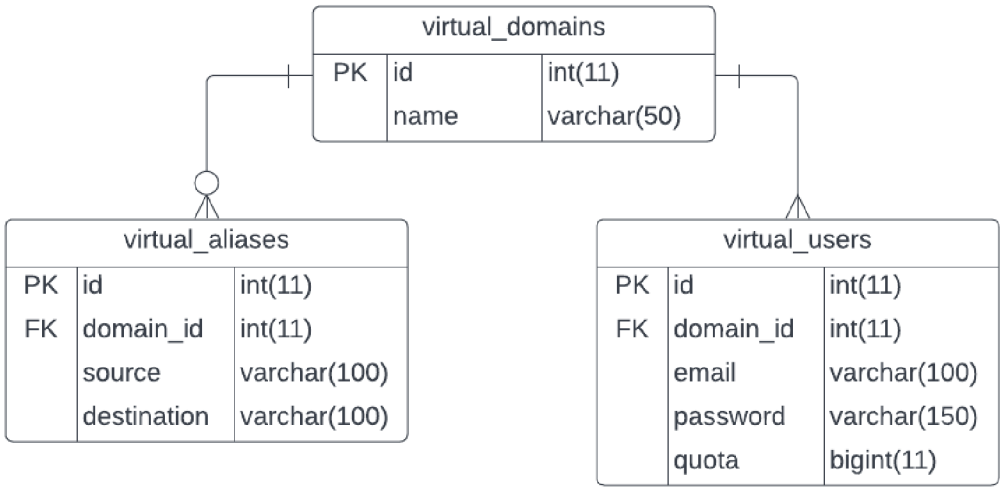
2.2.1 Webový server a komunikace

Nejprve je potřeba si nachystat webový server, na kterém poběží webový emailový klient. Pro webový server byl zvolen software Apache, který v základu podporuje pouze základní internetový protokol HTTP. Ten je však nezabezpečený,

neboť nepodporuje šifrování. Veškerá data jsou přenášena jako čistý text. Jakýkoli útočník, který by odposlouchával komunikaci mezi uživatelem a serverem, může vidět všechna přenášená data, včetně příhlašovacích údajů. Z tohoto důvodu je nezbytně nutné zajistit šifrování a ochránit data před odposlechnutím a zneužitím. Řešení zabezpečení je uvedeno v kapitole 2.3.1 Bezpečnost klient-server komunikace.

2.2.2 Databáze

Pro ukládání údajů je použita databáze MariaDB. Pro snadnější práci s databází skrze webový prohlížeč je navíc použit program Adminer. Jak je uvedeno na obrázku 1, v databázi jsou celkem tři tabulky uchovávající údaje o doménách, uživatelích a aliasech.



Obrázek 1: Schéma databáze

2.2.3 Napojení Postfix na MariaDB

Dalším krokem je napojení Postfix na vytvořenou databázi pro zajištění přijímání emailů. Nejprve napojení na tabulku domén, aby mohl Postfix rychle a snadno rozhodnout, zda příjemce zná. Pokud příjemcovu doménu Postfix pozná, podívá se do tabulky aliasů. Nalezne-li shodu s některým aliasem, zjistí, na jakou skutečnou adresu má email doručit. Jako poslední na řadě je tabulka s uživateli. To je finální krok určení, zda příjmce existuje a je možné mu email doručit. Pokud uživatel nebyl nalezen, Postfix email odmítne, v opačném případě předá email Dovecot, tedy lokálnímu MDA, který se stará o doručení a uložení emailu konkrétnímu příjemci.

2.2.4 Dovecot

Dalším nezbytně důležitým prvkem je Dovecot. Zajišťuje přijetí emailu od Postfix a uložení emailu do adresáře příslušného uživatele. Dále vyhodnocuje uživatelské filtry pro příchozí poštu, kontroluje zaplnění jednotlivých schránek a přístup uživatelů k jejich emailům pomocí protokolu IMAP nebo POP3. Na tomto serveru je po dohodě s vedoucím povolen přístup pouze pomocí protokolu IMAP, protokol POP3 není nastaven.

Dovecot také zajišťuje autentizaci uživatele. Proto je nutné mu nastavit, kde najde databázi s uživatelskými údaji. K ověření uživatele používá dva dotazy. Jeden zjišťující uživatelské jméno, id, cestu ke schránce a kapacitu schránky. Druhý dotaz slouží pro zjištění hesla. Pro uložení emailu do schránky uživatele se na serveru používá protokol LMTP (Local Mail Transport Protocol). Tento protokol neustále běží v jednom vláknu a dokáže zpracovat více emailů současně. Je to v podstatě lokální SMTP ve zjednodušené podobě.

Uživatelské filtry nabízí možnost automatizovaně zpracovávat příchozí poštu. Tyto filtry lze sestavit pomocí tzv. sieve pravidel. Sieve je jazyk pro filtrování emailů. [3] Pravidla napsaná tímto jazykem porovnávají každý příchozí email a pokud email splňuje zadaná kritéria, provede se zvolená akce. Sieve pravidla se vyhodnocují už na serveru, takže je možné je aplikovat na každý příchozí email v čas jeho přijetí a nečeká se, až se uživatel přihlásí do své schránky.

2.2.5 Kvóty

Schránky uživatelů na serveru aktuálně nemají nastavené žádné omezení kapacity. Přesto je však připraven systém kvót, který dokáže s omezenou kapacitou schránek pracovat. Omezení kapacity je možné provést pro schránky všech uživatelů nebo pouze pro vybrané uživatele. O schránky uživatelů se stará Dovecot, který zná i velikost všech emailů v jednotlivých schránkách. Přijetí emailu však zpracovává Postfix. Aby vše správně fungovalo, je třeba zajistit:

- aby si Dovecot udržoval přehled zaplnění všech schránek uživatelů,
- Postfix odmítal přijetí emailu v případě, že příjemce nemá dostatek volného místa.

Jakmile Postfix ověří, že příjemce emailu je uživatel na tomto serveru, zeptá se Dovecot, zda má uživatel dostatečné místo ve své schránce pro přijetí emailu. V reakci na to mu dá Dovecot kladnou nebo zápornou odpověď. Pokud je email možné doručit, Postfix jej předá Dovecot a ten se pak postará o doručení emailu do správné schránky i složky. V případě, že email nemůže být z důvodu nedostatku místa doručen příjemci, musí být odmítnut a to již na straně Postfix. Kdyby Postfix email přijal a odmítnutí provedl až Dovecot, email by zůstal ve frontě na serveru a Postfix by se pak v určitých intervalech snažil o opětovné doručení. Server by se tak zbytečně zatěžoval a odesílatel by nedostal informaci o tom, že jeho email nebyl uživateli doručen.

Aby si uživatel nemusel sám kontrolovat obsazenost své schránky, jsou připraveny automatizované emaily s upozorněním na zaplnění schránky. Dovecot zašle uživateli email s upozorněním při překročení 80% a 95% zaplnění jeho schránky.

2.2.6 RoundCube

Všechny nezbytné kroky pro přijetí emailu jsou splněny. Před nastavením odesílání pošty provedeme zprovoznění webového poštovního klienta. K tomu je zvolen software RoundCube, který je multiplatformní, kompatibilní s ostaním použitým softwarem a navíc zcela zdarma. Jelikož se jedná o webového klienta, je potřeba jej propojit s již připraveným webových serverem Apache. V základu nabízí RoundCube všechny potřebné prvky pro emailovou komunikaci. Pro rozšíření možností uživatele byly do RoundCube doinstalovány a povoleny následující pluginy, nebo-li rozšíření:

- Password - přidává možnost si změnit přihlašovací heslo.
- Managesieve - možnost nastavení vlastních filtrů pro příchozí poštu přímo v RoundCube. Jejich vytvoření navíc není nutné psát přímo v jazyce Sieve.
- Markasjunk - přidává do panelu nástrojů tlačítko pro označení / odznačení emailu jako spam.
- Userinfo - vypíše základní informace o právě přihlášeném uživateli.
- Zipdownload - pro emaily, které obsahují více než jednu přílohu, přidává možnost stáhnout všechny přílohy najednou jako jeden zip soubor.
- Archive - přidává do panelu nástrojů tlačítko pro přesunutí emailu do složky archiv.
- Contextmenu - přidává kontextovou nabídku při pravém kliknutí na prvky RoundCube a na emaily.
- Emoticons - možnost přidat emotikony do zprávy typu html.
- Keyboard_shortcuts - možnost používat přednastavené klávesové zkratky.
- Message_highlight - přidává možnost barevného označení emailu podle nastaveného odesílatele, příjemce, kopie nebo předmětu.
- Thunderbird_labels - přidává možnost stejného barevného označení emailu jako je tomu v emailové aplikaci Thunderbird.
- Spam_filter - přidává možnost výběru pravidel použitých pro filtraci spamu. Plugin byl vytvořen v rámci této práce, více podrobností v samostatné kapitole 2.4.3 Spam filter plugin.

2.2.7 Postfix

Odesílání emailu zajišťuje Postfix a probíhá ve dvou krocích, od uživatele k serveru a od serveru k dalšímu serveru. Každý krok se provádí rozdílným způsobem.

Server musí přijmout a odeslat pouze emaily od ověřených uživatelů. K ověření musí uživatel společně s emailem poslat i své přihlašovací údaje. Server tyto údaje zkontroluje a v případě jejich správnosti email odešle na další emailový server. Pokud jsou uvedené údaje chybné nebo je uživatel nezaslal, musí server email odmítnout. Hlavním důvodem je, že pokud není ověřena identita uživatele, pak není jisté, kdo je skutečným odesílatelem. Útočníci by mohli server zneužít pro odesílání spamu pod libovolnou emailovou adresou. Ostatní servery by pak tento server označily za zdroj rozesílání spamu a umístily by jej na černou listinu. Emaily, i od ověřených uživatelů, pocházející ze serveru umístěném na černé listině, jsou pak ve většině případů odmítány ostatními servery.

Ačkoli Postfix samotný nabízí ověření uživatele, je ověření přenecháno na Dovecot, který má již vše nakonfigurováno. Stačí tedy Postfix nastavit tak, aby před odesláním emailu předal patřičné údaje Dovecot, ten provede ověření uživatele a výsledek vrátí Postfix. Ten pak podle vráceného výsledku email přijme a odešle nebo zamítne.

2.2.8 Rspamd

S aktuálním nastavením již server dokáže přijímat a odesílat emaily. To však nestačí, neboť většinu přijatých emailů bude tvořit spam. Z tohoto důvodu je nezbytné spam filtrovat. K tomu je využit Rspamd, který pro každý email spočítá tzv. spam skóre. Čím je skóre vyšší, tím pravděpodobněji se jedná o spam a naopak. Jak Rspamd funguje, jaké nabízí možnosti a jak je na serveru implementován, jsou uvedeny v samostatné kapitole 2.4 Filtrace spamu.

Postfix propojíme s Rspamd a nastavíme tak, aby mu každý email předal ke kontrole. Rspamd provede řadu testů a vrátí email Postfix. Ten pak podle výsledku pokračuje stanoveným postupem. Kontrola se provádí u všech příchozích emailů i odchozích emailů. Zaprvé není žádnoucí, aby některý z uživatelů na serveru mohl rozesílat spam, ať už cíleně či nevědomě. Pokud bychom uživatele nezastavili, byla by to chyba serveru a ten by se rychle dostal na černou listinu. Za druhé Rspamd pozná, že se jedná o odchozí email a má k takovému emailu odlišný přístup. Například neprovádí kontrolu DKIM (DomainKeys Identified Mail) podpisu, naopak emailu DKIM podpis přidá. Rspamd také nabízí webové rozhraní pro snadný přístup a přehled a nabízí některá nastavení změnit.

2.3 Bezpečnost

Internet sám o sobě není bezpečný a stejně tak ani email a vše kolem něj. To se však se správným nastavením dá změnit. Chceme zajistit bezpečnost samotného serveru, komunikace mezi klientem a serverem a samozřejmě také email samotný.

2.3.1 Bezpečnost klient-server komunikace

Uživatel (klient) komunikuje se serverem při zobrazení a odeslání emailu. Pokud využívá našeho webového klienta RoundCube, pak navíc komunikuje i s naším webovým serverem. Všechny uvedené komunikace je nutné zabezpečit.

Komunikace s webovým serverem

Pro zabezpečení komunikace mezi uživatelem a webovým serverem je nutné přenášená data šifrovat. Přesně to umožňuje protokol HTTPS, k jehož fungování se musí nejprve vytvořit certifikát. K tomu je využit program Certbot, který si o certifikát zažádá na serveru Let's Encrypt a po jeho získání jej uloží na náš server. Certifikát má tříměsíční platnost. To je z důvodu průběžného mazání starých, již nepoužívaných, certifikátů. Certbot se stará i o znovuobnovování certifikátu. Vždy jeden měsíc před vypršením mu obnoví platnost na další tři měsíce. S platným certifikátem je již možné provádět komunikaci i přes protokol HTTPS. Aby byla jistota, že žádný uživatel nebude komunikovat na portu 80, kde se používá nezabezpečený protokol HTTP, je nastaveno automatické přesměrování z portu 80 na port 443. Komunikace na portu 443 probíhá vždy skrze protokol HTTPS.

Zobrazení emailu

Aby si uživatel mohl zobrazit emaily ve webovém klientovi, případně v nainstalované aplikaci, musí se emaily ze serveru nějak získat. K tomuto účelu slouží dva protokoly, IMAP a POP3. Protokol IMAP (Internet Message Access Protocol) pouze zobrazuje emaily, které jsou uloženy na serveru. POP3 (Post Office Protocol verze 3) emaily ze serveru stáhne do zařízení, kde se uživatel přihlásil a ze serveru je smaže. Na našem serveru je k dispozici pouze protokol IMAP, takže si uživatel může zobrazovat emaily na různých zařízeních, protože samotný email je uložen stále na serveru.

IMAP v základu nevyžaduje a nepoužívá šifrování, takže je tu potencionální riziko odposlechnutí nezašifrovaných dat. Na server byla proto přidána i zabezpečená verze protokolu IMAPS, kde komunikace probíhá celou dobu šifrovaně pomocí SSL/TLS. Protože někteří emailový klienti nepodporují protokol IMAPS, je možné pro komunikaci využít i protokol IMAP. Na serveru je však nastaveno, aby v případě využití IMAP protokolu, bylo ihned po nazávání spojení, ještě před výměnou jakýchkoli dat, zahájeno šifrování pomocí STARTTLS. Ve výsledku jsou tedy v případě obou protokolů IMAP a IMAPS všechna přenášená data šifrovaná.

Odesílání emailu

Posledním krokem je zabezpečení odesílání emailu. K odeslání emailu uživatelem se používá protokol Submission nebo SMTPS, k odeslání emailu serverem pak protokol SMTPS nebo SMTP (Simple Mail Transfer Protocol). Způsob zabezpečení komunikace protokoly Submission a SMTPS je stejný jako u IMAP

a IMAPS. Komunikace protokolem SMTPS již při navázání spojení vyžaduje šifrování pomocí SSL/TLS protokolu a musí zůstat šifrovaná až do ukončení spojení. Submission protokol nevyžaduje při navázání spojení šifrování. Opět je však nastaveno, aby po navázání spojení, ještě před výměnou dat, bylo zahájeno šifrování pomocí STARTTLS. Data jsou v obou případech šifrována a zabezpečena.

Protokol SMTP na portu 25 slouží pro odesílání emailů pouze mezi servery a pouze v případě, že jeden ze serverů nepodporuje komunikaci přes protokol SMTPS. Komunikace mezi uživatelem a serverem není přes protokol SMTP na portu 25 povolena.

2.3.2 Bezpečnost serveru

Server je připojen k internetu a má veřejnou IP adresu. To stačí k tomu, aby se útočníci, potažmo jejich boti, snažili získat přístup na tento server. Pokud by se jim to podařilo, mohli by se dostat k uloženým emailům uživatelů, smazat či zašifrovat všechna data na serveru nebo jej zneužít pro rozesílání spamu.

Nejjednodušší a také nejefektivnější obranou je firewall, díky němuž je možné uzavřít všechny porty, které server nepotřebuje k postykování emailové služby. V tabulce 1 je uveden seznam všech otevřených portů na serveru, včetně dostupných adres, běžících služeb a jejich popisu.

Na serveru je také nainstalovaný software fail2ban, který zajišťuje ochranu před bruce force attack (útok hrubou silou). Jedná se o typ útoku, kdy se útočník snaží opakovaně přihlásit do systému a postupně zkouší různé přihlašovací údaje. Například v případě Redis může být provedeno až milion pokusů za sekundu. Útok je zpravidla prováděn z jedné IP adresy a přesně toho se využívá při obraně.

Fail2ban neustále kontroluje neúspěšné pokusy o přihlášení. Pokud zjistí více nezdařených pokusů z jedné IP adresy v krátkém časovém úseku, danou IP adresu dočasně zablokuje. Při aktuálním nastavení je po 5 neúspěších daná IP adresa zablokována na 10 minut. Útočník sice může i nadále útočit, počet jeho pokusů je však natolik omezen, že je pro něj prakticky nemožné se pouhou metodou brute force dostat do systému.

2.3.3 Bezpečnost emailové služby

Ačkoli je server i komunikace mezi uživatelem a serverem dobře zabezpečena, poslední důležitý krok stále chybí. Kdokoli se může vydávat za uživatele naší domény na jakémkoli jiném emailovém serveru. Nikde totiž není uvedeno, že naše doména běží pouze na našem serveru. To zajistíme přidáním patřičných DNS záznamů naší doméně.

Sender Policy Framework (SPF)

SPF záznam udává IP adresy serverů, které spravují danou doménu. Každý odeslaný email má v hlavičce záznam o návratové emailové adrese. Na tuto adresu

Tabulka 1: Seznam otevřených portů

Port	Adresa	Služba	Popis
22/tcp	`public`	`SSH`	Vzdálený přístup na server.
25/tcp	`public`	`SMTP`	Komunikace server-server při odesílání emailu.
80/tcp	`public`	`HTTP`	Komunikace klient-server při přístupu k webovému klientovi.
143/tcp	`public`	`IMAP`	Komunikace klient-server při zobrazování emailu.
443/tcp	`public`	`HTTPS`	Zabezpečená komunikace klient-server při přístupu k webovému klientovi.
465/tcp	`public`	`SMTPS`	Zabezpečená komunikace klient-server nebo server-server při odesílání emailu.
587/tcp	`public`	`Submission`	Komunikace klient-server při odesílání emailu.
993/tcp	`public`	`IMAPS`	Zabezpečená komunikace klient-server při zobrazování emailu.
3306/tcp	`localhost`	`MySQL`	Databáze domén, emailových adres a aliasů.
4190/tcp	`public`	`Sieve`	Nastavování vlastních filtrů pomocí sieve pravidel.
6379/tcp	`localhost`	`Redis`	Databáze pro Rspamd při testování emailů.
11332/tcp	`localhost`	`Rspamd`	Komunikae s Postfix.
11333/tcp	`localhost`	`Rspamd`	Testování emailů.
11334/tcp	`localhost`	`Rspamd`	Řadič Rspamd.

se zasílá upozornění o nedoručení a odmítnutí emailu. Při kontrole SPF se vezme doména z návratové adresy a vyžádají se DNS záznamy pro tuto doménu. Je-li pro doménu uveden SPF záznam, pak obsahuje IP adresy, ze kterých se mohou emaily dané domény posílat. Provede se porovnání IP adres ze SPF záznamu s IP adresou serveru, ze kterého byl email skutečně odeslán. V případě shody IP adres proběhla validace SPF v pořádku. Neúspěšná validace může nastat z důvodu chybějícího nastavení SPF pro danou doménu nebo podvržení emailu.

DomainKeys Identified Mail (DKIM)

Pro používání DKIM je potřeba vytvořit soukromý a veřejný klíč. Soukromý klíč je uložen na serveru a pomocí něj se šifruje podpis každého odeslaného emailu. Veřejný klíč se umístí mezi DNS záznamy domény. Při přijetí emailu s DKIM podpisem se vyžádají DNS záznamy pro doménu odesílatele. Ze záznamů se vezme veřejný klíč a použije se pro dešifrování a ověření podpisu.

Domain-based Message Authentication (DMARC)

DMARC využívá kombinaci obou výše zmíněných validací. Navíc provádí kontrolu porovnáním domén. Pro SPF porovná doménu odesílatele s doménou návratové adresy. Pro DKIM porovná doménu odesílatele s doménou uvedenou v DKIM záznamu v DNS záznamech. Aby byla validace pomocí DMARC úspěšná, musí projít validace SPF i DKIM a navíc musí úspěšně projít alespoň jedno porovnání.

Co se stane s emailem, kde některá z validací SPF, DKIM nebo DMARC selhala, záleží na nastavení jednotlivých emailových serverů.

2.4 Filtrace spamu

Zajištění filtrace spamu je nedílnou součástí každého emailového serveru. V rámci této práce jsou navíc požadavky na automatické učení se spam filtru od akcí uživatelů. Každý uživatel má mít také možnost si vybrat, jakým způsobem chce spam filtrovat.

2.4.1 Princip fungování

Filtraci spamu na serveru zajišťuje Rspamd. Ten má neustále spuštěná vlákna pro komunikaci s Postfix a testování emailů. Rspamd má již z instalace celou řadu symbolů rozdělených do jednotlivých skupin. Příklad některých skupin: subject, statistics, header, url, spf. Skupina subject obsahuje symboly:

- SUBS_ALL_CAPS,
- LONG_SUBJ,
- URL_IN_SUBJECT.

Pro rozlišení názvů skupin a symbolů zavedl Rspamd syntaxi, kde se názvy skupin se píší malými písmeny a názvy symbolů velkými písmeny. Stejná syntaxe je dodržena i v této práci.

Rspamd má stovky předvytvořených symbolů rozdělených do desítek skupin. Jednotlivé symboly a skupiny symbolů je možné povolit či zakázat. Při kontrole emailu pak vypočítá Rspamd za základě všech povolených symbolů tzv. spam skóre. Čím je skóre vyšší, tím spíše se dá email považovat za spam a naopak. Výsledné skóre a symboly, které se pro výpočet použily, jsou přidány do hlavičky emailu.

Každý symbol má stanovenou hodnotu a podmínku. Pokud je podmínka symbolu splněna, hodnota symbolu se použije pro výpočet celkového spam skóre emailu. Například podmínka symbolu URL_IN_SUBJECT testuje, zda předmět emailu obsahuje url adresu. Hodnota tohoto symbolu jsou 4 body. Pokud je tedy

podmínka splněna, spam skóre emailu bude symbolem URL_IN_SUBJECT navýšeno o 4 body.

Rspamd umožňuje přídání i vlastních symbolů a skupin. Všem symbolům, předvytvořeným i uživatelsky vytvořeným, je možné upravit jejich hodnotu. Všechny symboly mají pevně stanovené hodnoty, které se v případě splnění podmínky aplikují v plné výši, nebo se v případě nesplnění podmínky ignorují. Existuje však jedna výjimka a tou jsou symboly ve skupině statistics. Tyto symboly nemají pevně stanovenou podmínku, ale je zde využita statistika na založena na Bayesově klasifikátoru. [4] Ten vychází z Bayesovy věty o podmíněných pravděpodobnostech. Bayesův klasifikátor minimalizuje pravděpodobnost chybné klasifikace. [5] V případě filtrace emailů se používá pro snížení pravděpodobnosti chybného označení emailu jako ham nebo spam. Ham je označení pro dobrou poštu.

Kontrolovaný email se porovnává s již dříve doručenými a zkontrolovanými emaily na základě tokenů. Token udává pravděpodobnost výskytu spamu. Celkově se pak pravděpodobnost určuje podle pravděpodobností všech tokenů a podle pravděpodobnosti výskytu jednotlivých tokenů. Výsledná pradvěpodobnost pak udává, jak moc se jedná o ham nebo spam. Hodnoty symbolů ze skupiny statistics udávají minimum a maximum. Podle výsledné pravděpodobnosti pak může email od těchto symbolů dostat jakoukoli hodnotu mezi minimem a maximem.

2.4.2 Učení se spam filtru a metriky

V základu jsou ve skupině statistics pouze dva symboly BAYES_HAM a BAYES_SPAM, udávající dolní (ham) a horní (spam) hranici skóre. Požadavkem však je, aby uživatel mohl k filtraci využívat globální nebo uživatelský filtr, tedy statistiku všech uživatelů nebo pouze svou vlastní. K zajištění tohoto požadavku bylo proto potřeba vytvořit další dva symboly, BAYES_HAM_USER a BAYES_SPAM_USER. V nastavení Rspamd se stanovilo, aby se pro původní symboly užívala jedna společná statistika od všech uživatelů a pro nové symboly s _USER se pro každého uživatele zvlášť vytvořila nová statistika.

Pomocí Dovecot je pak zajištěno, že když uživatel označí email jako spam, je tato akce zanesena vždy do globální statistiky a navíc do statistiky daného uživatele. Stejně je tomu naopak. Pokud uživatel email označený jako spam odznačí, pak se do statistik zanese, že se tento email považuje za ham. Navíc Rspamd nabízí možnost automatického učení stanovením hranice pro učení se hamu a spamu. Pokud je výsledné spam skóre emailu nižší nebo rovno hranici hamu, pak se do statistik automaticky uloží, že se email považuje za ham, aniž by uživatel provedl jakoukoli akci. Stejně tak je tomu u hranice spamu, kde se za spam považuje email, jehož spam skóre je vyšší nebo rovno hranici spamu.

Rspamd má stanovené metriky, podle kterých je možné provést s jednotlivými emaily zvolené akce. Metriky jsou čtyři uspořádané od nejvyšší priority:

1. reject,
2. rewrite subject,
3. add header,
4. greylist.

Pro každou metriku je možné uvést její hodnotu. Pokud je hodnota překročena, aplikuje se akce dané metriky. Pokud je překročena hodnota více metrik současně, aplikuje se metrika s nejvyšší prioritou. Hodnota metrik se porovnává se spam skóre emailu.

Reject

Metrika *reject* udává, že při jejím překročení má být příchozí email odmítnut. V požadavcích práce je však stanoveno, že mají být všechny emaily uživateli doručeny. Protože metrika *reject* musí mít nějakou hodnotu, byla hodnota nastavena na 150 bodů. Takovou hranici žádný email nikdy nepřekročí, takže je metrika prakticka vypnutá.

Rewrite subject

Další metrikou je *rewrite subuject*, která při překročení přepisuje předmět emailu. Přepisování předmětu není vyžadováno, proto je metrika vypnutá.

Add header

Velmi důležitá a nezbytná metrika. Při překročení přidává do hlavičky emailu záznam *X-Spam: Yes*. Dovecot má nastaveno, aby každý email před uložením do uživatelovi schránky zkontroloval na přítomost záznamu *X-Spam: Yes*. Pokud se záznam v hlavičce vyskytuje, pak email uloží do složky spam. V opačném případě jej standardně uloží do složky příchozí pošta. Uložení emailu jiným způsobem je možné provést nastavením vlastních filtrů. Hodnota metriky *add header* je aktuálně nastavena na 6 bodů, takže jakýkoli email s výsledným spam skóre rovno nebo výše 6 bodů je označen jako spam a není-li uvedeno jinak, doručen do složky spam.

Greylist

V případě aplikace této metriky se příchozí email dočasně odmítne. Pokud proběhne pokus o znovudoručení emailu, pak je již email přijat. Dočasné odmítnutí má sloužit jako základní ochrana proti spamu. Většina emailových serverů, které rozesílají spam, nemá frontu odesílaných emailů, takže pokud je odeslaný email odmítnut jiným serverem, daný email se již znovu nepokusí odeslat. Nevýhodou je, že i emaily, které nejsou spam, mohou být dočasně odmítnuty a pokus

o jejich znovu doručení může proběhnout i za několik hodin. Za jakou dobu proběhne další pokus o doručení emailu záleží na nastavení odesílajícího serveru. Z uvedeného důvodu se metrika greylist na tomto serveru nepoužívá.

2.4.3 Spam filter plugin

Aby si uživatel mohl navolit, jakou filtraci spamu chce používat, byl vytvořen plugin pro RoundCube přidávající volby výběru spam filtru. Plugin je dostupný v Nastavení - Vlastnosti - Spam filtr. Pro filtraci se pak používá kombinace uživatelských, globálních a statických pravidel. Uživatelská pravidla využívají statistiku pouze konkrétního uživatele, zatímco globální pravidla využívají statistiku všech uživatelů na serveru. V rámci pluginu si uživatel může vybrat, jaká pravidla chce mít zapnutá. Jaké symboly pravidla povolují je uvedeno v tabulce 2.

Tabulka 2: Seznam možností spam filtru

Možnost	Popis
Globální pravidla	Povoluje globální symboly BAYES_SPAM a BAYES_HAM.
Uživatelská pravidla	Povoluje uživatelské symboly BAYES_SPAM_USER a BAYES_HAM_USER.
Statická pravidla	Povoluje všechny ostatní symboly.

Poté, co si uživatel zvolí, jaká pravidla chce používat a klikne na tlačítko uložit, spustí se na serveru skript. Ten zjistí, jaká pravidla byla vybrána a jakým uživatelem. Skript následně spustí program, který patřičným způsobem upraví nastavení Rspamd. Na závěr skript provede restart Rspamd, aby se aplikovalo nové nastavení.

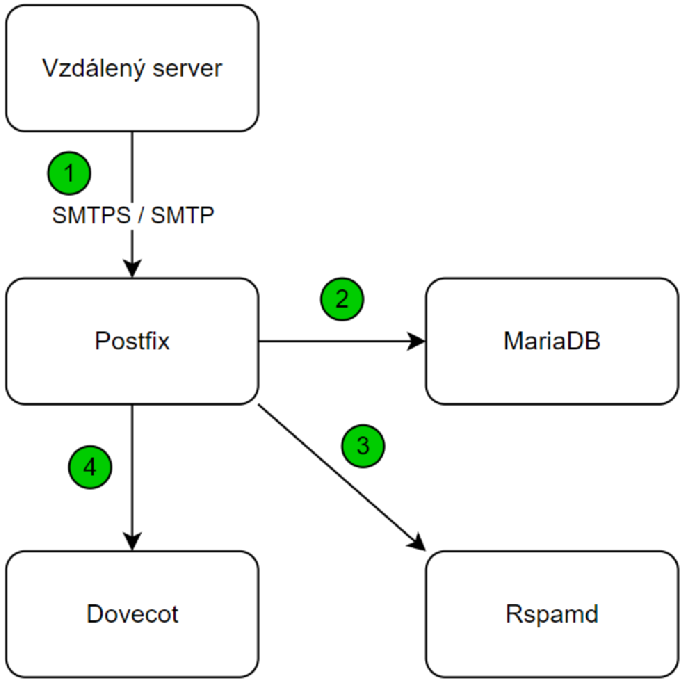
2.5 Průběh emailové komunikace

Všechny prvky, které komunikaci zajišťují, již byly samostatně popsány. Nyní se podíváme na komunikaci jako na celek. Jak probíhá přijetí emailu od uživatele na jiném serveru, jakým způsobem se odesílá email uživatelem na našem serveru a co vše musí proběhnout, když si chce uživatel prohlédnout své emaily.

2.5.1 Přijetí emailu

Na obrázku 2 je znázorněn postup při přijetí emailu z jiného serveru. V prvním kroku naváže vzdálený server spojení s naším serverem pomocí protokolu SMTPS na portu 465. Pokud jeden nebo oba servery nepodporují protokol SMTPS, pak je spojení navázáno na portu 25 protokolem SMTP. Přijetí emailu na našem serveru řeší Postfix. Ten se podívá do databáze a zjistí, zda příjemce emailu je uživatel na

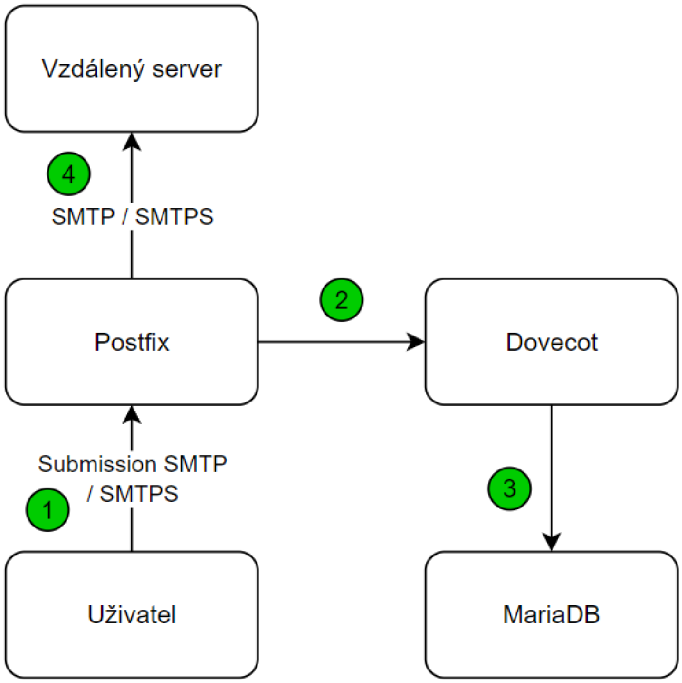
našem serveru. Je-li tomu tak, pak předá email Rspamd. Ten provede kontrolu emailu podle uživatelem zvolených pravidel, vypočítá spam skóre a vrátí email Postfix. V posledním kroku je email předán Dovecot, aby jej uložil do schránky správného uživatele a do příslušné složky.



Obrázek 2: Přijetí emailu

2.5.2 Odeslání emailu

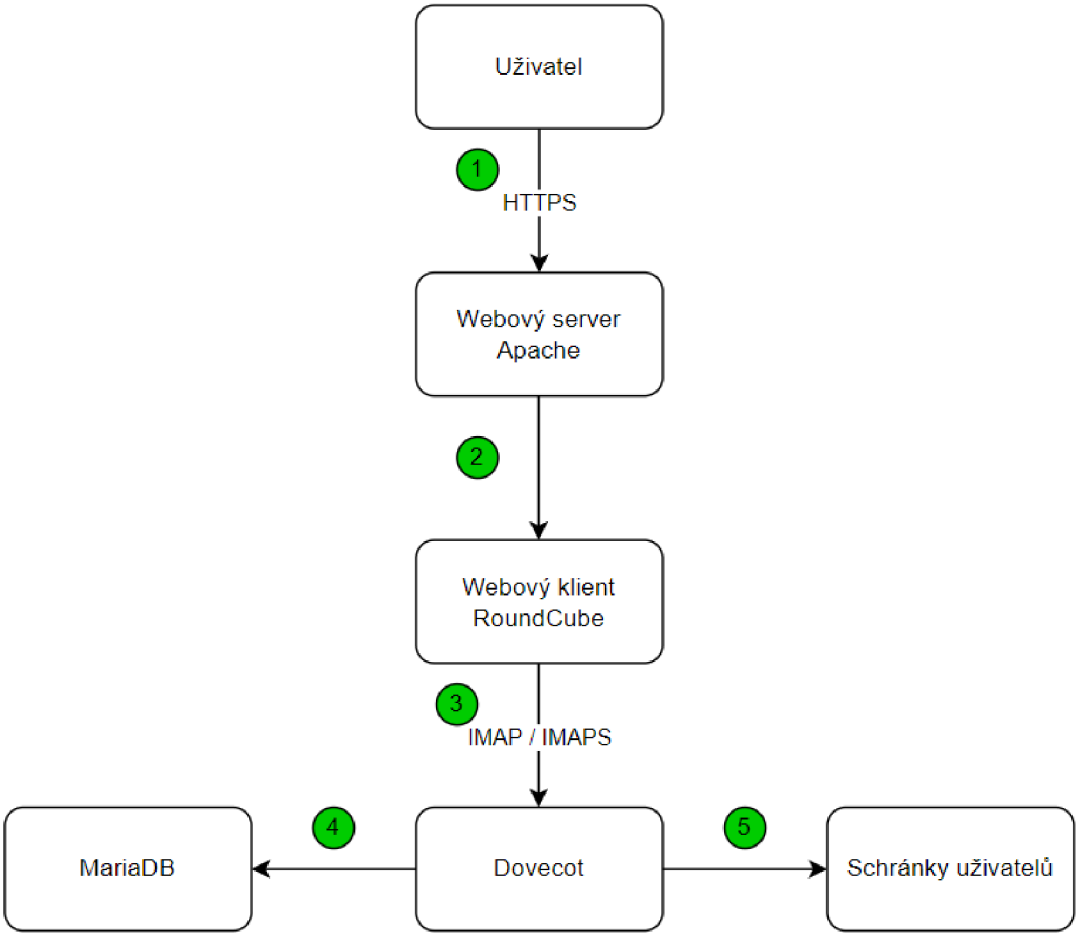
Postup při odeslání emailu je vyobrazen na obrázku 3. Uživatel naváže komunikaci s Postfix pomocí protokolu Submission na portu 587 nebo SMTPS na portu 465. Společně s emailem posílá také své přihlašovací údaje. Postfix předá uživatelské údaje k ověření Dovecot, který je porovná s údaji uloženými v databázi. Pokud vše souhlasí, naváže Postfix spojení se vzdáleným serverem pomocí protokolu SMTPS na portu 465, případně SMTP na portu 25, a email odešle.



Obrázek 3: Odeslání emailu

2.5.3 Zobrazení emailu

Na obrázku 4 je vizualizován postup pro zobrazení emailů uživatele skrze našeho webového klienta RoundCube. Nejprve uživatel zahájí komunikaci s webovým serverem na portu 443 protokolem HTTPS. Webový server spustí skripty RoundCube a vyzve uživatele k zadání přihlašovacích údajů. Ve třetím kroku probíhá komunikace mezi RoundCube a Dovecot skrze protokol IMAPS na portu 993. Zde žádá webový klient o zobrazení emailů pro daného uživatele. Dovecot nejprve zkontroluje zadané přihlašovací údaje s údaji v databázi. Pokud kontrola proběhne úspěšně, najde schránku uživatele a předá RoundCube potřebná data k zobrazení všech emailů, která jsou na serveru ve schránce uživatele dostupná.



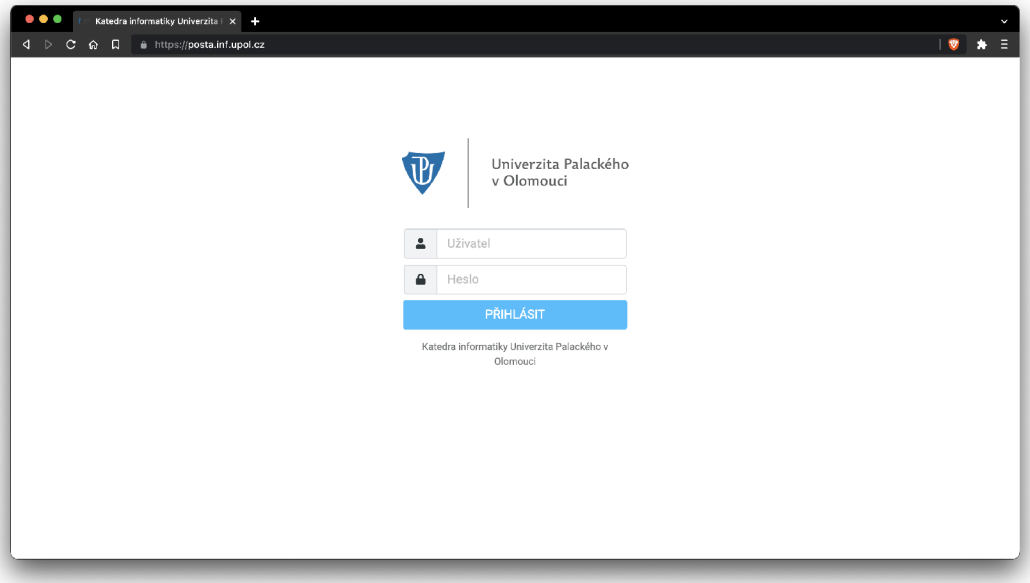
Obrázek 4: Zobrazení emailu

3 Uživatelská příručka

V kapitole je popsáno přihlášení a základní používání emailové schránky v přednastaveném webovém klientovi RoundCube. Uveden je také postup, jak si uživatel může změnit některá nastavení, navolit si svůj spam filtr a také vytvořit pravidla pro filtraci příchozí pošty. Poslední podkapitola je zaměřena na přístup a používání schránky pomocí jiného emailového klienta.

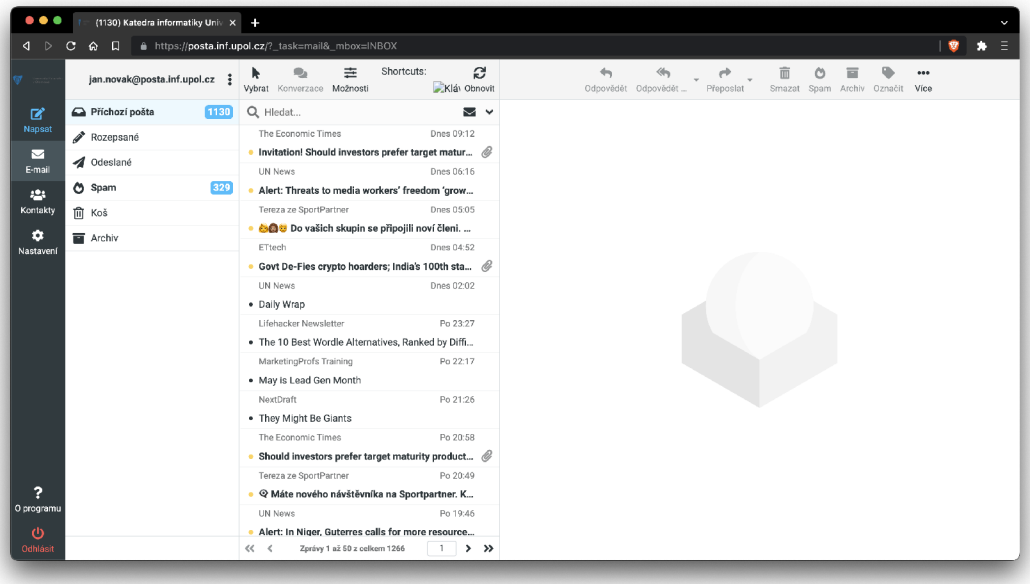
3.1 Používání emailové schránky

Přihlášení se ke své schránce je možné provést skrze webového emailového klienta RoundCube na adrese `https://posta.inf.upol.cz/`. Na obrázku 5 je zobrazena úvodní přihlašovací stránka. Pro přihlášení je nezbytné vyplnit uživatelské jméno včetně domény (např. *jan.novak@posta.inf.upol.cz*) a heslo. Po vyplnění obou údajů je možné se přihlásit kliknutím na tlačítko „Přihlásit“. V rámci této práce se používá doména *posta.inf.upol.cz*, v případě ostrého nasazení serveru by se pak doména změnila na *inf.upol.cz*.



Obrázek 5: Přihlašovací stránka RoundCube

Po přihlášení se zobrazní přehled doručených emailů, jak je vidět na obrázku 6. Úplně vlevo se nachází menu, které nabízí šest možností. Postupně od shora dolů jsou možnosti: napsat nový email, přehled všech emailů, seznam kontaktů, nastavení schránky, bližší informace o programu a odhlášení uživatele.



Obrázek 6: Přehled emailů

Ve sloupci mezi menu a přehledem emailů je seznam všech aktivních složek. Přehled a popis, k čemu jednotlivé složky slouží, je uveden v tabulce 3. Na obrázku i v tabulce jsou uvedeny pouze předvytvořené složky, uživatel má možnost si vytvořit a přidat i vlastní složky, jak je uvedeno dále.

Tabulka 3: Seznam složek

Složka	Popis
Příchozí pošta	Pokud není stanoveno jinak, jsou zde ukládány všechny příchozí emaily.
Rozepsané	Pokud uživatel vytvoří nový email, ale neodešle jej, uloží se email do této složky pro pozdější dopsání a odeslání.
Odeslané	Přehled emailů, které uživatel odeslal.
Spam	V této složce se nachází emaily, které byly spam filtrem vyhodnoceny jak spam. Emaily v této složce se automaticky trvale smažou po 30 dnech.
Koš	Všechny smazané emaily jsou přesunuty do složky Koš. Po 30 dnech se automaticky trvale smažou.
Archiv	Zde jsou uloženy archivované emaily. Archivace se provádí ručním přesunem emailu do složky Archiv.

S každým emailem je možné provést několik akcí, například odpovědět, přeposlat, smazat, označit jako spam, archivovat. Zvolení akce je možné provést v horní liště nad zobrazeným emailem nebo v kontextové nabídce, která se zobrazí při pravém kliknutí na konkrétní email.

3.2 Základní nastavení

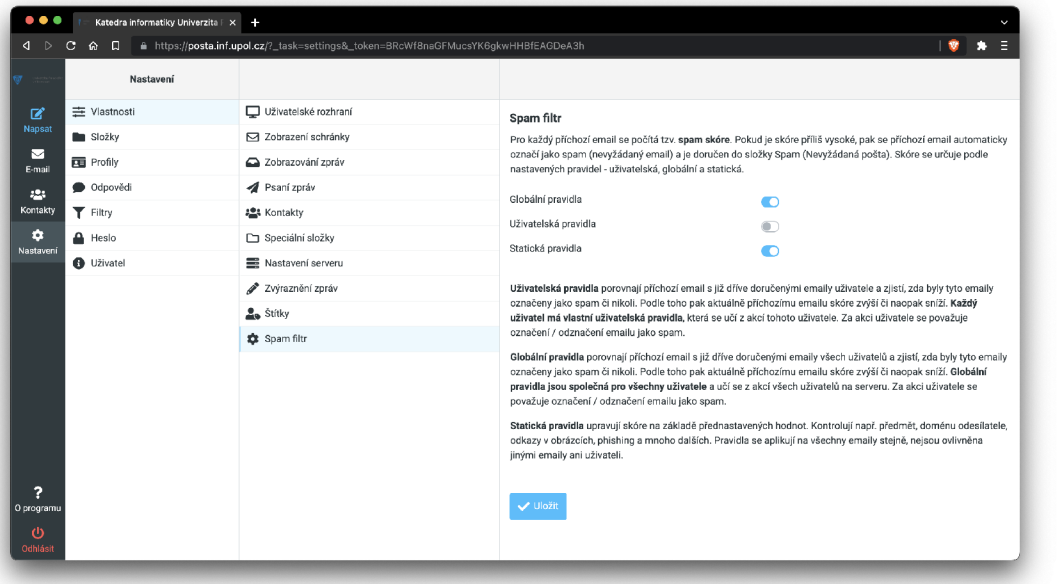
Kliknutím na položku *Nastavení* v levém menu se zobrazí všechna dostupná nastavení emailové schránky. V záložce *Vlastnosti* v sekci *Uživatelské rozhraní* má uživatel možnost si změnit jazyk i vzhled celé schránky.

V záložce *Složky* je k dispozici vytváření a mazání uživatelských složek a také možnost smazat všechny emaily z vybrané složky.

Změna přihlašovacího hesla se provádí v záložce *Heslo*. Nastavení nového hesla se provede zadáním nového hesla a zopakováním nového hesla. Všem uživatelům se důrazně doporučuje, aby si po prvním přihlášení své heslo změnili.

3.3 Nastavení filtrace spamu

V *Nastavení* v záložce *Vlastnosti* se nachází sekce *Spam filtr*. V této sekci je možné vybrat, která pravidla má spam filtr používat při kontrole všech příchozích emailů. Celkem jsou k dispozici tři druhy pravidel - globální, uživatelská a statická. Je možné si nastavit jejich libovolnou kombinaci, včetně vypnutí všech pravidel. Obrázek 7 znázorňuje nastavení pro používání globálních a statických pravidel.



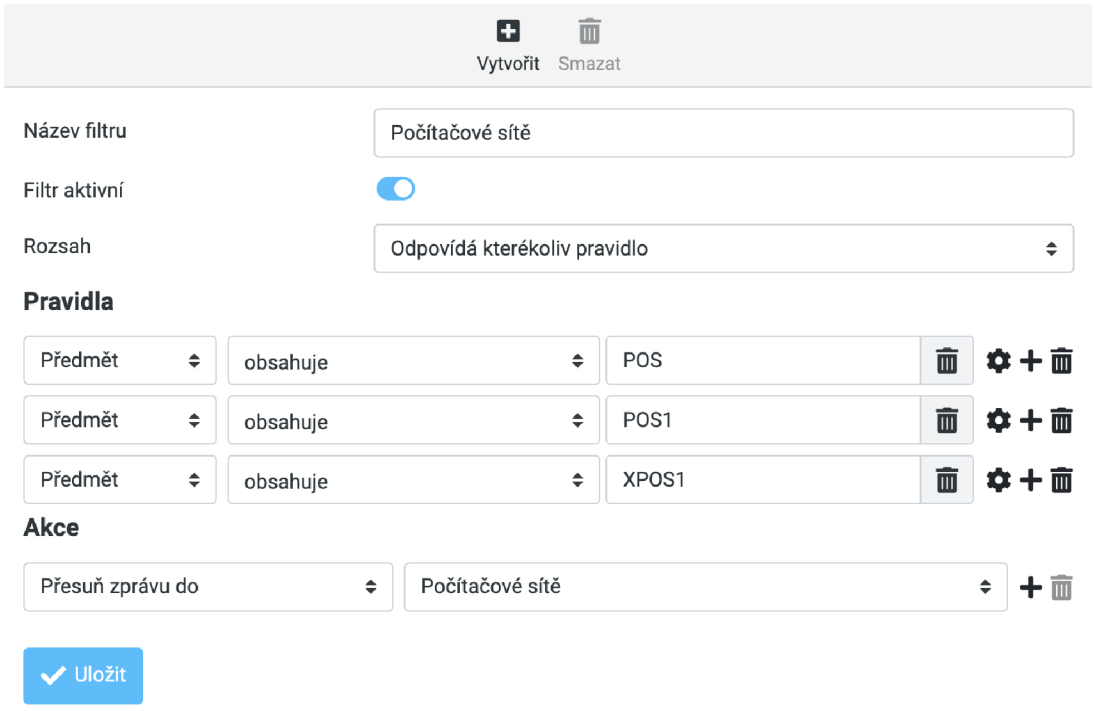
Obrázek 7: Nastavení spam filtru

Na základě používaných pravidel se pro každý příchozí email vypočítá spam skóre. Pokud je skóre příliš vysoké, email je označen jako spam a při doručení je automaticky umístěn do složky *Spam*. Globální a uživatelská pravidla se navíc automaticky s každým příchozím emailem učí, zda se jednalo o dobrý email nebo spam. Pokud filtr chybně označí email jako spam, nebo jej naopak neoznačí když má, uživatel má možnost zasáhnout. Stačí daný email ručně označit nebo odznačit jako spam, tím se email zařadí do příslušné složky. Zároveň si filtr, respektive globální a uživatelská pravidla, tuto akci zapamatují. Vypnutí všech pravidel má za následek nefiltrování příchozích emailů a jejich ukládání do složky *Příchozí pošta*.

Výsledné spam skóre lze vyčíst z hlavičky každého emailu. V hlavičce jsou také uvedeny symboly, které se použily pro výpočet skóre.

3.4 Vlastní filtry pro příchozí poštu

Pro příchozí emaily lze nastavit vlastní filtry podle zadaných pravidel. Díky těmto filtrům je možné ukládat příchozí emaily do jiné než výchozí složky, přesměrovat emaily na jinou emailovou adresu, nastavit automatickou odpověď a další. Předhled filtrů, jejich vytvoření a editace jsou možné v záložce *Filtry*.



Obrázek 8: Nastavení vlastního filtru

Obrázek 8 znázorňuje příklad vytvoření filtru s názvem *Počítačové sítě*. Jako rozsah je zvolena varianta *Odpovídá kterékoliv pravidlo*, tedy aby se akce filtru

vykonala, musí být splněno alespoň jedno z uvedených pravidel. Nastavená pravidla kontrolují, zda se kdekoli v předmětu emailu nachází výraz *POS*, *POS1* nebo *XPOS1*. Pokud se v předmětu vyskytuje alespoň jeden z uvedených výrazů, provede se zvolená akce. V tomto případě akce udává, že se má příchozí email uložit do složky *Počítačové sítě*.

3.5 Přístup z emailového klienta

Emailovou schránku je možné používat i v emailových klientech jako např. Thunderbird, Spark nebo Outlook. Pro zprovoznění je nutné nastavit servery pro příchozí a odchozí poštu. Tabulka 4 uvádí, jaké jsou možnost nastavení serverů. Pro každý server je možné zvolit jednu ze dvou možností. Někteří emailový klienti nemusí podporovat obě možnosti. Do atributu *username* se uvádí email uživatele, např. *jan.novak@posta.inf.upol.cz*.

Tabulka 4: Nastavení serverů pro příchozí a odchozí poštu

	Server pro příchozí poštu		Server pro odchozí poštu	
	Možnost 1	Možnost 2	Možnost 1	Možnost 2
Protokol	IMAP	IMAP	-	-
Hostname	posta.inf.upol.cz		posta.inf.upol.cz	
Port	143	993	587	465
Connection security	STARTTLS	SSL/TLS	STARTTLS	SSL/TLS
Authentication method	Normal passwrod		Normal passwrod	
Username	[username]@posta.inf.upol.cz		[username]@posta.inf.upol.cz	

Využívání emailových klientů má však svá omezení. Nelze v nich změnit přihlašovací heslo, upravit pravidla spam filtru ani vytvářet a editovat vlastní filtry pro příchozí poštu. Pro změnu těchto parametrů je nutné se přihlásit přímo do webového klienta RoundCube. Učení spam filtru v emailových klientech funguje stejně jako v RoundCube, tedy přesunem emailu do/ze složky *Spam*.

4 Programátorská dokumentace

V úvod kapitoly popisuje nejdůležitější konfiguraci použitého softwaru. Další část se věnuje programům a skriptům, vytvořených pro zajištění veškeré funkčnosti, kterou práce vyžaduje. Závěr je zaměřen na rozšíření RoundCube, jenž bylo vytvořeno za účelem výběru spam filtru ze strany uživatele.

4.1 Konfigurace serveru

Server vytvořený v rámci této práce využívá již hotový software, ten je však potřeba správně nakonfigurovat, aby splňoval stanovené požadavky. Všechny změněné konfigurační soubory jsou uvedeny na přiloženém datovém médiu. Zde je zmínka pouze o vybraných důležitých nastavení.

Chceme, aby veškerou autentizaci prováděl Dovecot. Zdrojový kód 1 ukazuje, jak se tento požadavek aplikuje v konfiguračním souboru */etc/dovecot/conf.d/10-master.conf*. První část *service auth* udává, že má Dovecot provádět autentizaci i pro Postfix. Dále je nutné nastavit, kde bude Dovecot s Postfix komunikovat. To bylo nastaveno ve druhé části *service lmtp*. Postfix běží pouze v prostředí */var/spool/postfix/*, proto i všechny sokety, které chtějí s Postfix komunikovat, musí být umístěny někde v tomto adresáři. [6]

```
service auth {
  # Postfix smtp-auth
  unix_listener /var/spool/postfix/private/auth {
    mode = 0660
    user = postfix
    group = postfix
  }
}
service lmtp {
  unix_listener /var/spool/postfix/private/dovecot-lmtp {
    group = postfix
    mode = 0600
    user = postfix
  }
}
```

Zdrojový kód 1: Konfigurace Dovecot pro komunikaci a autentizaci s Postfix

Poslední krok bylo nastavit Postfix tak, aby předával emaily Dovecot k ověření. K tomu stačilo zadat do terminálu jeden příkaz:

```
postconf virtual_transport=lmtp:unix:private/dovecot-lmtp
```

Nastavení Roundcube v souboru */etc/roundcube/config.inc.php* zobrazeném ve zdrojovém kódu 2 udává způsob přihlášení uživatele k jeho emailové schránce.

Navíc také určuje SMTP server pro odesílání pošty. Je vybrán port 465 s protokolem SMTPS a šifrováním SSL/TLS.

```
// The IMAP host chosen to perform the log-in.
$config['default_host'] = 'ssl://posta.inf.upol.cz';

// SMTP server host (for sending mails).
$config['smtp_server'] = 'ssl://posta.inf.upol.cz';

// SMTP port. 465 for Implicit TLS
$config['smtp_port'] = 465;

// SMTP username
$config['smtp_user'] = '%u';

// SMTP password
$config['smtp_pass'] = '%p';
```

Zdrojový kód 2: Nastavení přihlašování do RoundCube a odesílajícího SMTP serveru

Poslední uvedené nastavení se týká filtrace spamu, kterou zajišťuje Rspamd. Aby bylo možné z hlavičky emailu zjistit výsledné spam skóre a jaké symboly se ve výpočtu použily, bylo potřeba hlavičku rozšířit. K tomu stačilo vytvořit soubor */etc/rspamd/local.d/milter_headers.conf* a do něj uvést jednoduchý parametr [7]:

```
extended_spam_headers = true;
```

Vytvoření vlastních symbolů pro učení se hamu a spamu jednotlivých uživatelů, bylo provedeno rozšířením již existující skupiny *statistics*. Pro rozšíření bylo potřeba vytvořit soubor */etc/rspamd/local.d/statistics_group.conf* a v něm uvést nové symboly. Pro každý symbol je nutné uvést jeho název a hodnotu (weight). Ukázka rozšíření je uvedena ve zdrojovém kódu 3.

4.2 Programy v C#

Pro správné fungování spam filtru jednotlivých uživatelů je potřeba mít vždy aktuální soubor */etc/rspamd/local.d/settings.conf* s nastavením Rspamd [8]. Pro úpravu tohoto souboru byly vytvořeny tři programy napsané v jazyce C#, které jsou umístěné v adresáři */etc/roundcube/plugins/spam_filter/cs*. Program *AddUserSetting.cs* vytvoří nový záznam se zadaným uživatelem a pravidly. Naopak *RemoveUserSetting.cs* najde a odstraní záznam se zadaným uživatelem. Poslední program *UpdateUserName.cs* aktualizuje jméno uživatele. Má dva vstupní parametry, staré a nové uživatelské jméno.

```
symbols = {
   "BAYES_SPAM_USER" {
      weight = 5.1;
      description = "Message probably spam, probability: ";
   }
   "BAYES_HAM_USER" {
      weight = -3.0;
      description = "Message probably ham, probability: ";
   }
}
```

Zdrojový kód 3: Vytvoření nových symbolů Rspamd pro uživatelskou filtraci

V ukázkovém zdrojovém kódu 4 je zobrazena funkce, která pro zadaného uživatele zjistí čísla řádků, kde jeho záznam začíná a kde končí. Využívá toho, že každý záznam uživatele je uložen ve stylu *user_jmenouzivatele*. Hranice záznamu uživatele se pak využívají při mazání záznamu. Nový záznam se přidává vždy na konec souboru.

Aby bylo možné tyto programy spustit, byly převedeny do spustitelných souborů, které jsou uloženy v adresáři */etc/roundcube/plugins/spam_filter/exe*. Klasické *exe* soubory nejsou na Linuxu spustitelné, proto byl převod souborů proveden pomocí softwaru mono-complete.

4.3 Skripty

Všechna přidávání, editování a mazání uživatelů a jejich spam filtrů zajišťují bash skripty v adresáři */etc/roundcube/plugins/spam_filter/sh*. Tyto skripty mají většinou jen pár řádků a slouží vždy pro jeden účel. Pokud je potřeba provést složitější akci, pak je nutné postupně spustit více skriptů. Spuštění skriptů provádí buď samotný uživatel skrze Spam filter plugin v RoundCube nebo správce serveru spuštěním vlastních správcovských skritpů.

Správce serveru má k dispozici celkem tři skripty umístěné v samostatném adresáři */etc/roundcube/plugins/spam_filter/admin_sh*, které zajistí vše, co je potřeba. Skript *add_user.sh* přidá nového uživatele do databáze a nastaví mu výchozí hodnotu spam filtru, což jsou zapnutá globální a statická pravidla. Pro změnu uživatelského jména slouží skript *update_user_name.sh*. Ten změní jméno uživatele v dababázi, zajistí také změnu jména v nastavení Rspamd a v posledním kroce přejmenuje adresář s emaily uživatele, aby se uživatel i po změně jméno dostal ke svým původním emailům. Poslední dostupný skript *delete_user.sh* slouží pro smazání uživatele. Vždy se vymaže uživatel z databáze a vymaže se také jeho nastavení v Rspamd. Volitelně při zadání druhého nepovinného parametru je pak možné smazat i adresář s emaily uživatele.

```
1  public static int[] FindUserSettingBorders(string user, string[]
       lines)
2  {
3     int[] result = { int.MaxValue, int.MaxValue };
4
5     for (int i = 0; i < lines.Length; i++)
6     {
7        if (lines[i].Contains(user))
8        {
9           // start of user's setting
10          result[0] = i;
11          break;
12       }
13    }
14    if (result[0] != int.MaxValue) {
15       for (int i = result[0] + 1; i < lines.Length; i++) {
16          if (lines[i].Contains("user_"))
17          {
18             // end of user's setting
19             result[1] = i;
20             break;
21          }
22       }
23    }
24    return result;
25 }
```

Zdrojový kód 4: Funkce pro nalezení hranic záznamu uživatele v nastavení Rspamd

4.4 Spam filter plugin

Pro výběr pravidel filtrace spamu byl vytvořen plugin, nabízející úpravu přímo v uživatelské schránce v rámci RoundCube. Spam filter plugin je napsán v jazyce PHP. Jak vytvořit plugin pro RoundCube jsem se nechal inspirovat veřejnými již hotovými pluginy. Nejvíce informací jsem čerpal z pluginu Thunderbird Labels. [9] Spam filter plugin je umístěný v adresáři */var/lib/roundcube/plugins/spam_filter*, kde se také nachází adresář *localization* obsahující jednotlivé jazykové mutace. Plné lokalizace jsou v českém, slovenském a anglickém jazyce. V dalších vybraných jazycích je pak překlad pro název pluginu a názvy pravidel spam filtru.

Akce, které se provádí poté, co uživatel klikne na uložení vybraných pravidel ve spam filter pluginu, jsou uvedeny ve zdrojovém kódu 5. Nejprve se vypočítá číslo filtru (není uvedeno v ukázce) a poté se zjistí uživatelské jméno právě přihlášeného uživatele. V poslední fázi se zavolá skript pro aktualizaci nastavení Rspamd a jako parametry se použijí uživatelské jméno a číslo filtru.

```
1  $rcmail = rcmail::get_instance();
2  $user   = $rcmail->user;
3  $identity = $user->get_identity();
4  $email = $identity['email'];
5
6  $username = substr($email, 0, strpos($email, '@'));
7
8  chdir('/etc/roundcube/plugins/spam_filter/sh');
9  $command = './update_user_setting.sh ' . $username . ' ' .
       $filter_number;
10 shell_exec($command);
```

Zdrojový kód 5: Volání skriptu s parametry po uložení pravidel spam filtru

4.5 Ruční učení spam filtru

V základu se spam filtr učí sám a z akcí uživatelů. Správce serveru má však navíc možnost ručně naučit globální nebo i uživatelský filtr hamu či spamu.

Na adrese https://posta.inf.upol.cz/rspamd/ je k dispozici webové rozhraní Rspamd, které nabízí možnost vložit email v plain textu a nechat jej otestovat. Nabízí také možnost využít vložený email pro naučení globálního filtru hamu nebo spamu. Jedná se o jednoduchou a pohodlnou možnost učení globálního filtru, ovšem pouze po jednom emailu.

Druhou možností je učení globálního nebo uživatelského filtru přímo na serveru skrze terminál. Velkou výhodou oproti webovému rozhraní je, že můžeme filtr učit více emailů současně. Emaily je možné zadat jako seznam parametrů, nebo je možné zvolit adresář, ve kterém se emaily nachází. Synopsis má tvar *rspamc [options] [command] [input-file]...*, kde do *options* je možné uvést konrétního uživatele, v případě učení globálního filtru se parametr vynechá. Za *command* se pak podle potřeby vybere jedna z variant *learn_ham* nebo *learn_spam*. Na místo *input-file* je pak možné uvést jeden či více emailů za sebou, nebo adresář s emaily.[10]

Závěr

Výsledná práce poskytuje plnohodnotné řešení emailové služby, včetně zajištění základní bezpečnosti. Nabízí také, dle mého názoru, velmi slušnou úroveň filtrace spamu a to jak z globálního tak uživatelského hlediska. Celkově práce splňuje má očekávání a požadavky, které jsem si na začátku stanovil. Největší komplikací bylo ukládat akce uživatelů do globální a uživatelské statistiky současně. Obtížné však bylo pouze přijít na řešení, samotná implementace je již triviální.

Jednoho nedostatku jsem si vědom, a to, že novému uživateli jsou při jeho vytvoření na straně serveru zapnuta globální a statická pravidla, ovšem uživatel po přihlášení do RoundCube tato pravidla nevidí zapnutá. Jakmile si však uživatel nějaká pravidla vybere a uloží, pak se již vše zobrazuje podle očekávání.

Možným rozšířením do budoucna by mohlo být přidání možnosti měnit váhu uživatelských symbolů pro ham a spam emaily. Uživatel by také mohl mít možnost si vyresetovat svůj spam filtr nebo jej naopak naučit všechny emaily, které již zná globální filtr.

Conclusions

The resulting work provides a full-fledged email service solution, including basic security. It also offers, in my opinion, a very decent level of spam filtering from a global and user perspective. Overall, the work meets my expectations and requirements that I set out at the beginning. The biggest complication was storing user actions in the global and user statistics at the same time. However, it was only difficult to come up with a solution, the implementation itself is trivial.

One shortcoming I am aware of is that a new user has global and static rules enabled on the server side when they are created, but the user does not see these rules enabled when they log into RoundCube. However, once the user selects and saves some rules, then everything is displayed as expected.

A possible future extension could be to add the ability to change the weight of user symbols for ham and spam emails. The user could also be able to reset their spam filter or, on the contrary, teach it all emails that the global filter already knows.

A Obsah přiloženého datového média

Popis souborů a adresářů umístěných na přiloženém datovém médiu.

`bin/`
: Zde jsou umístěny všechny upravené konfiguračné soubory, spustitelné soubory zdrojových kódů a skripty. Nachází se zde také adresář obsahující kompletní spam filter plugin.

`doc/`
: Adresář obsahuje text práce ve formátu PDF a soubory použité pro jeho vygenerování.

`src/`
: Obsahem adresáře jsou všechny zdrojové kódy vytvořené pro fungování spam filter pluginu.

`readme.txt`
: Soubor obsahující instrukce pro zprovoznění serveru, včetně seznamu všech softwarů, které je potřeba na server nainstalovat. Jsou zde uvedeny také všechny přístupové údaje spojené s implementací v rámci této práce.

Literatura

[1] Email Statistics Report, 2015-2019. The Radicati Group, Inc. [online]. The Radicati Group, 2021 [cit. 2022-05-09]. Dostupné z: https://www.radicati.com/wp/wp-content/uploads/2015/02/Email-Statistics-Report-2015-2019-Executive-Summary.pdf

[2] HAAS, Christoph. Workaround.org: ISPmail guide for Debian 11 "Bullseye" [online]. 2021 [cit. 2022-05-09]. Dostupné z: https://workaround.org/ispmail/bullseye/

[3] GUENTHER, Philip; Tim SHOWALTER. Sieve: An Email Filtering Language. RFC Editor [online]. RFC Editor, 2008 [cit. 2022-05-09]. Dostupné z: https://www.rfc-editor.org/info/rfc5228

[4] STAKHOV, Vsevolod. Rspamd Statistics. Rspamd spam filtering system [online]. 2008 [cit. 2022-05-09]. Dostupné z: https://rspamd.com/doc/configuration/statistic.html

[5] DEVROYE, Luc; László GYÖRFI; Gábor LUGOSI. A probabilistic theory of pattern recognition. New York: Springer, c1996. ISBN 0-3879-4618-7.

[6] SIRAINEN, Timo; Solar DESIGNER; Andrey PANIN; et al. Service configuration. Dovecot manual [online]. 2002 [cit. 2022-05-09]. Dostupné z: https://doc.dovecot.org/configuration_manual/service_configuration/

[7] VSEVOLOD, STAKHOV. Milter headers module. Rspamd spam filtering system [online]. 2008 [cit. 2022-05-09]. Dostupné z: https://rspamd.com/doc/modules/milter_headers.html

[8] VSEVOLOD, STAKHOV. Rspamd user settings. Rspamd spam filtering system [online]. 2008 [cit. 2022-05-09]. Dostupné z: https://rspamd.com/doc/configuration/settings.html

[9] KEFEDER, Michael. Thunderbird Labels. Thunderbird Labels [online]. [2011] [cit. 2022-05-09]. Dostupné z: https://github.com/mike-kfed/roundcube-thunderbird_labels

[10] Rspamc - rspamd command line client. Debian [online]. 2007 [cit. 2022-05-09]. Dostupné z: https://manpages.debian.org/testing/rspamd/rspamc.1.en.html