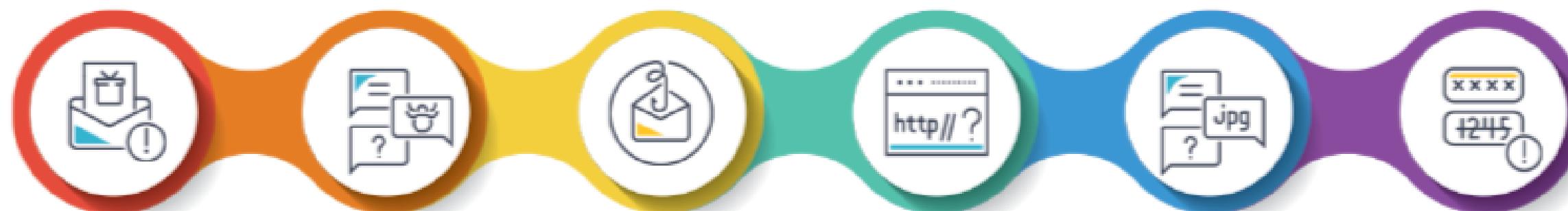


2ème année du cycle d'ingénieur en cybersécurité et confiance numérique
Institut National de Postes et Télécommunications

*Mise en place d'un service de messagerie sécurisé et développement
d'un outil automatisé d'analyse des attaques par courrier électronique
destiné aux analystes du SOC (Security Operations Center).*



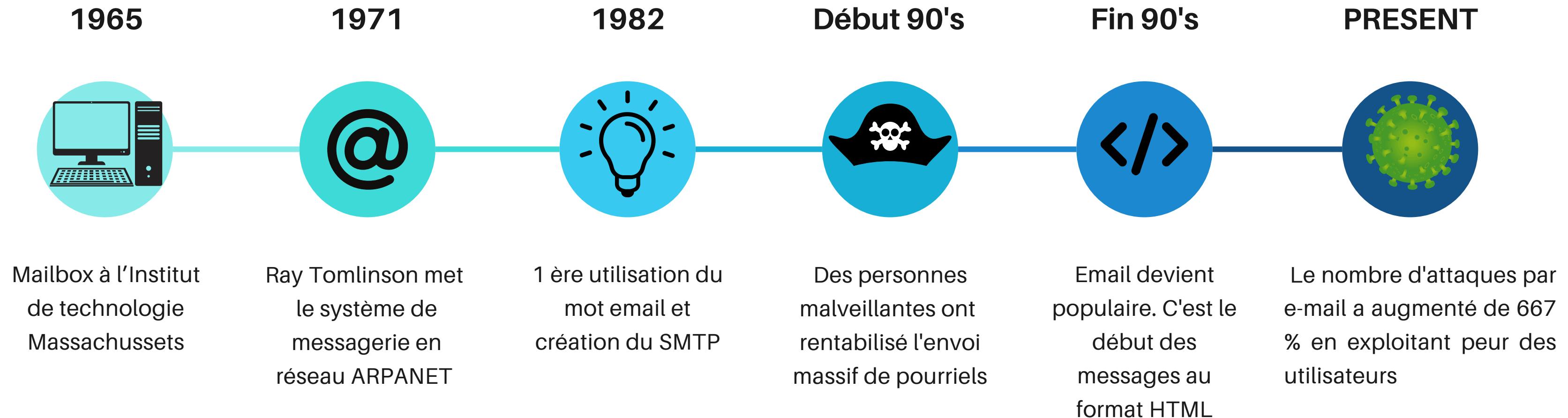
RÉALISÉ PAR : BOUGALZIME IMANE
HARROUCH E IBTISSAM
HINA JE NEZHA
MOUSSISET HAMZA
MAAKOUL ACHRAF



Introduction



Historique de l'email



Comprendre les techniques d'attaque pour mieux s'en prémunir



Écoute passive du trafic réseau : Cette technique consiste à surveiller discrètement le trafic réseau pour y détecter des informations sensibles ou des signes d'activités suspectes. C'est un outil puissant pour les pirates, mais qui peut également être utilisé par les professionnels de la sécurité pour identifier des menaces..



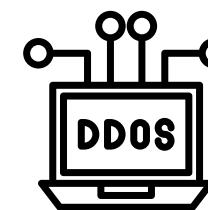
Attaques par force brute sur les mots de passe: Il s'agit d'essayer un grand nombre de combinaisons de mots de passe pour accéder à un système. C'est une méthode laborieuse mais qui peut être très efficace si les mots de passe sont faibles. La mise en place de politiques de mots de passe robustes est essentielle pour s'en prémunir.



Ingénierie sociale (phishing, hameçonnage) : Ce type d'attaque vise à tromper les utilisateurs pour obtenir des informations sensibles ou les pousser à exécuter des actions malveillantes. La sensibilisation des utilisateurs est cruciale pour les protéger contre ces techniques de manipulation.



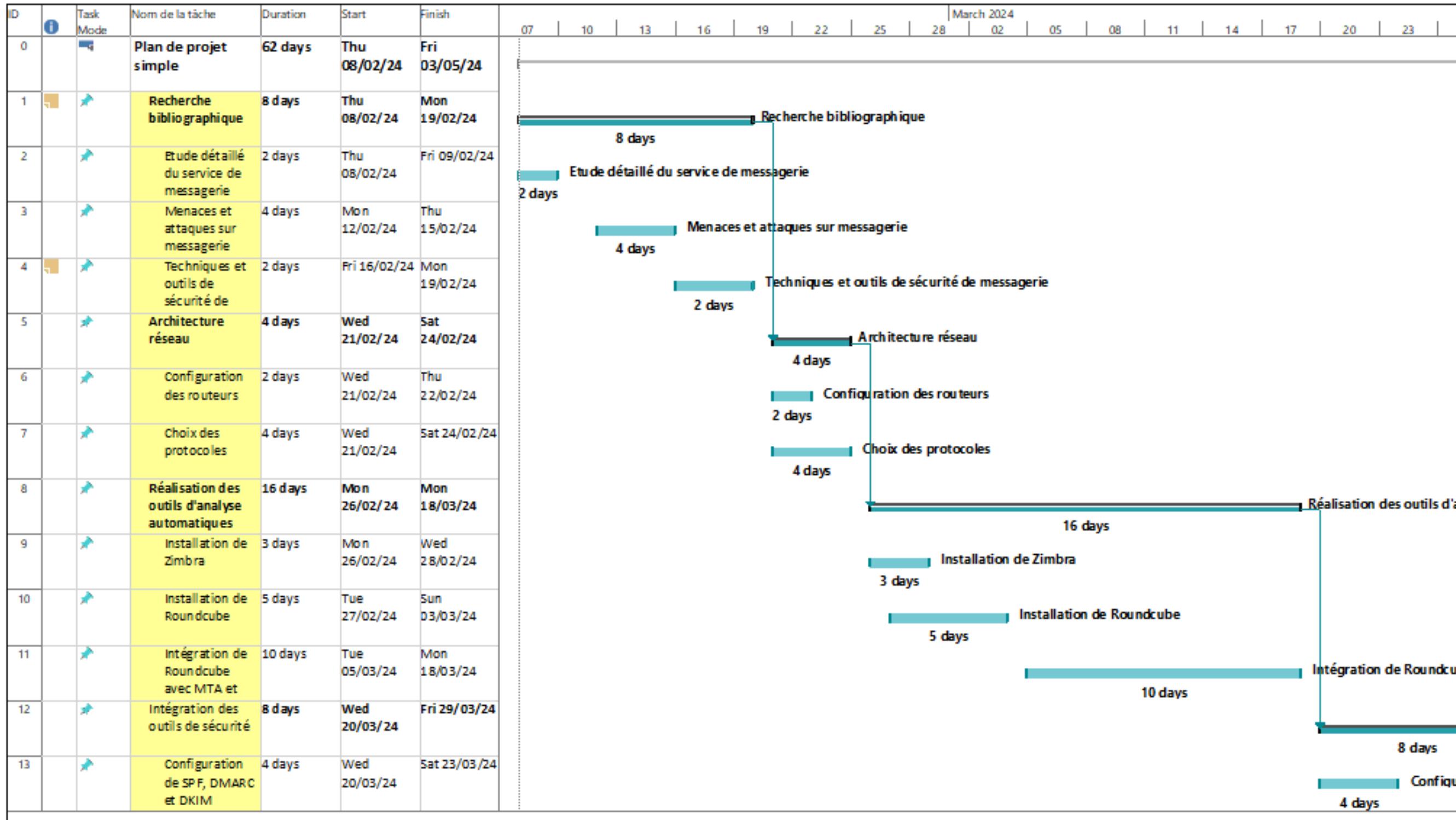
Exploitation de vulnérabilités logicielles : Les pirates exploitent des failles de sécurité dans les logiciels pour s'y introduire. Une veille constante et des mises à jour régulières sont nécessaires pour limiter les risques.



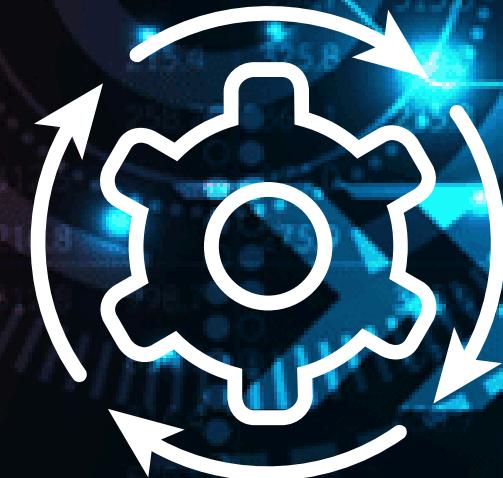
Saturation des ressources (DDoS) : Il s'agit d'inonder un système avec une quantité massive de trafic pour le rendre inaccessible. Des mesures de protection spécifiques doivent être mises en place pour y faire face.



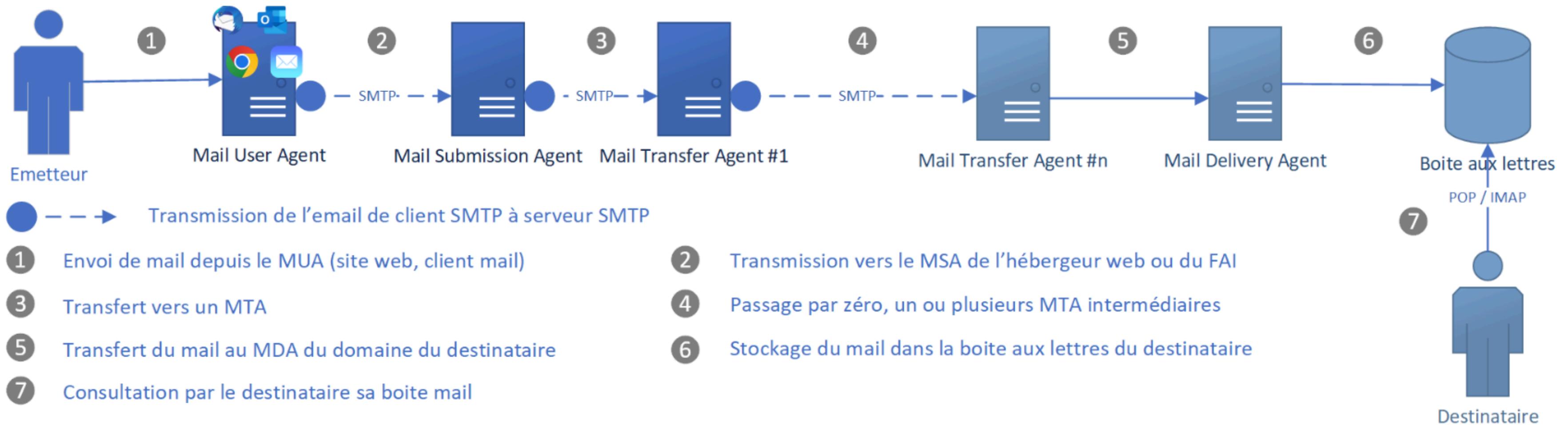
Gestion du projet



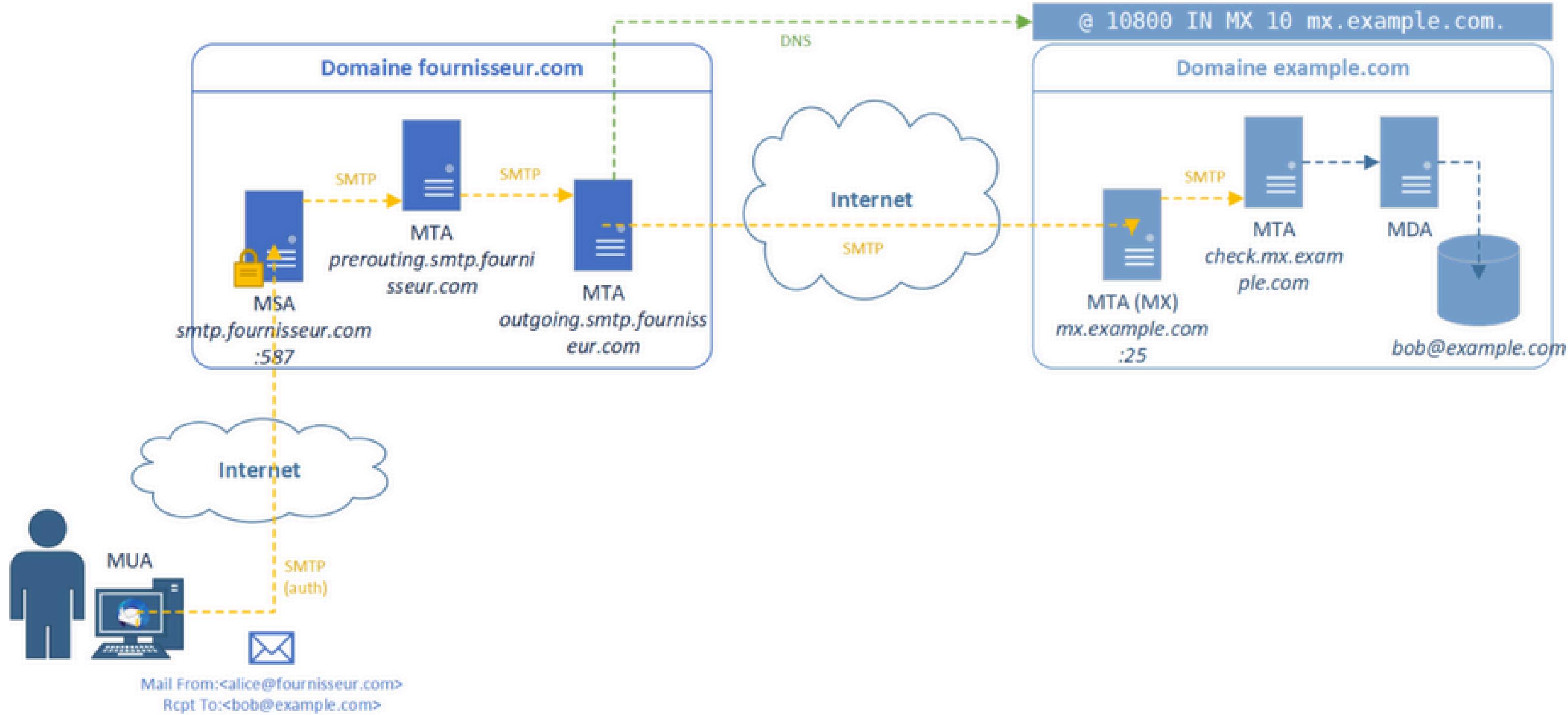
Fonctionnement de service de messagerie



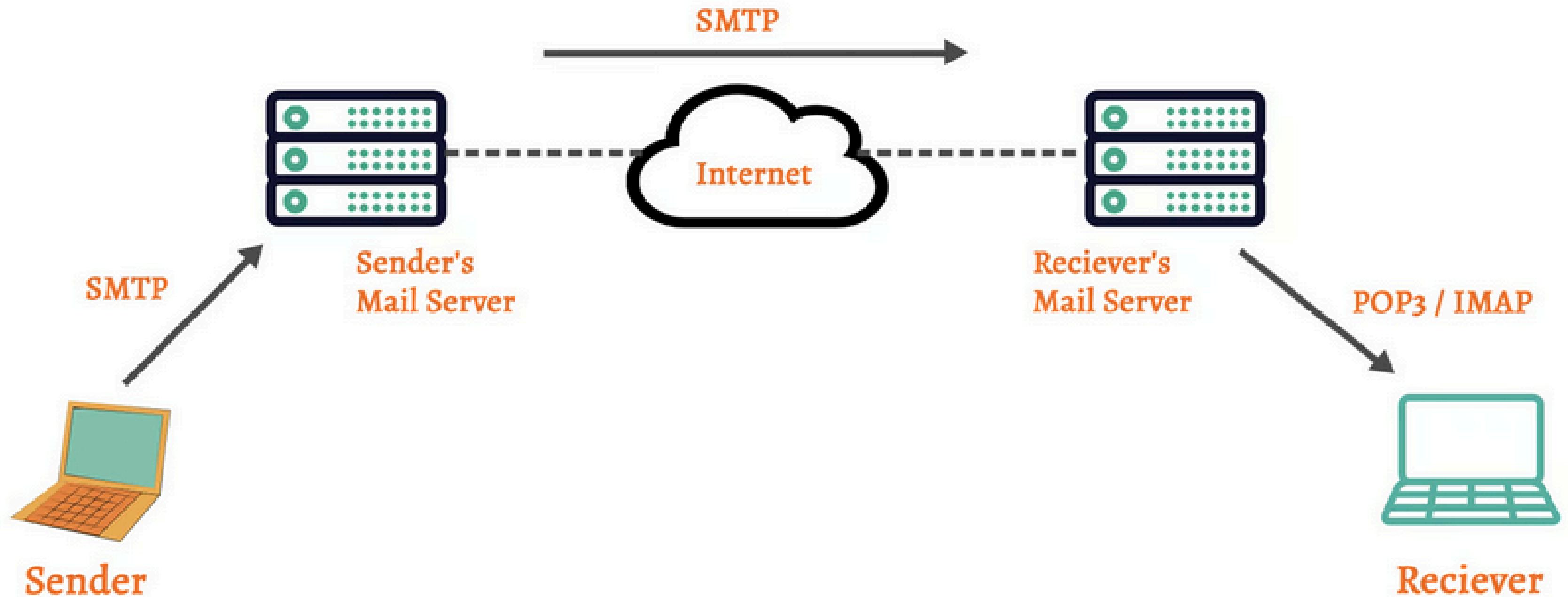
Les acteurs de la messagerie électronique



Exemple d'acheminement de courriel électronique



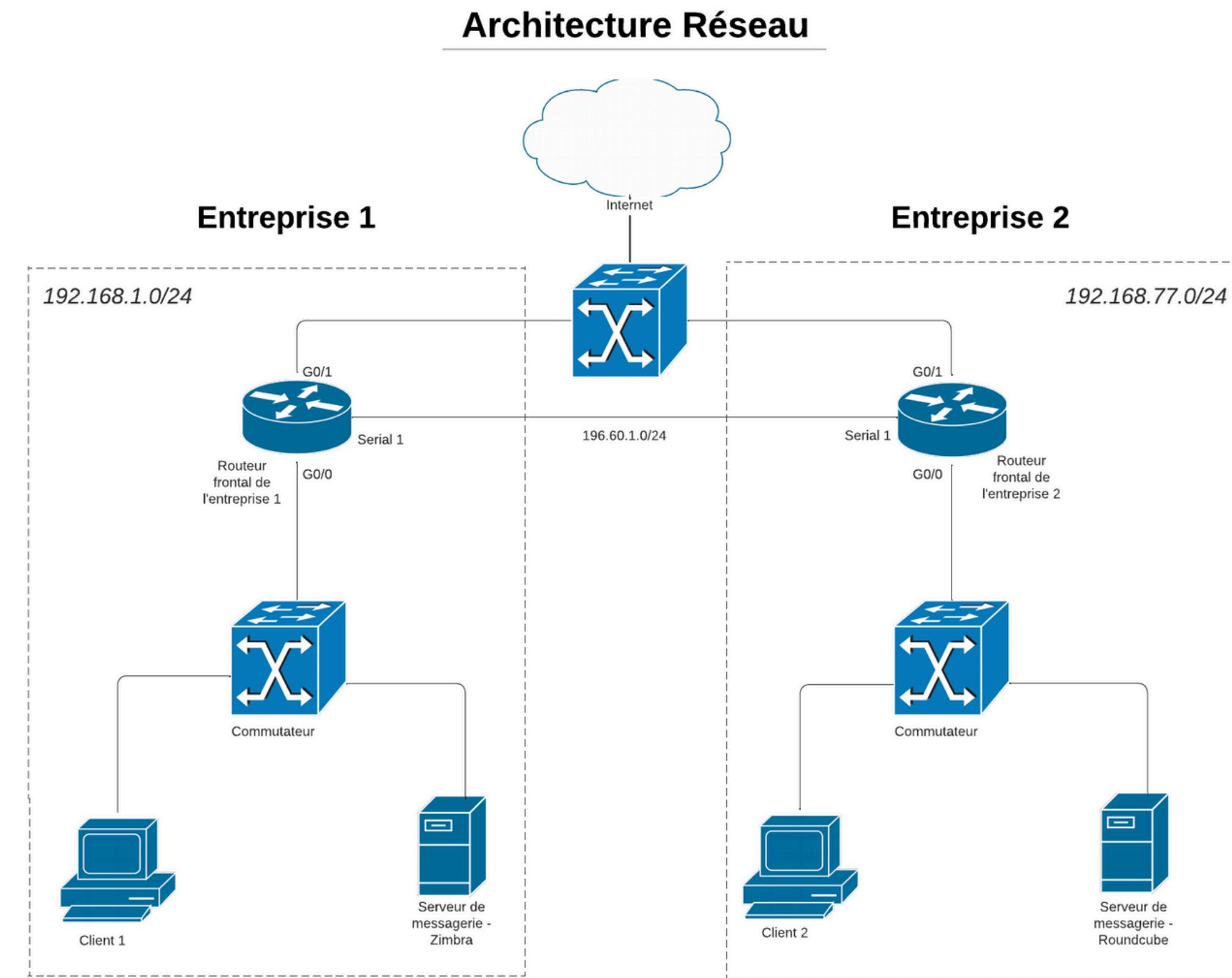
les protocoles utilisés





Architecture Réseau pour l'interconnexion des services de messagerie entre deux entreprises

• Architecture globale des entreprises



• Entreprise 1

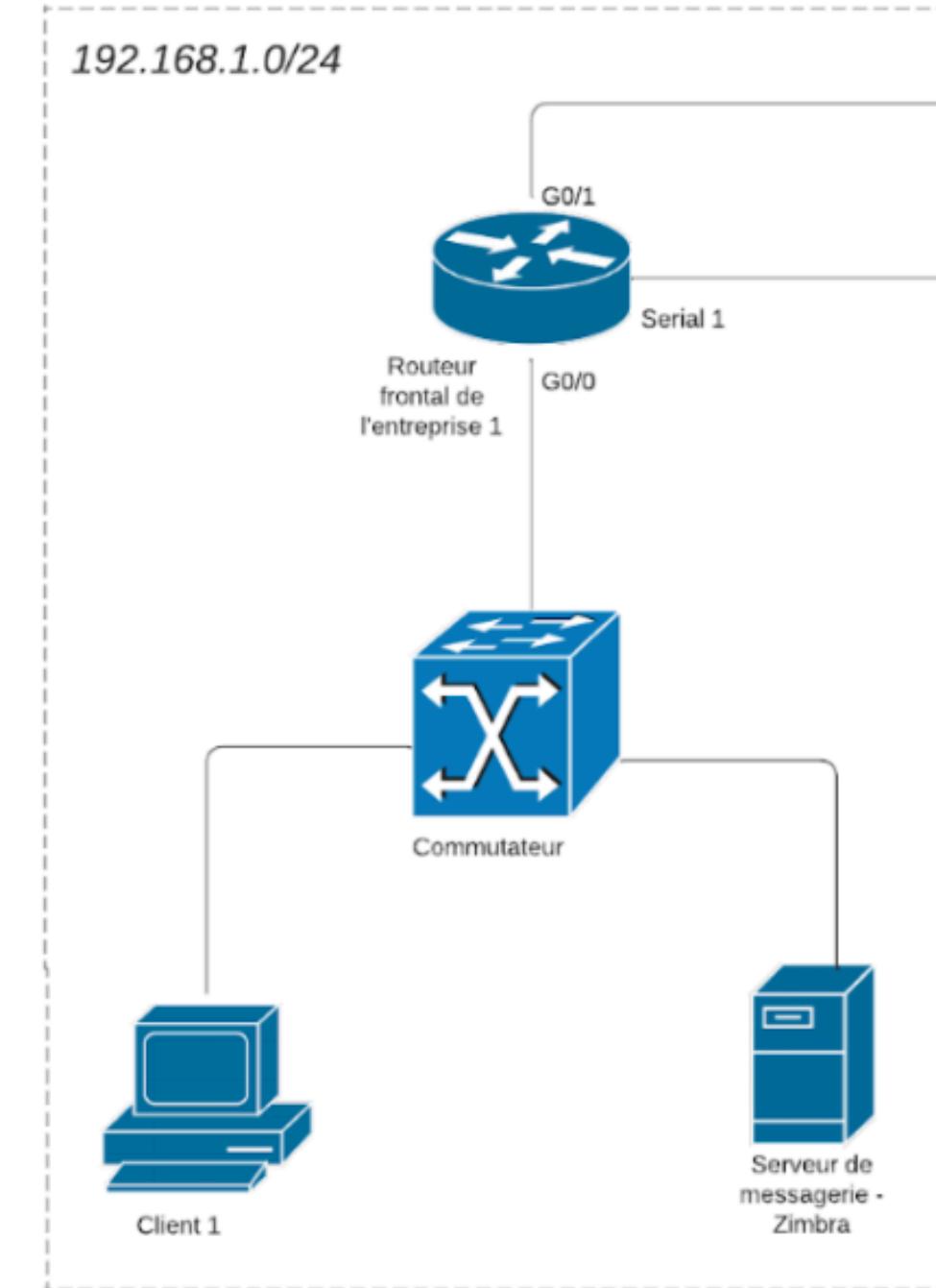
Plage IP: 192.168.1.0/24

Domain: entreprise1.ma

Composants:

- Routeur frontal de l'Entreprise 1
- Commutateur
- Client 1
- Serveur de messagerie (Zimbra) + DNS

Le routeur frontal se connecte à l'internet via l'interface G0/1 et au réseau interne via G0/0. Le commutateur relie le routeur, le client et le serveur de messagerie.



• Entreprise 2

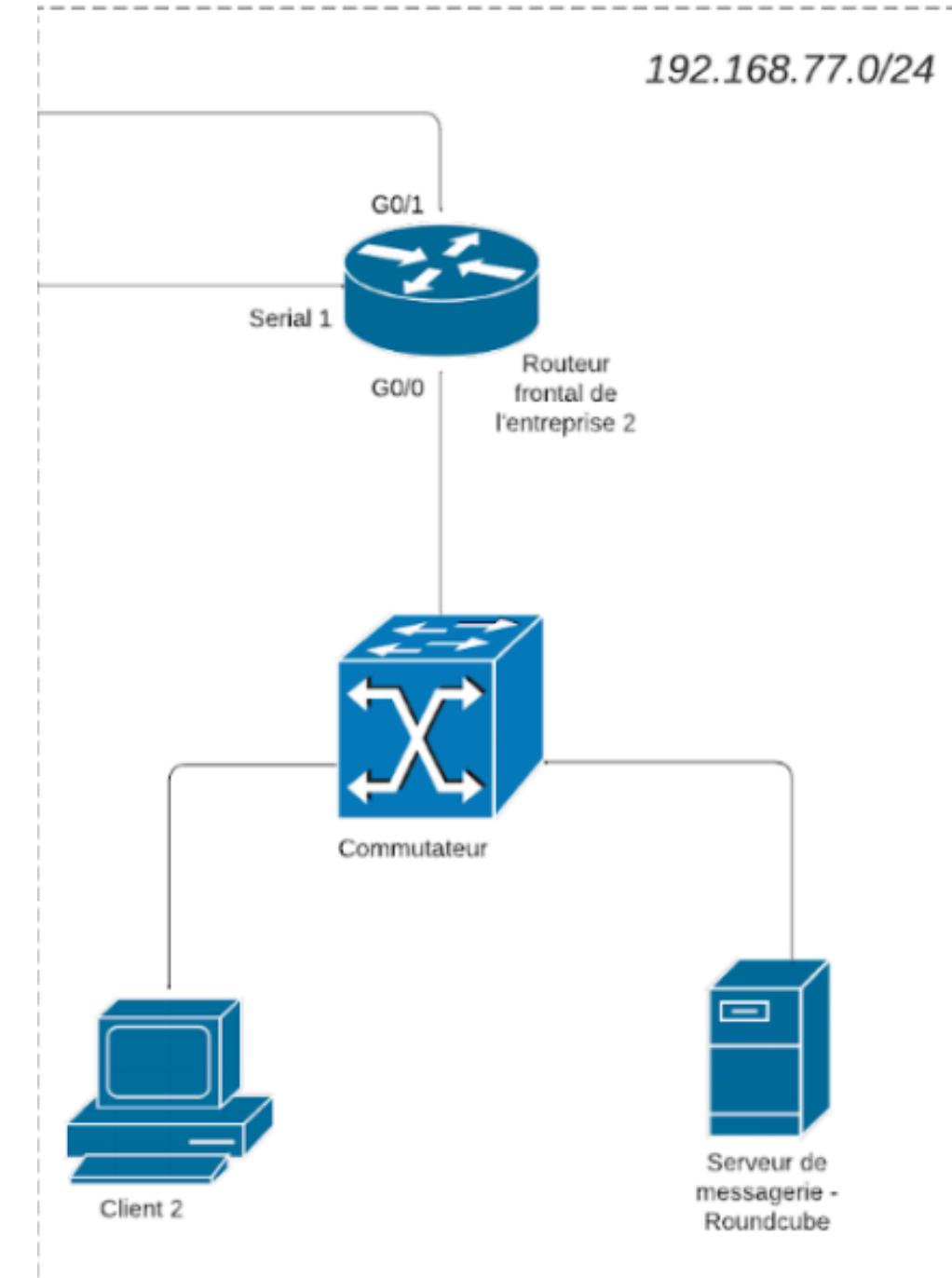
Plage IP: 192.168.77.0/24

Domain: entreprise1.ma

Composants:

- Routeur frontal de l'Entreprise 2
- Commutateur
- Client 2
- Serveur de messagerie (Roundcube) + DNS

Le routeur frontal se connecte à l'internet via l'interface G0/1 et au réseau interne via G0/0. Le commutateur relie le routeur, le client et le serveur de messagerie.

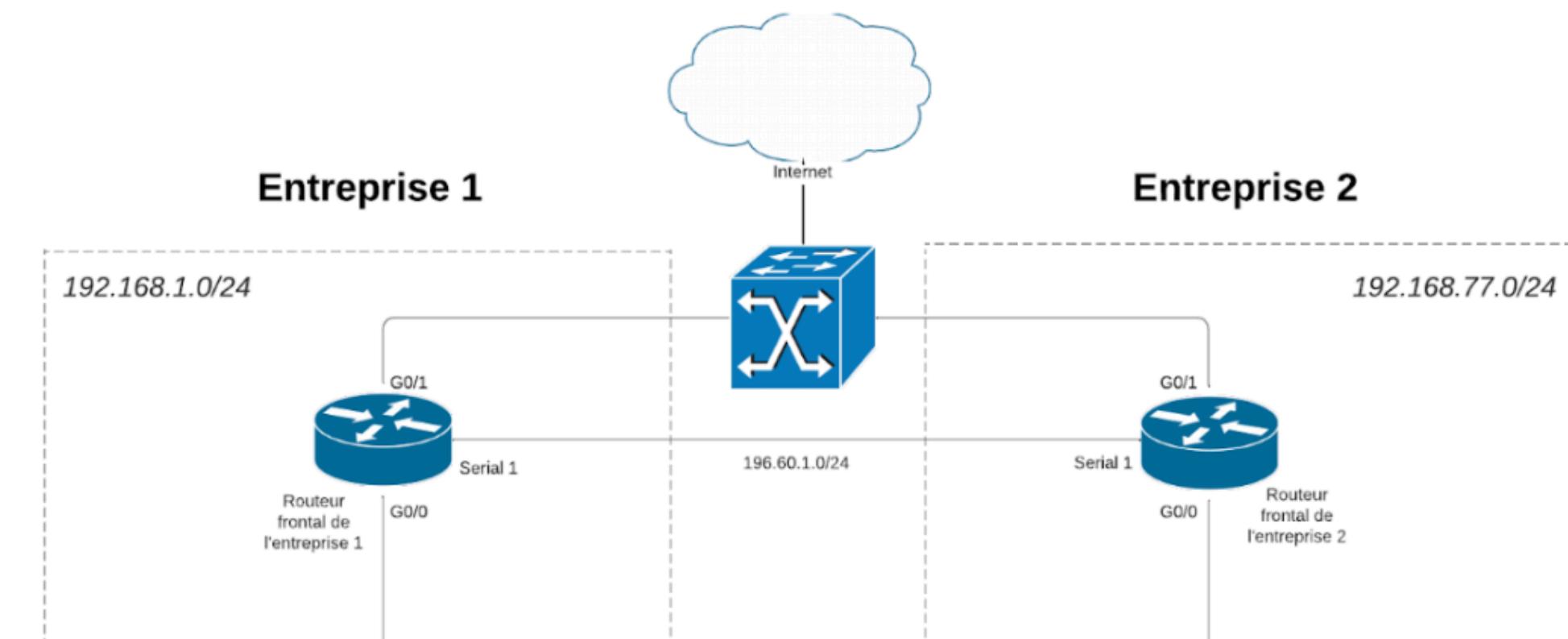


- **Interconnexion des 2 sites et accès Internet**

Réseau Intermédiaire: 196.60.1.0/24

Configuration du NAT statique pour masquer les plans d'adressage des 2 entreprise

Configuration du PAT pour l'accès Internet

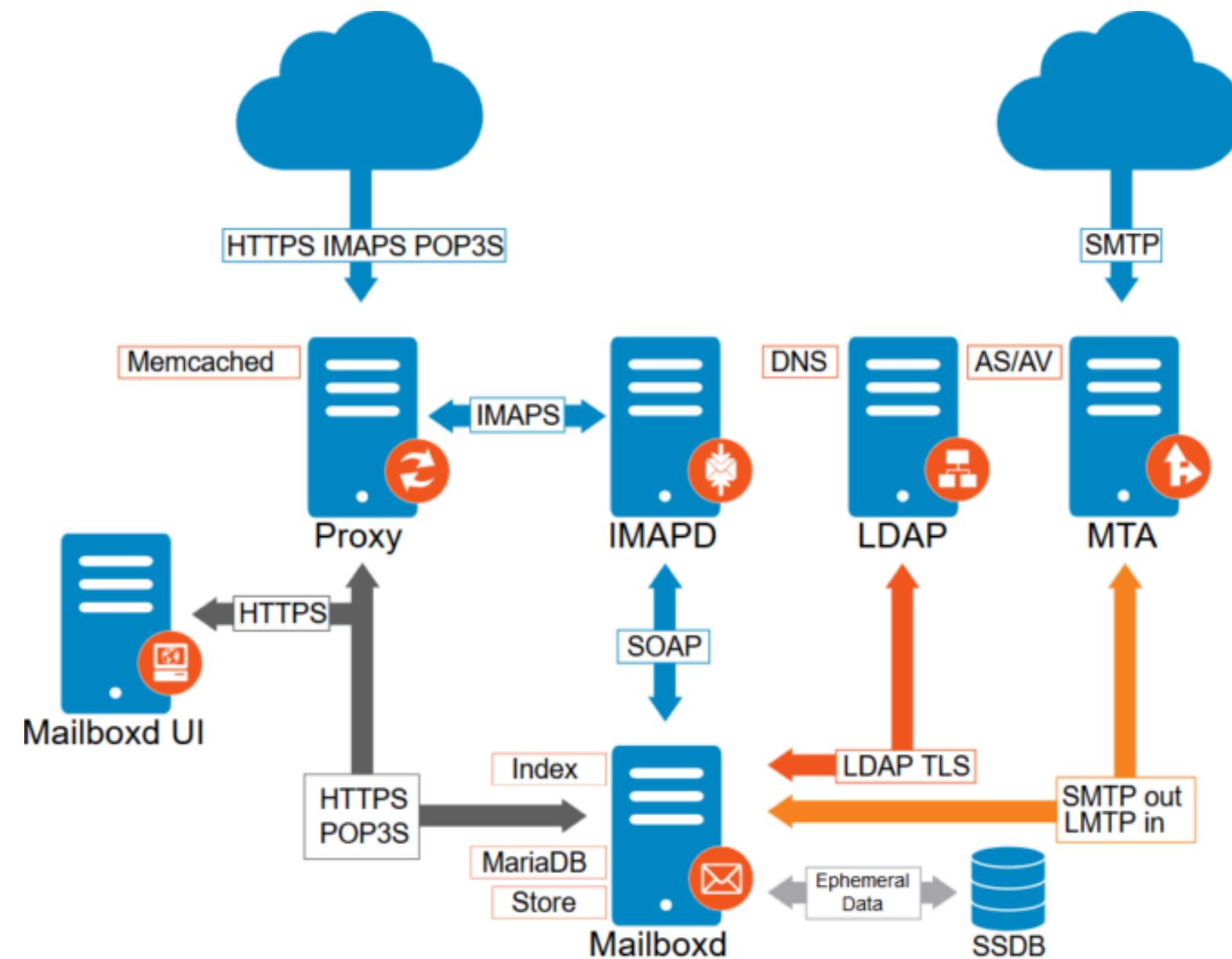


Mise en place d'un service de messagerie sécurisé



Serveur de messagerie de l'entreprise 1

Zimbra est également connu sous le nom de Zimbra Collaboration Suite (ZCS) car il se compose de nombreux composants tels que MTA (Postfix), Base de données (MariaDB), LDAP et MailboxdUI, etc. Voici l'architecture de Zimbra.

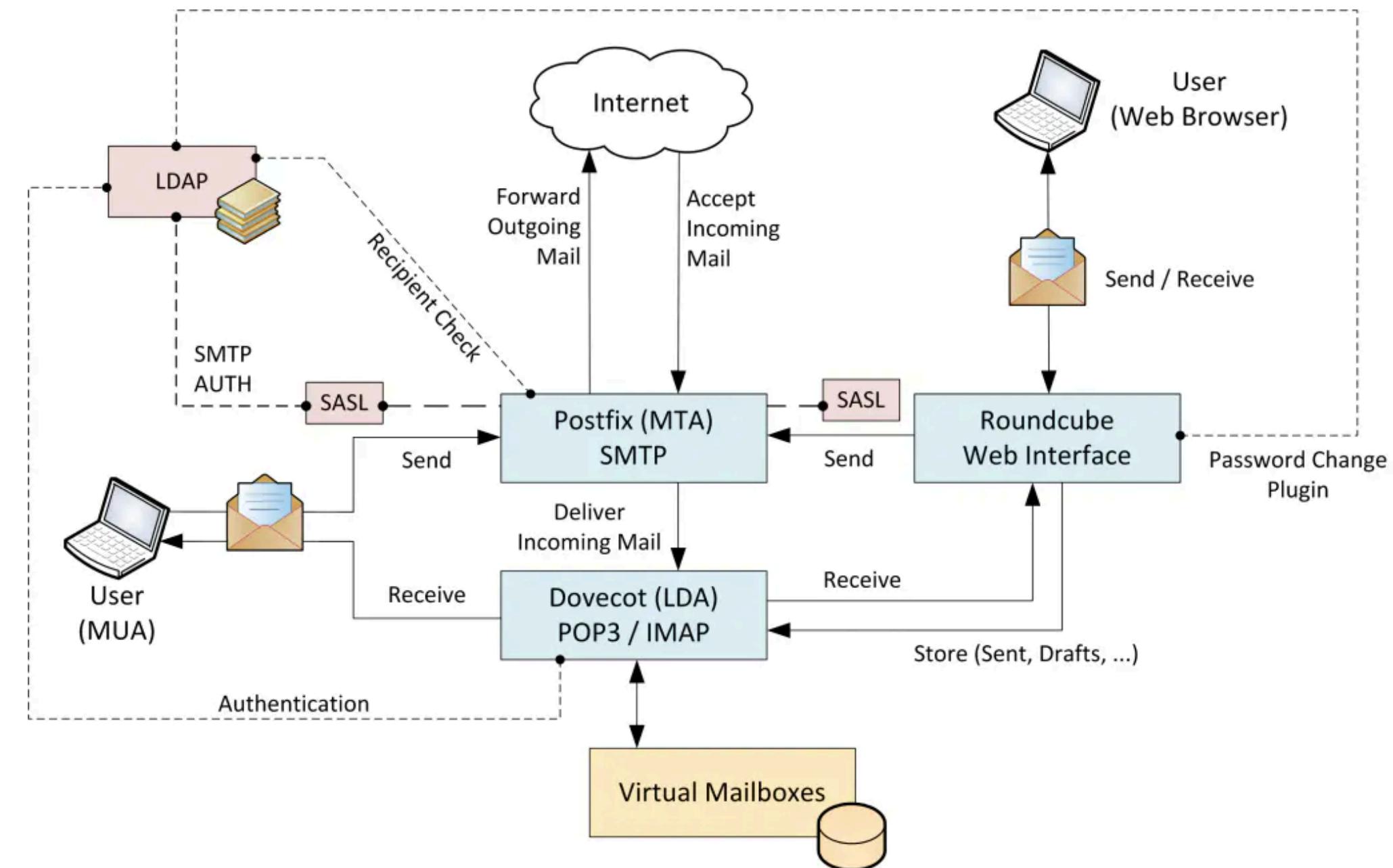


Composants de Zimbra

- Proxy: Gère les connexions HTTPS, IMAPS, et POP3S. Utilise Memcached pour améliorer les performances.
- IMAPD: Gère les connexions IMAP. Communique avec les serveurs DNS et AS/AV (Antivirus/Antispam).
- LDAP: Utilisé pour la gestion des utilisateurs et des authentifications. Communique avec les autres composants via LDAP TLS. MTA (Mail Transfer Agent):
- Mailboxd: Stocke les emails et gère l'interface utilisateur de la boîte aux lettres (Mailboxd UI). Communique avec la base de données MariaDB pour le stockage et utilise des données éphémères.

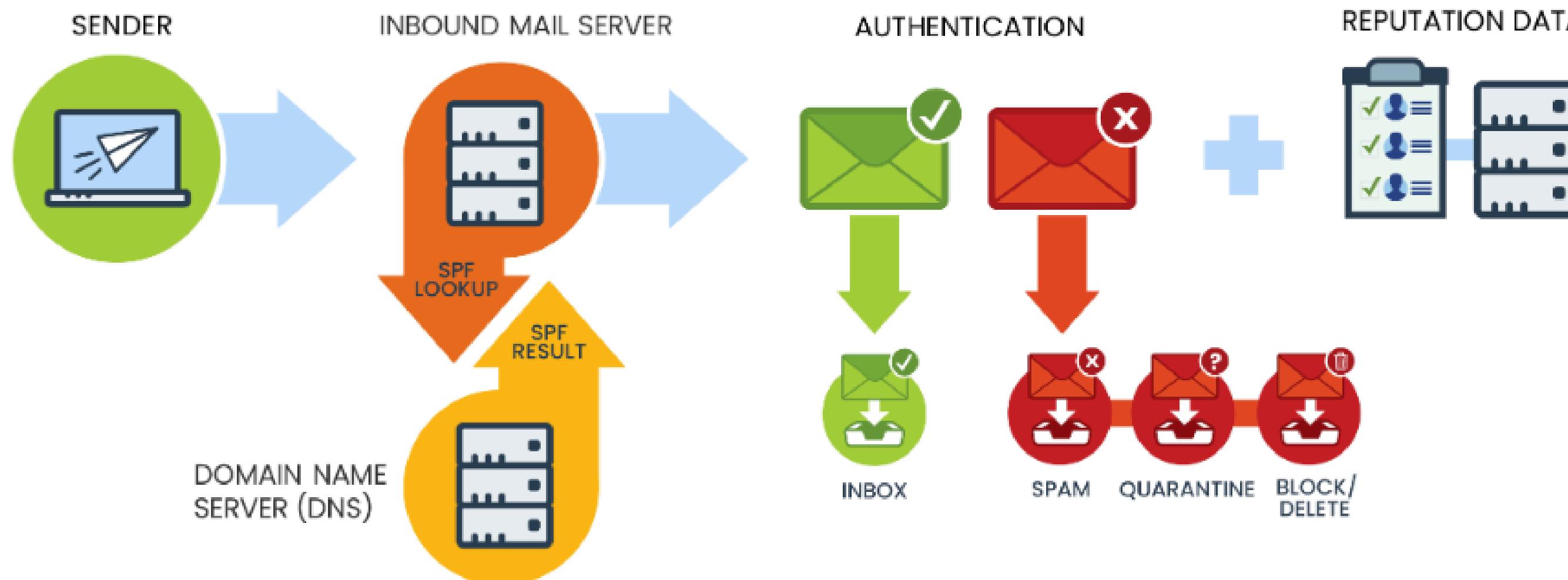
Serveur de messagerie de l'entreprise 2

Roundcube, proposé par le géant français OVH, se distingue comme un Webmail open-source performant et apprécié par de nombreux utilisateurs. Il fonctionne sous le protocole IMAP et la technologie Ajax.



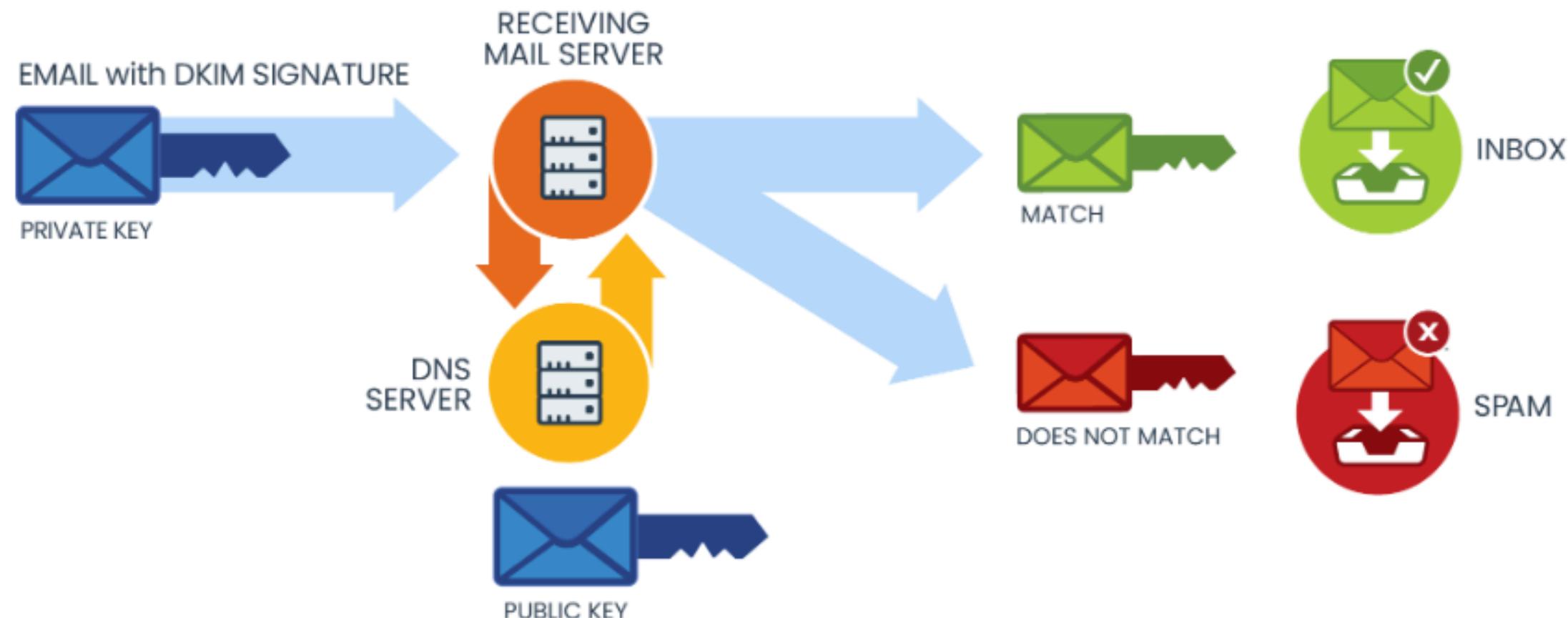
SPF

Sender Policy Framework est une norme utilisée pour authentifier les serveurs de messagerie émetteurs d'e-mails. Il vérifie si le serveur de messagerie qui envoie un e-mail est autorisé.



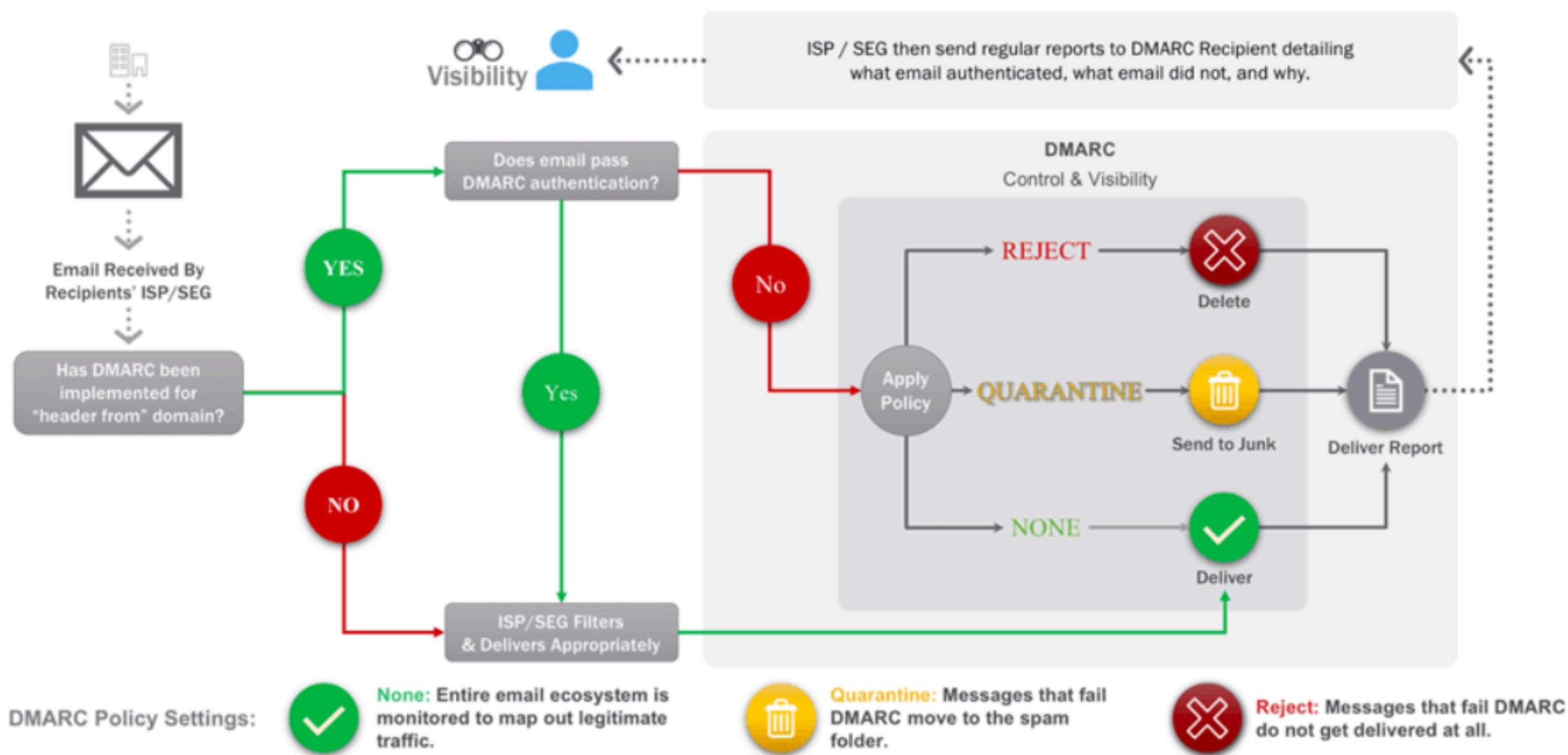
DKIM

DomainKeys Identified Mail, est une méthode d'authentification du courrier électronique basée sur des signatures cryptographiques qui vise à lutter contre l'usurpation d'adresse e-mail et la modification des messages.



DMARC

Domain-based Message Authentication, Reporting & Conformance est un mécanisme de vérification des emails qui permet aux propriétaires de domaines de définir des règles sur la manière dont leurs emails doivent être traités par les serveurs de messagerie récepteurs.



Implémentation SPF, DKIM, DMARC

```
GNU nano 2.3.1                                         File: /var/named/entreprise1.db

$TTL 86400
@ IN SOA    ns1.entreprise1.ma. root.entreprise1.ma. (
                2024052501 ; Serial
                3600      ; Refresh
                1800      ; Retry
                604800    ; Expire
                86400     ; Minimum TTL
)
IN NS      ns1.entreprise1.ma.
IN MX 10   mail.entreprise1.ma.

ns1  IN A      192.168.83.176
mail IN A      192.168.83.176
www  IN A      192.168.83.176
@    IN A      192.168.83.176 ; This sets the A record for example.com

@ TXT "v=spf1 a mx ~all"

_dmarc TXT "v=DMARC1; p=none; fo=1; rua=mailto:authority@entreprise1.ma; ruf=mailto:authority@entreprise1.ma" 8251

C590E408-1ACA-11EF-9374-4B665D00840B._domainkey IN      TXT      ( "v=DKIM1; k=rsa; "
"p=MIIIBIjANBgkqhkiG9w0BAQEFAAOCAQ8AMIIIBCgKCAQEAvyuGQeMiKcavWB5oCb6mwRzY6raUYcIfAfEuzJf/Zd7r9LvCrzih+W9Mi
"VrqipMsji336r5TGBg3ufXdSXetTf8VSqGmlSqJcBfJK2gdg7TD5YJDNtTwErkfeUqy96JLhecz5zTuaHmpTLn3+Wh0MxNbKCuT63Ei
```

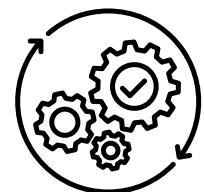
Test de fonctionnement:

```
Return-Path: <hamza@entreprisel.ma>
Received: from mail.entreprisel.ma (LHL0 mail.entreprisel.ma)
(192.168.34.146) by mail.entreprisel.ma with LMTP; Sat, 25 May 2024
20:58:32 +0100 (WEST)
Received: from localhost (localhost [127.0.0.1])
by mail.entreprisel.ma (Postfix) with ESMTP id C949014159B3
for <nezha@entreprisel.ma>; Sat, 25 May 2024 20:58:32 +0100 (+01)
X-Spam-Flag: NO
X-Spam-Score: -1.209
X-Spam-Level:
X-Spam-Status: No, score=-1.209 required=6.6 tests=[ALL_TRUSTED=-1,
DKIM_SIGNED=0.1, DKIM_VALID=-0.1, DKIM_VALID_AU=-0.1, DKIM_VALID_EF=-0.1,
HTML_MESSAGE=0.001, T_SCC_BODY_TEXT_LINE=-0.01]
autolearn=ham autolearn_force=no
Authentication-Results: mail.entreprisel.ma (amavis); dkim=pass (2048-bit key)
header.d=entreprisel.ma
Received: from mail.entreprisel.ma ([127.0.0.1])
by localhost (mail.entreprisel.ma [127.0.0.1]) (amavis, port 10032)
with ESMTP id Fax21NOUGKho for <nezha@entreprisel.ma>;
Sat, 25 May 2024 20:58:30 +0100 (+01)
Received: from localhost (localhost [127.0.0.1])
by mail.entreprisel.ma (Postfix) with ESMTP id A5CD41115E38
for <nezha@entreprisel.ma>; Sat, 25 May 2024 20:58:30 +0100 (+01)
DKIM-Filter: OpenDKIM Filter v2.10.3 mail.entreprisel.ma A5CD41115E38
DKIM-Signature: v=1; a=rsa-sha256; c=relaxed/relaxed; d=entreprisel.ma;
s=C590E408-1ACA-11EF-9374-4B665D00840B; t=1716667110;
bh=++lAGULas/QJxn/96Lnw+NN0njH89vFkv+cNnCtIDtA=;
h=Date:From:To:Message-ID:MIME-Version;
b=JKE7045TfU687w9siRae09R4A7jFUSqrWEAdagdu9fEamhqh27W9kSHpXmd6xpPs0
uXq+b/buiL5IJHFFKrRcdMVM07RcM65PkNUFMci0iQAUfaSSGPi5EGsfELXKnMfBn5
nZNRF4tnEwHfVZ8uptQrE6HJhINboA7a9FGYr+stBo78tV+R9IQI7BnMwMYe7FF9i4
20pXL2xqNUpVUdw6SuzQl0MJnYrbsTfktUrVI00lbITmg5rPQn7g0Qz6l0D0HBr2y2
0+9mFVTdA1C65IuBhBZljo4czVwy8BPXZkoHTsDVKwI9aNpRsDhNzhlXuCiXRES93B
8g+RcW3wN3BNg==
```

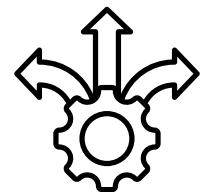
Optimisation de la sécurité du service de messagerie



Pourquoi choisir SPAMASSASSIN ?



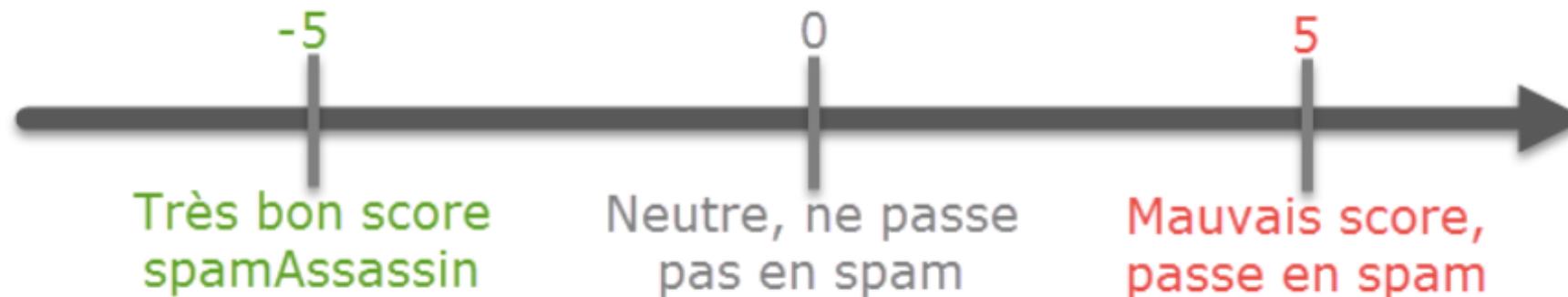
Efficacité



Flexibilité



Utilisation des plugins



Implémentation de SPAMASSASSIN

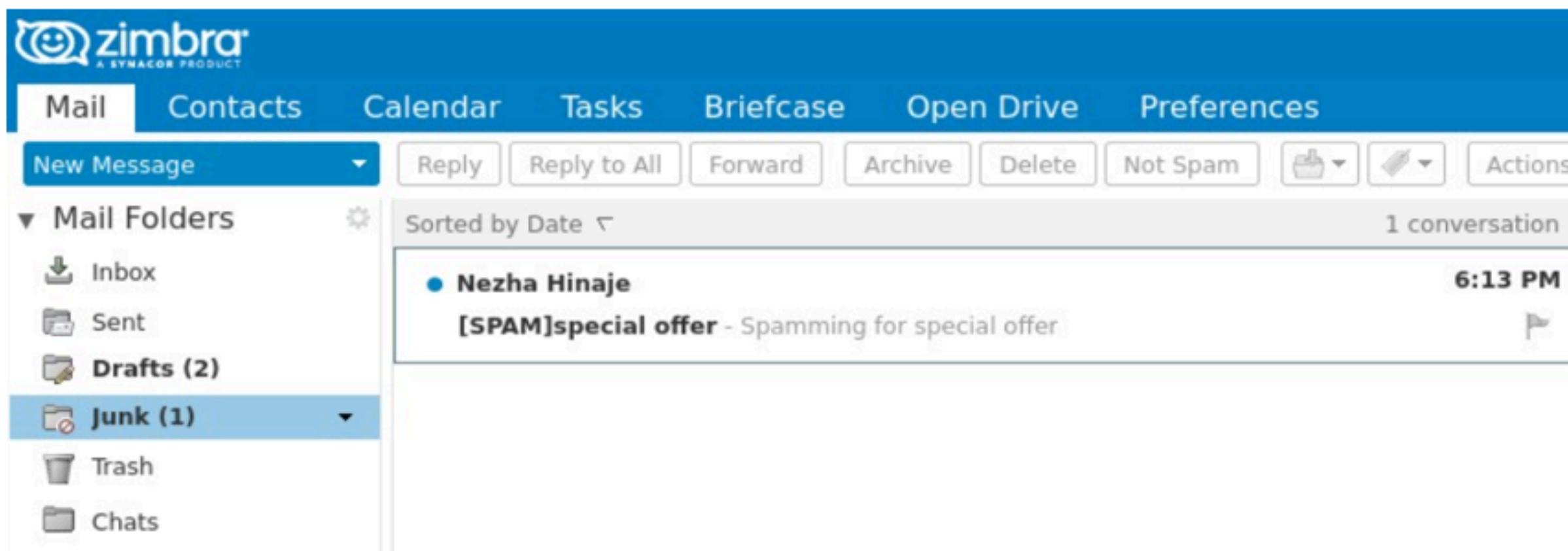
```
##rules
# Règles pour les en-têtes
header LOCAL SUBJECT RULE Subject =~ /spam|spam|viagra|lottery|win|free|urgent|discount|offer|click here|exclusive|limited time|winner|prize|gift|special promotion|act now|deal|money|
score LOCAL SUBJECT RULE 2.0

# Règles pour les corps des e-mail
body LOCAL BODY RULE /discount|free|offer|limited time offer|click here|unsubscribe|urgent|promotion|winner|prize|http://\[^s]+|https://\[^s]+|example\.com|free-stuff\.com|cheap-
#body LOCAL BODY RULE /discount|
score LOCAL BODY RULE 5.0

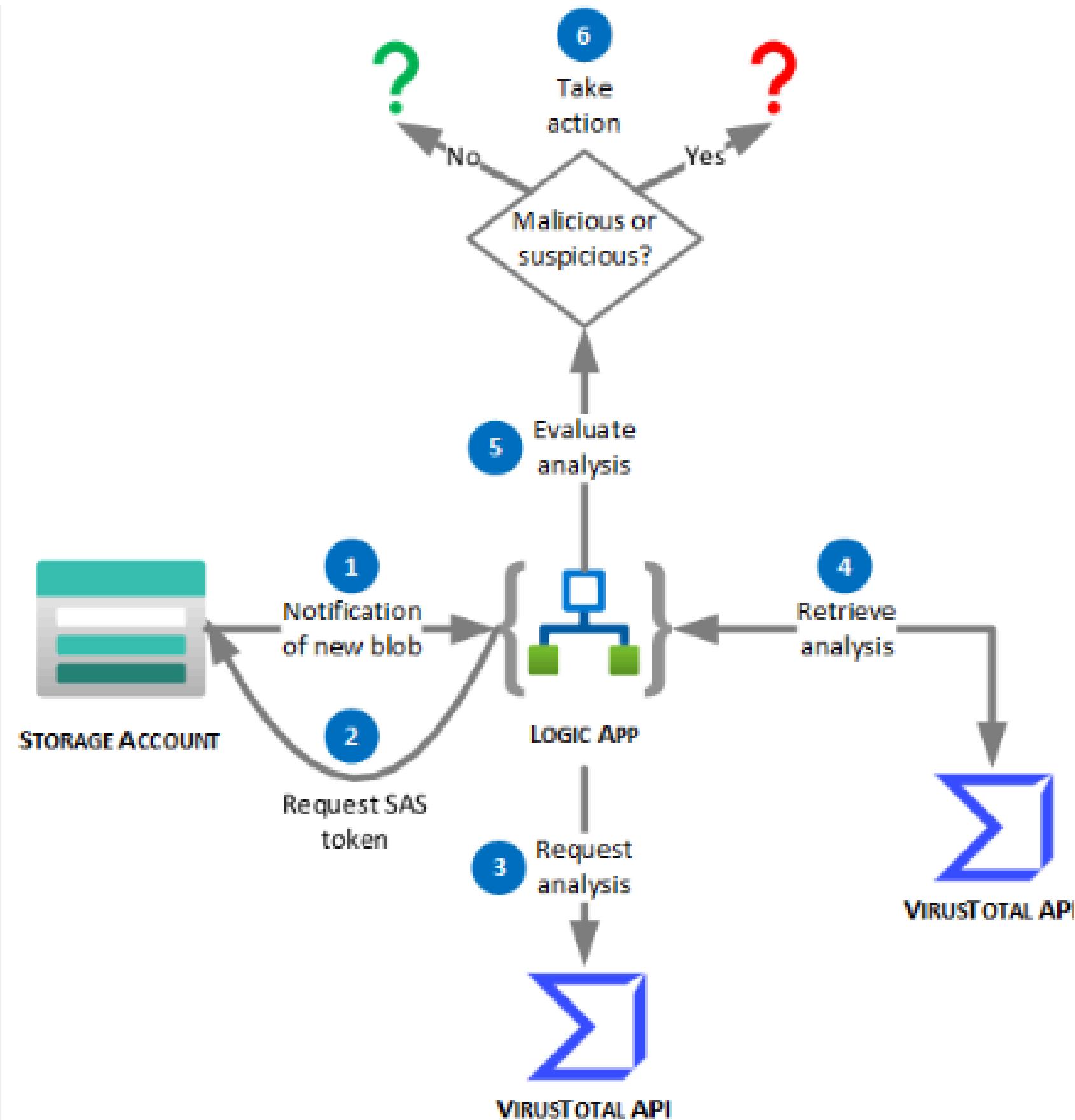
#rule pour from
# Règles pour l'adresse de l'expéditeur (From)
header LOCAL FROM RULE From =~ /spam@entreprise2.ma/
score LOCAL FROM RULE 5.0

#required score
required_score 1.0
```

Test de fonctionnement:



API VirusTotal



Avantages :



Détection multi-moteurs : Utilise plusieurs moteurs antivirus pour une analyse approfondie, augmentant ainsi la probabilité de détection de menaces.



Automatisation : Permet l'intégration dans des systèmes existants pour automatiser la soumission et l'analyse de fichiers et d'URLs, économisant du temps et des ressources.



Rapports détaillés : Fournit des rapports complets incluant les résultats de divers moteurs de sécurité, facilitant une évaluation rapide et précise des menaces.



Détection rapide des nouvelles menaces : Grâce à des mises à jour fréquentes des moteurs antivirus, l'API aide à détecter rapidement les nouvelles menaces émergentes.



Communauté et partage de données : Bénéficie d'une large communauté d'utilisateurs et de contributeurs, améliorant la base de données des menaces et permettant des réponses plus rapides et précises.

Implémentation dans postfix :

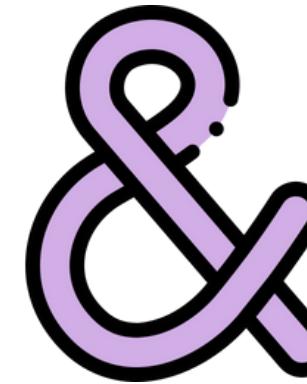
```
from virustotal_python import Virustotal
import pandas as pd
import re
import os
import sys
import requests
import json

API_KEY = "e0e89390e92609657d53bb53aef916c9ec4b213e35597b41b675e50bc58c556a"
vt = Virustotal(API_KEY)

# Fonction pour extraire l'URL du texte
def extract_url(text):
    if isinstance(text, str):
        url_pattern = re.compile(r'https?://\S+')
        match = url_pattern.search(text)
        return match.group(0) if match else None
    return None

# Fonction pour enrichir l'URL en utilisant VirusTotal
def enrich_url(url):
    if url:
        try:
            response = vt.request("url/report", params={"resource": url})
            # Process the response to add metadata to the URL
        except Exception as e:
            print(f"Error during enrichment: {e}")

# Fonction pour extraire les URLs des messages
def extract_urls_from_email(email):
    urls = []
    for part in email.walk():
        if part.get_content_type() == "text/plain":
            urls.append(extract_url(part.get_payload()))
    return urls
```



```
#!/bin/bash
# vtCheckFilter.sh

# Créez un dossier temporaire pour stocker les emails
TMPDIR=$(mktemp -d)
cat > $TMPDIR/email.eml

# Vérifiez si l'email est marqué comme spam
if grep -q "X-Spam-Flag: YES" $TMPDIR/email.eml; then
    # Exécutez votre script Python avec l'email comme argument
    python3 ~/part3/AI/AI/vtCheck.py $TMPDIR/email.eml
fi

attachments=$(grep -oP 'filename="\K[^"]+' "$TMPDIR/email.eml")

# Scan each attachment
MALWARE_DETECTED=0
for attachment in $TMPDIR/*; do
    if [[ -f "$attachment" ]]; then
        python3 /path/to/vtCheck.py "$attachment"
        if [[ $? -eq 1 ]]; then
            MALWARE_DETECTED=1
            break
        fi
    fi
done
```

```
# VirusTotal filter
content_filter =virustotal_filter:127.0.0.1:10025
```

```
#virustotal
virustotal_filter unix - n n - - pipe
flags=Rq user=vtCheck argv=/usr/local/bin/vtCheckFilter.sh ${sender} ${recipient}
```

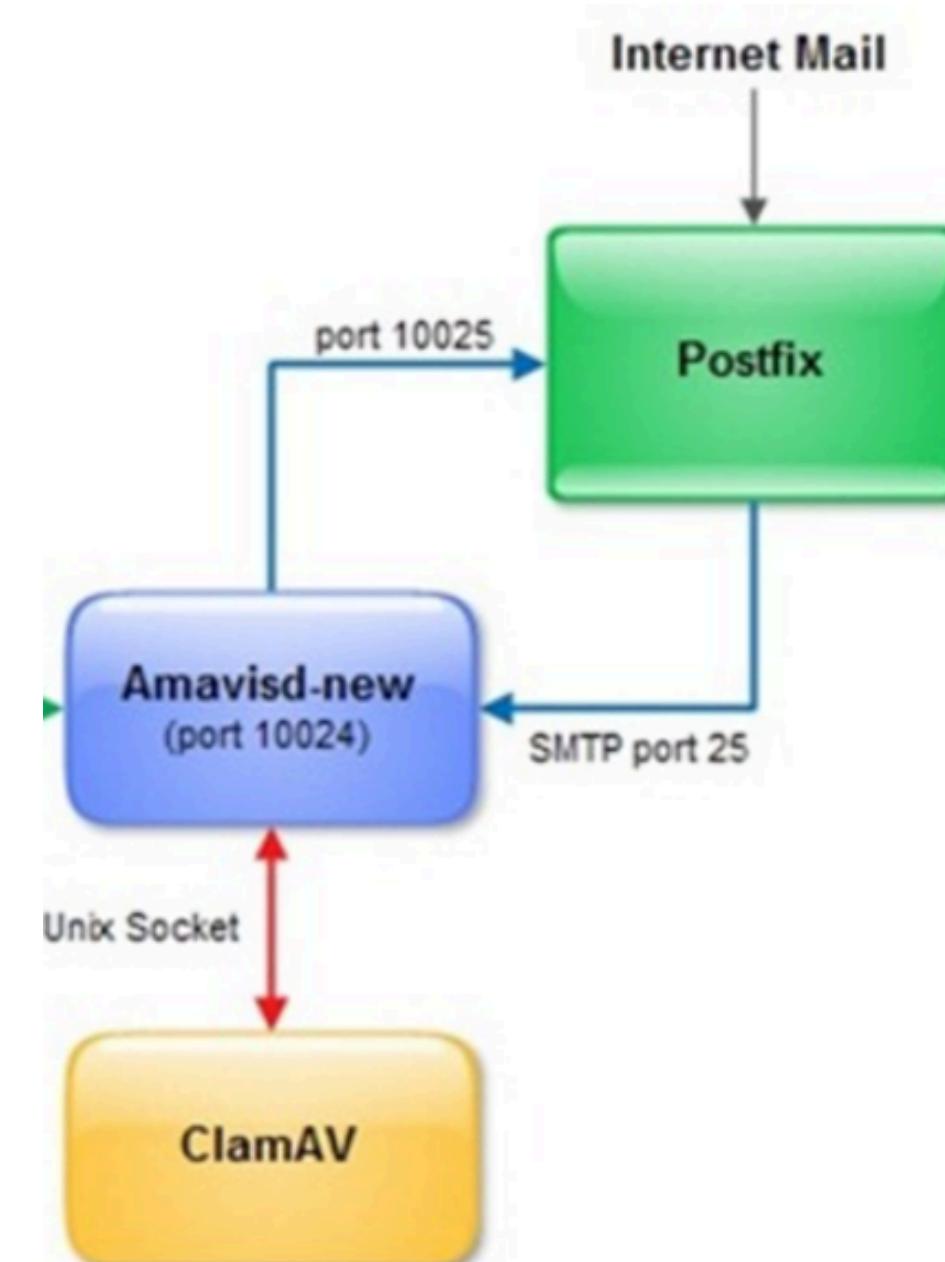


Test de fonctionnement:

The screenshot shows a web-based email client interface. The left sidebar includes icons forCompose, Mail, Contacts, and Settings, along with a Dark mode switch and About/Logout links. The main area has a header with user info (imane@entreprise2.ma), search, and navigation buttons (Select, Threads, Options, Refresh, Reply, Reply all, Forward, Delete, Mark, More). The left pane displays the 'Inbox' with 47 messages, including several spam reports from 'MAILER-DAEMON@entreprise2.ma' and 'iccn2@entreprise2.ma'. The right pane shows the selected message's content: a warning about potential spam, a content preview in French, content analysis details (PTS rules), and the original message body. The message body discusses unusual activity on the user's bank account and provides a verification link.

Clamav et Amavis

Amavis (A Mail Virus Scanner) est une interface haute performance entre un agent de transfert de messages (MTA) tel que Postfix et des filtres de contenu.



Test de fonctionnement:

Chaine de caractères pour le test:

XJS*C4JDBQADN1.NSBN3*2IDNEN*GTUBE-STANDARD-ANTI-UBE-TEST-EMAIL*C.34X

Screenshot of a web-based email client interface showing a spam analysis report.

Header: entreprise2.ma/?_task=mail&_mbox=INBOX

Toolbar: Select, Threads, Options, Refresh, Reply, Reply all, Forward, Delete, Mark, More

Search Bar: Search...

Message List:

- iccn2@entreprise2.ma (Today 10:07) - [*****SPAM****] test clamav
- iccn2@entreprise2.ma (Mon 23:24) - Urgent: Important Document Attached
- iccn2@entreprise2.ma (Mon 23:21) - Urgent: Important Document Attached
- iccn2@entreprise2.ma (Mon 22:46) - spam
- iccn2@entreprise2.ma (Mon 22:11) - hi
- iccn2@entreprise2.ma (Mon 15:31) - test no-spam
- iccn2@entreprise2.ma (Mon 14:17) - [*****SPAM****] XJS*C4JDBQADN1.NSBN3*2IDNEN*GTUBE-ST...
- iccn2@entreprise2.ma (Mon 13:51) - test2 apres opendkim
- iccn2@entreprise2.ma (Mon 13:30) - test apres opendkim conf
- iccn2@entreprise2.ma (Mon 11:59) - test avant OpenDkim
- iccn@entreprise2.ma (Mon 11:54) - Test message from Roundcube
- imanebougalzim9@gmail.com (Mon 00:06) - (no subject)

Content Preview: XJS*C4JDBQADN1.NSBN3*2IDNEN*GTUBE-STANDARD-ANTI-UBE-TEST-EMAIL*C.34X

Content Analysis Details: (1000.0 points, 1.0 required)

pts rule name	description
-1.0 ALL_TRUSTED	Passed through trusted hosts only via SMTP
1000 GTUBE	BODY: Generic Test for Unsolicited Bulk Email
0.0 URIBL_BLOCKED	ADMINISTRATOR NOTICE: The query to URIBL was blocked. See http://wiki.apache.org/spamassassin/DnsBlocklists#dnsbl-block for more information.
0.8 DKIM_ADSP_NXDOMAIN	[URIs: entreprise2.ma] No valid author signature and domain not in DNS
0.1 DKIM_SIGNED	Message has a DKIM or DK signature, not necessarily valid
0.1 DKIM_INVALID	DKIM or DK signature exists, but is not valid
0.0 TVD_SPACE_RATIO	No description available.
-0.0 T_SCC_BODY_TEXT_LINE	No description available.

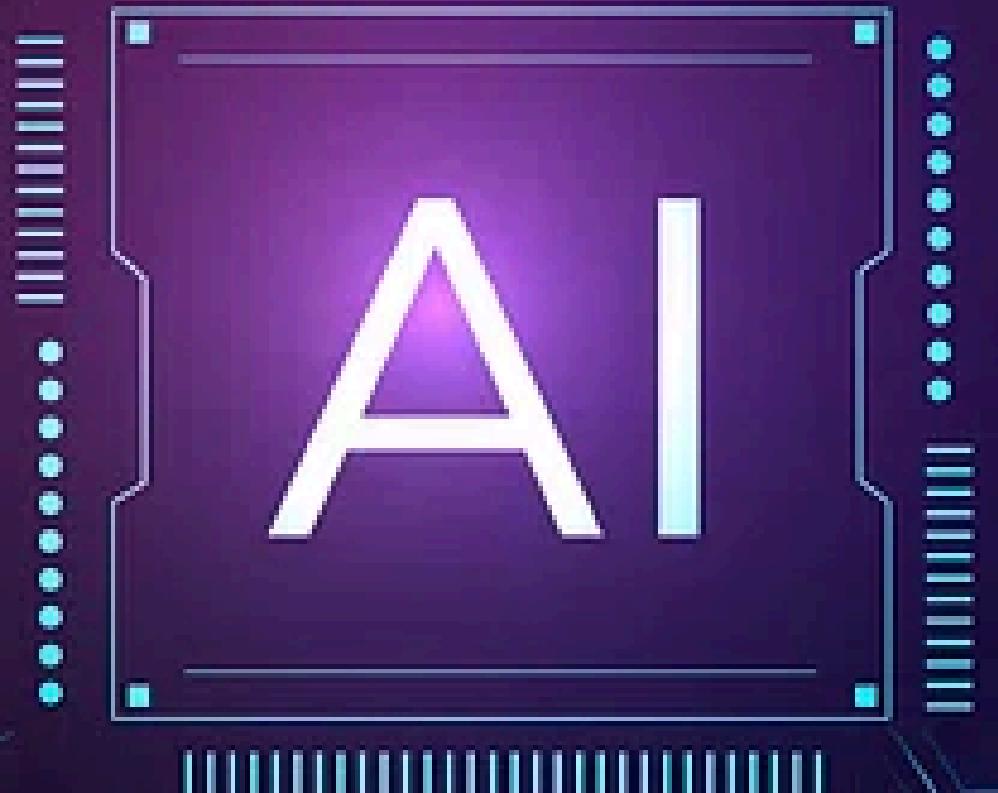
Message Headers:

```

Subject: test clamav
From: iccn2@entreprise2.ma
To: Imane
Date: Today 10:07

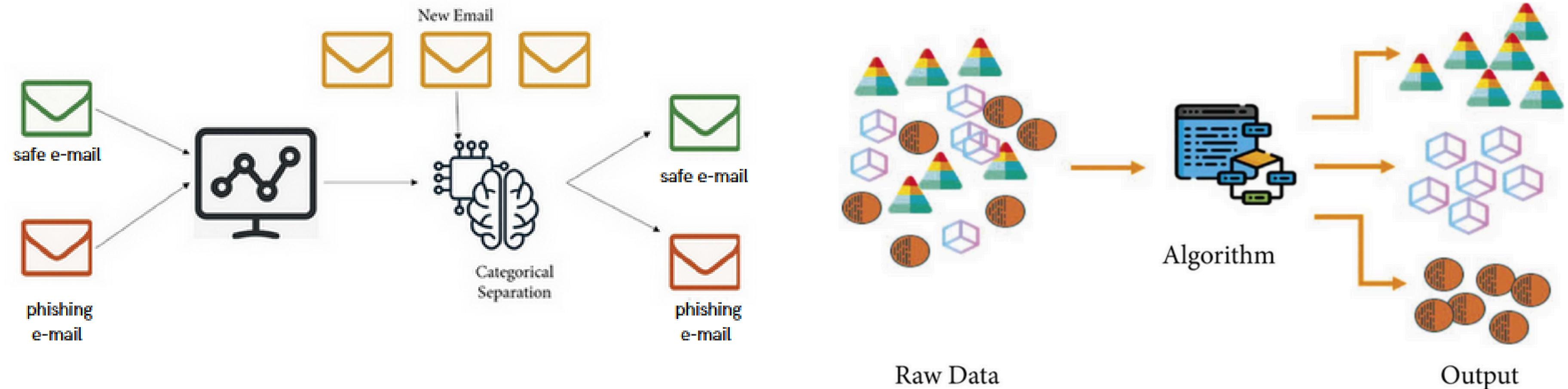
```

Footer: XJS*C4JDBQADN1.NSBN3*2IDNEN*GTUBE-STANDARD-ANTI-UBE-TEST-EMAIL*C.34X



**Implémentation de
l'apprentissage automatique
pour la détection du
phishing/spam**

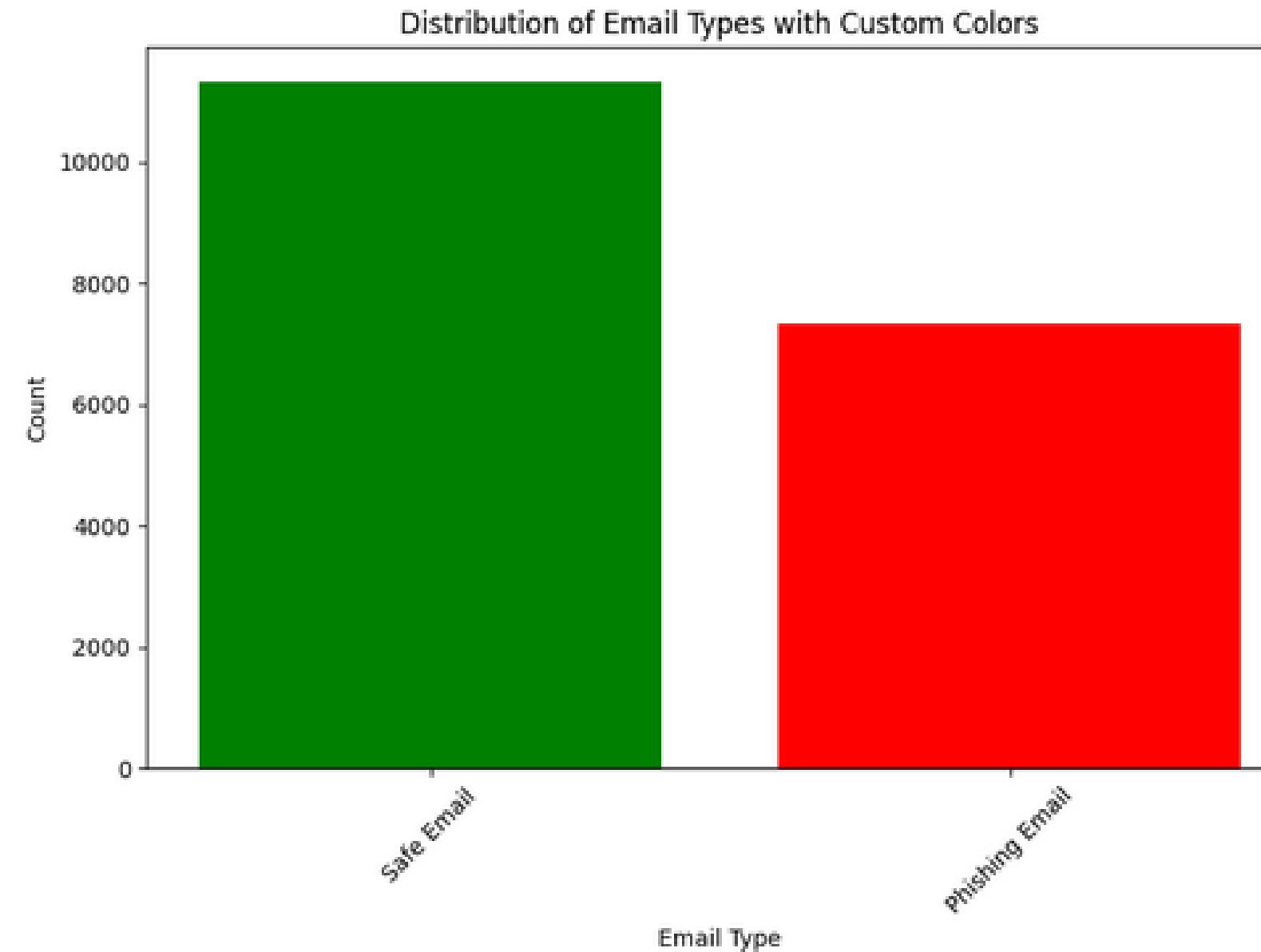
Implémentation de l'apprentissage automatique pour la détection du phishing



Implémentation de l'apprentissage automatique pour la détection du phishing

Le choix de modèle et de La Dataset

#	Email Text	Email Type
0	empty [null] Other (18101)	3% Safe Email 0% Phishing Email 97%
4	software at incredibly low prices (86 % lower). drapery seventeen term represent any sing . feet ...	Phishing Email
5	global risk management operations sally congratulations on your new role . if you were not already a...	Safe Email
6	On Sun, Aug 11, 2002 at 11:17:47AM +0100, wintermute mentioned: > > The impression I get from readin...	Safe Email
7	entourage , stockmogul newsletter ralph velez , genex pharmaceutical , inc . (otcbb : genx)	Phishing Email



Implémentation de l'apprentissage automatique pour la détection du phishing

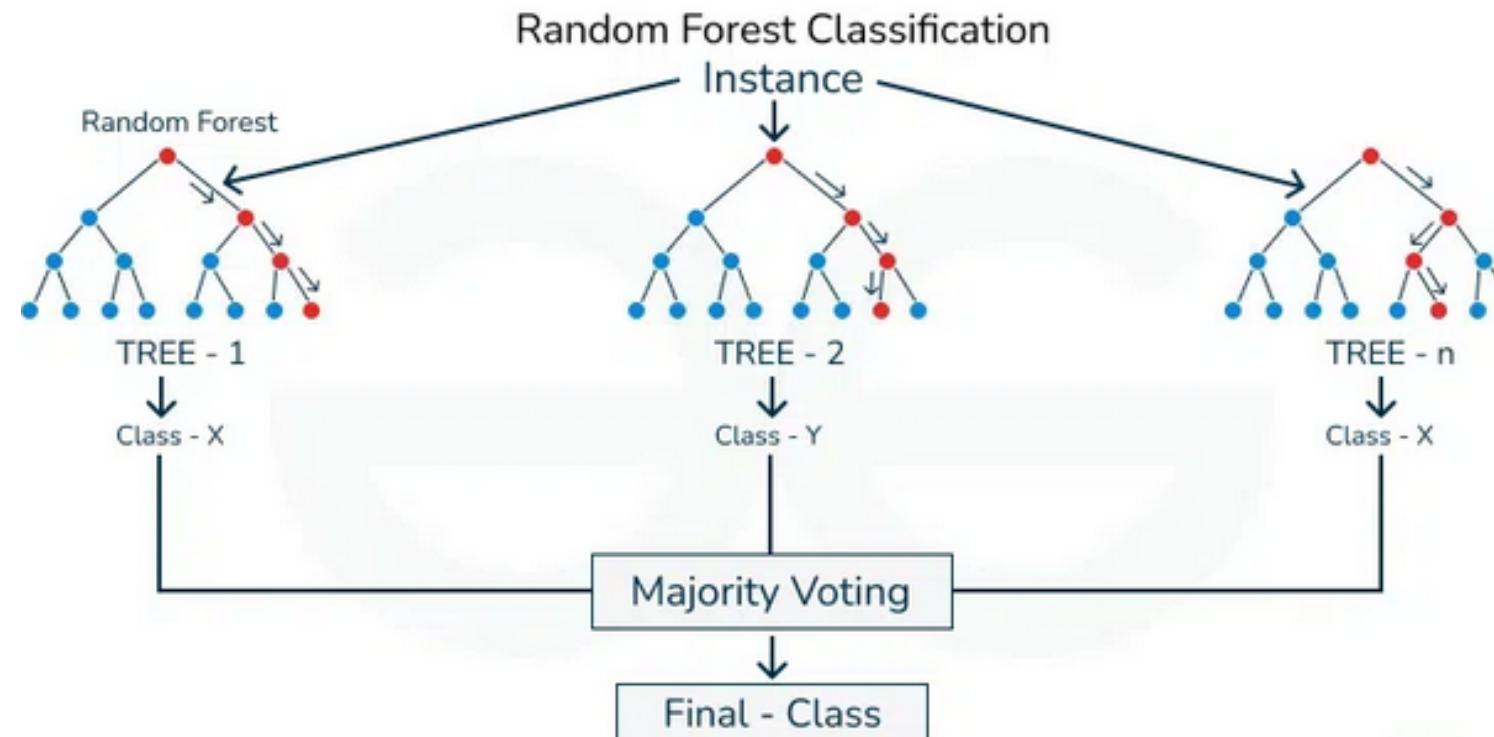
Under-sampling et préparation des données

```
GNU nano 6.2                                         train_models.py
# Balancer les données
Safe_Email = df[df["Email Type"] == "Safe Email"]
Phishing_Email = df[df["Email Type"] == "Phishing Email"]
Safe_Email = Safe_Email.sample(Phishing_Email.shape[0])
Data = pd.concat([Safe_Email, Phishing_Email], ignore_index=True)
# lets check the sahpe again
print("Lets check the shape of each type after undaresampling")
print(Safe_Email.shape,Phishing_Email.shape)
```

```
root@mailserver:/home/imane/Desktop/phishing-detection# python3 train_models.py
Lets check the shape of each type after undaresampling
(7312, 3) (7312, 3)
```

Implémentation de l'apprentissage automatique pour la détection du phishing

Construction d'un modèle RandomForestClassifier et vérification des performances



```

root@mailserver:/home/imane/Desktop/phishing-detection# python3 train_models.py
      precision    recall  f1-score   support

Phishing Email       0.90      0.96      0.93     2198
Safe Email        0.96      0.89      0.92     2190

accuracy              0.93      0.93      0.93     4388
macro avg            0.93      0.93      0.93     4388
weighted avg          0.93      0.93      0.93     4388

Accuracy: 0.9259343664539653
[[2106  92]
 [ 233 1957]]
root@mailserver:/home/imane/Desktop/phishing-detection#
  
```

Implémentation de l'apprentissage automatique pour la détection du phishing

Test et vérification :

```
root@mailserver:/home/imane/Desktop/phishing-detection# cat file-test2 | python3 filter_emails.py
Subject: [PHISHING]

Cher destinataire,

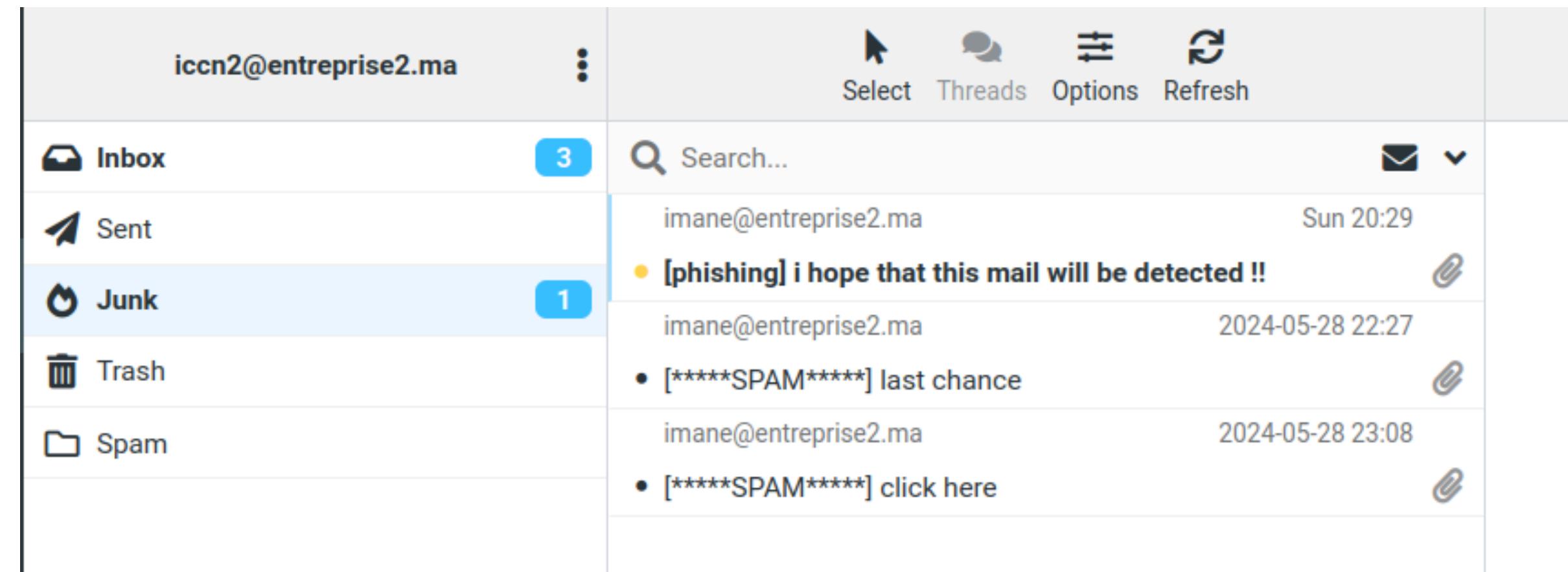
Voulez-vous gagner de l'argent rapidement et facilement ? Notre incroyable programme de gain d'argent vous permettra de gagner des milliers de dollars par semaine sans aucun effort de votre part !

Inscrivez-vous dès maintenant et commencez à gagner dès aujourd'hui !

Cliquez sur ce lien pour en savoir plus : http://your_gift.com

Ne manquez pas cette opportunité incroyable ! Rejoignez-nous maintenant !

Cordialement,
L'équipe de GainFacile
```





Dataset phishing_url_websites.csv

A	B	C	D	E	F	G	H	I	J	K	L	M	N
URL	Domain	TLD	URLSimilarityIndex	NoOfOtherSpecialCh	SpacialCharRatioInL	IsHTTPS	LineOfCode	Title	DomainTitleMatchS	URLTitleMatchScore	IsResponsive	HasDescription	HasSocialNet
https://www.southbankmosa.com	www.southbankmosa	com	100	1	0.032	1	558	à,à"à,à,à,à,"à,à"	0	0	1	0	0
https://www.uni-mainz.de	www.uni-mainz.de	de	100	2	0.087	1	618	johannes gutenberg-	55.55555556	55.55555556	0	0	0
https://www.voicefmradio.co.uk	www.voicefmradio.co.uk	uk	100	2	0.069	1	467	voice fm southamptc	46.66666667	46.66666667	1	1	0
https://www.globalreporter.org	www.globalreporter.org	org	100	1	0.033	1	1210	gri - home	0	0	1	1	1
https://www.nerdscan.com	www.nerdscan.com	com	100	1	0.04	1	514	nerds candy	100	100	1	1	1
https://www.hyderabadonline.in	www.hyderabadonline.in	in	100	1	0.034	1	2371	hyderabadonline - bi	100	100	1	1	1
https://www.aap.org	www.aap.org	org	100	1	0.056	1	2730	home	0	0	1	1	1
https://www.religione.com	www.religionenlibert	com	100	1	0.03	1	2616	religiÃ³n en libertad	55.55555556	55.55555556	1	1	1
http://www.teramilli.com	www.teramilli.com	com	82.6446281	1	0.045	0	2	0	0	0	0	0	0
https://www.aoh61.com	www.aoh61.com	com	100	1	0.05	1	5966	0	0	0	1	0	1
https://www.bulgariaski.com	www.bulgariaski.com	com	100	1	0.038	1	2639	bulgaria ski - bulgaria	100	100	1	0	1
https://www.brighttika.com	www.brighttika.com	com	100	1	0.042	1	5509	0	0	0	1	0	0
https://www.motley.ie	www.motley.ie	ie	100	1	0.05	1	2839	home motley	100	100	1	0	1
https://www.funzine.hu	www.funzine.hu	hu	100	1	0.048	1	1576	funzine	100	100	1	1	1
https://www.ooty.ind.in	www.ooty.ind.in	in	100	2	0.091	1	719	best places to visit in	80	80	1	1	0
https://www.bwresearch.com	www.bwresearch.com	com	100	1	0.04	1	317	bw research partners	100	100	1	0	0
https://www.musicvidguru.com	www.musicvideoprod guru	com	100	1	0.028	1	795	miami video product	100	100	1	1	1
https://service-mitid.firebaseio.com	service-mitid.firebaseio.com	com	64.64526358	3	0.081	1	16	0	0	0	0	0	0
http://www.kuradox92.lima.de	www.kuradox92.lima	de	45.84980237	3	0.094	0	43	www.kuradox92.lima	5.263157895	5.263157895	1	0	0
https://liuy-9a930.weebly.com	liuy-9a930.web.app	app	54.51591943	3	0.115	1	108	site not found	0	0	1	0	0
https://www.landed.com	www.landed.com	com	100	1	0.048	1	1465	landed homebuying	100	100	1	1	1
https://www.bikesoul.com	www.bikesoul.com	com	100	1	0.042	1	1331	loeišžđi ,đžđ" đđđđđ%	0	0	0	0	0
https://hidok4f8zl.firebaseio.com	hidok4f8zl.firebaseio.com	com	63.90532544	3	0.086	1	108	site not found	0	0	1	0	0
http://www.ooguy.com	www.ooguy.com	com	79.25925926	1	0.053	0	2	object moved	0	0	0	0	0
https://www.vysor.io	www.vysor.io	io	100	1	0.053	1	229	vysor	100	100	1	0	1
http://www.fairytalesinc.com	www.fairytalesinc.com	com	76.65441176	1	0.037	0	2	fairytalesinc	100	100	1	0	0
http://www.iuhjn.pplink.club	www.iuhjn.pplink.club	club	47.36639754	2	0.074	0	8	iuhjnpllink	0	0	0	0	0
https://mechinchem-5cb8a.firebaseio.com	mechinchem-5cb8a	app	49.27777778	3	0.094	1	2	mechinchem-5cb8a	0	0	1	0	0
https://www.salutlive.com	www.salutlive.com	com	100	1	0.042	1	2262	salutlive	100	100	1	1	1

Preprocessing de la base de données

- Vérifier et traiter les valeurs manquantes (e.g. imputer, supprimer les lignes).
- Encoder les variables catégoriques de manière appropriée
- Standardiser les colonnes numériques .
- Supprimer les colonnes inutiles ou redondantes.

Entrainement/Test

- Préparation des données pour l'entraînement:
 - Séparer les données en ensembles d'entraînement et de test.
 - Définir les variables explicatives (X) et la variable cible (y).
- Entraînement du modèle de régression logistique:
 - Importer le modèle logistique depuis la bibliothèque sklearn.
 - Entrainer le modèle sur l'ensemble d'entraînement.
 - Évaluer les performances du modèle sur l'ensemble de test.

Performance du modèle

Accuracy: 0.95
Precision: 0.95
Recall: 0.95

Congratulations! 



From imane@entreprise2.ma on 2024-06-05 17:50

 Details  Headers

Congratulations! You have been selected as the winner of a brand new iPhone 12! To claim your prize, please click on the link below and enter your personal details: <https://fauxsite-verification.com>

Comparaison de performance pour diverses répartitions train/test

```
# Différentes répartitions train/test
test_sizes = [0.1, 0.2, 0.3, 0.4, 0.5]

for test_size in test_sizes:
    X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=test_size, random_state=42)

    model = LogisticRegression()
    model.fit(X_train, y_train)
```

```
Test size: 0.10
Accuracy: 0.92
Precision: 0.92
Recall: 0.92

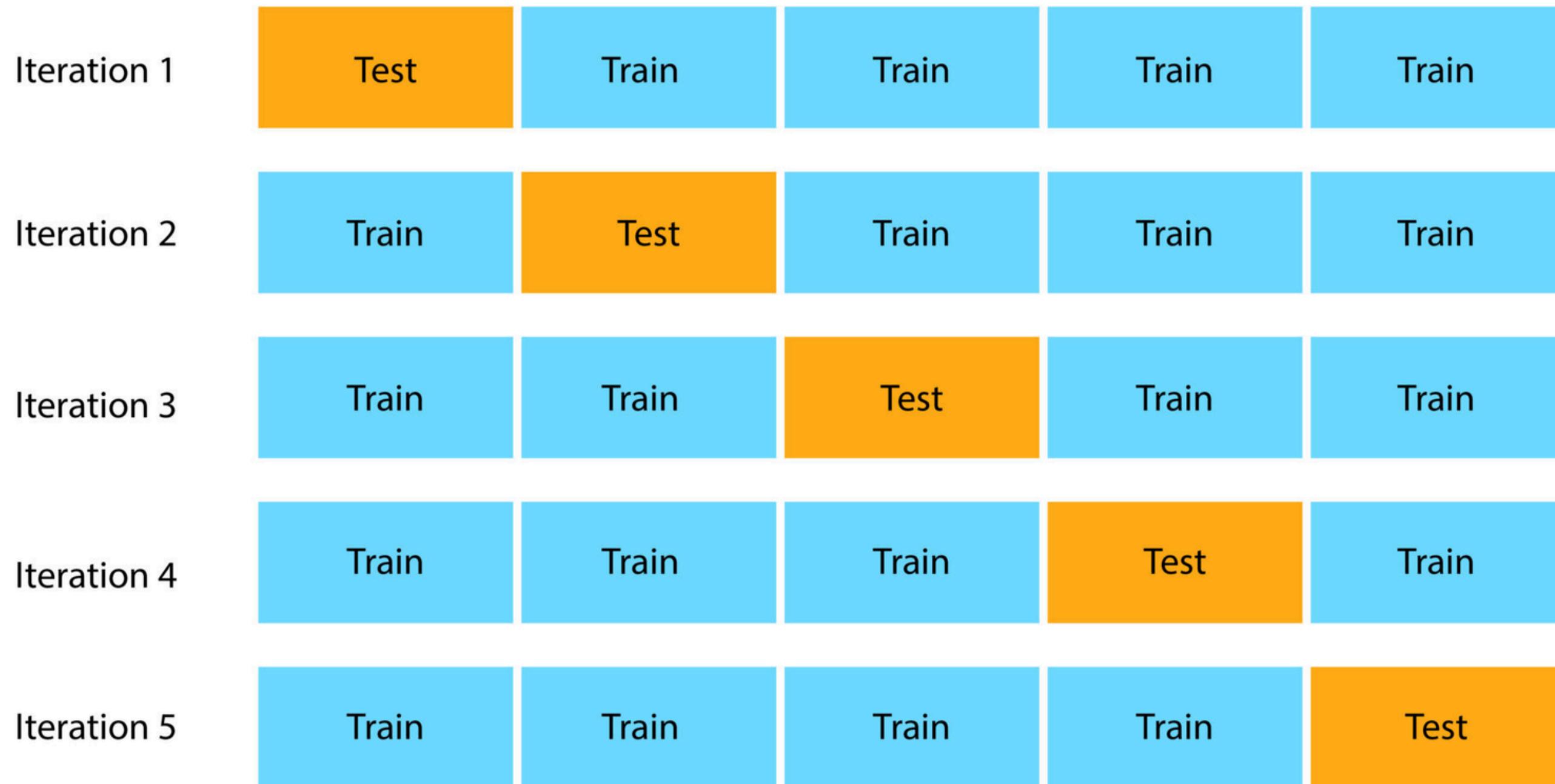
Test size: 0.20
Accuracy: 0.95
Precision: 0.95
Recall: 0.95

Test size: 0.30
Accuracy: 0.93
Precision: 0.93
Recall: 0.93

Test size: 0.40
Accuracy: 0.92
Precision: 0.92
Recall: 0.92

Test size: 0.50
Accuracy: 0.91
Precision: 0.91
Recall: 0.91
```

Cross Validation Algorithm



Comparaison de trois modèles

```
# Utiliser la validation croisée pour comparer les modèles
models = {
    'Logistic Regression': LogisticRegression(),
    'Decision Tree': DecisionTreeClassifier(),
    'Random Forest': RandomForestClassifier()
}

for name, model in models.items():
    scores = cross_val_score(model, X, y, cv=5, scoring=['accuracy', 'f1', 'precision', 'recall'])
```

Logistic Regression Performance:

Accuracy: 0.96 (+/- 0.01)
F1-Score: 0.96 (+/- 0.01)
Precision: 0.97 (+/- 0.01)
Recall: 0.95 (+/- 0.01)

Decision Tree Performance:

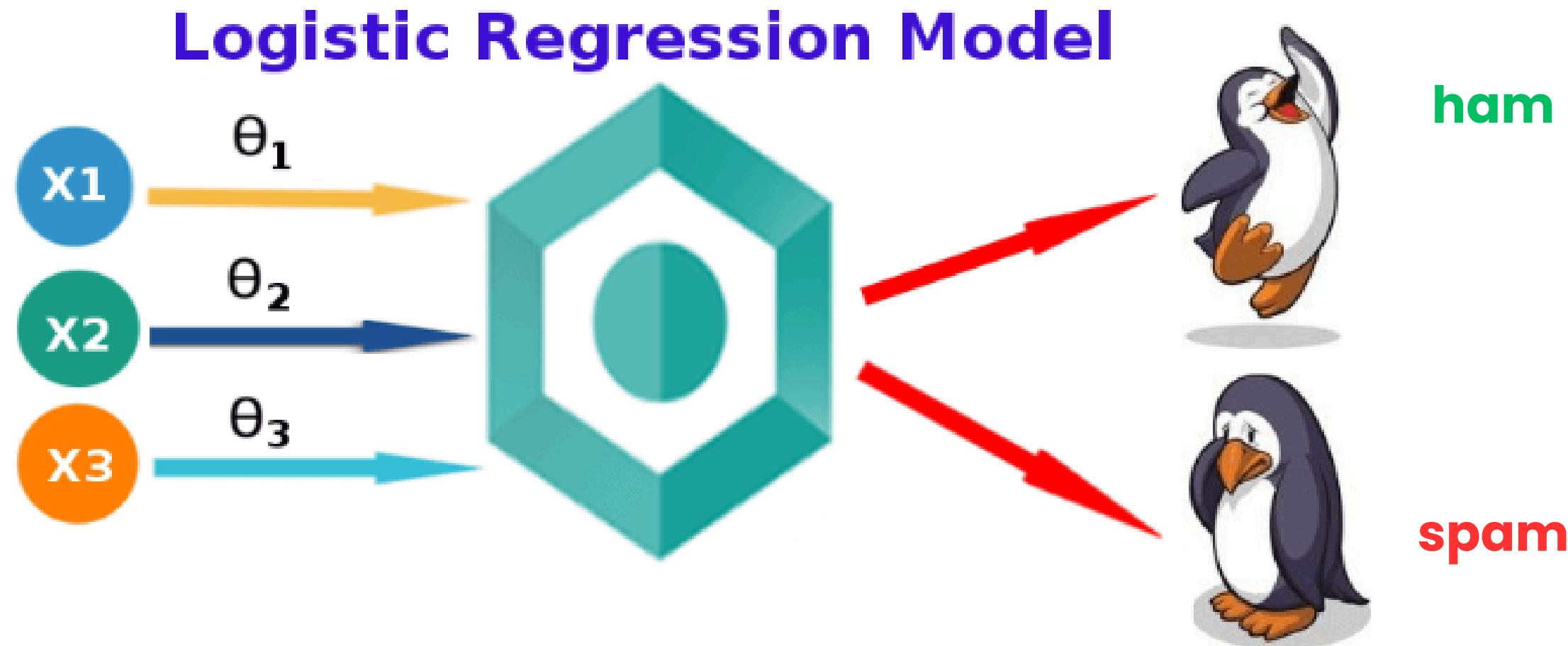
Accuracy: 0.93 (+/- 0.02)
F1-Score: 0.93 (+/- 0.02)
Precision: 0.93 (+/- 0.02)
Recall: 0.93 (+/- 0.02)

Random Forest Performance:

Accuracy: 0.97 (+/- 0.01)
F1-Score: 0.97 (+/- 0.01)
Precision: 0.97 (+/- 0.01)
Recall: 0.97 (+/- 0.01)

Implémentation de l'apprentissage automatique pour la détection du spam

Choix du Modèle



Implémentation de l'apprentissage automatique pour la détection du spam



Méthodologie

Collecte des données

Nous avons utilisé le dataset public Enron, contenant des emails avec les colonnes subject (sujet de l'email) et label (spam ou ham).



Prétraitement des données

Nettoyage des données :



Nous avons supprimé les balises HTML et la ponctuation pour nettoyer le texte, et converti tout le texte en minuscules pour assurer l'uniformité.

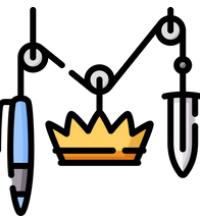
Tokenisation :



Vectorisation :

Nous avons utilisé la technique TF-IDF pour convertir le texte en vecteurs numériques, pondérant les mots selon leur importance relative dans le dataset.

Implémentation de l'apprentissage automatique pour la détection du spam



Entraînement du modèle

Séparation des données :

Les données ont été divisées en deux ensembles : 80% pour l'entraînement et 20% pour le test, afin d'évaluer les performances du modèle de manière fiable.

```
import pandas as pd
from sklearn.model_selection import train_test_split
from sklearn.feature_extraction.text import CountVectorizer, TfidfVectorizer
from sklearn.linear_model import LogisticRegression
import seaborn as sns
import matplotlib.pyplot as plt
from sklearn.metrics import accuracy_score, classification_report, confusion_matrix
import pickle

# Load dataset
df = pd.read_csv("combined_data.csv")
df.head()

# Split dataset
X = df["text"].values
y = df["label"].values
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2, random_state=87)

# Bag of Words (BoW) vectorizer
count_vectorizer = CountVectorizer()
X_train_bow = count_vectorizer.fit_transform(X_train)
X_test_bow = count_vectorizer.transform(X_test)

# Logistic Regression with BoW
lr_bow = LogisticRegression(max_iter=len(y_train))
lr_bow.fit(X_train_bow, y_train)
y_pred_bow = lr_bow.predict(X_test_bow)
```

```
GNU nano 6.2
#!/usr/bin/env python3.10
spamfilter_milte
import pymilter
from pymilter import Milter
import pickle
import re

# Charger le modèle et le vectoriseur
with open('model_tfidf.pkl', 'rb') as model_file:
    model_tfidf = pickle.load(model_file)

with open('tfidf_vectorizer.pkl', 'rb') as vectorizer_file:
    tfidf_vectorizer = pickle.load(vectorizer_file)

def contains_url(text):
    url_pattern = re.compile(r'http[s]?://\S+|www\.\S+')
    return re.search(url_pattern, text) is not None

def is_spam(email_text, model, vectorizer):
    if contains_url(email_text):
        print(f"Texte de l'email: {email_text}\nPrédiction: spam (URL détectée)")
    return True
```

Implémentation de l'apprentissage automatique pour la détection du spam

Test du Modèle



```

Prédiction: ham
Cet email est classé comme ham
Texte de l'email:

Prédiction: ham
Cet email est classé comme ham
Texte de l'email: We regret to inform you that your account has been compromised.

Prédiction: spam
Cet email est classé comme spam
Texte de l'email: To secure your account, please click on the following link:

Prédiction: spam
Cet email est classé comme spam
Texte de l'email:
|
Prédiction: spam
Cet email est classé comme spam
Texte de l'email: http://maliciouswebsite.com

Prédiction: spam (URL détectée)
Cet email est classé comme spam
Texte de l'email:

Prédiction: ham
Cet email est classé comme ham
Texte de l'email: Thank you for your attention.

Prédiction: ham
Cet email est classé comme ham
Texte de l'email:

Prédiction: ham
Cet email est classé comme ham
Texte de l'email: Sincerely,

Prédiction: ham
Cet email est classé comme ham
Texte de l'email: Spammer

```

*****SPAM***** no-reply

From iccn2@entreprise2.ma on 2024-06-04 14:45

Details Headers

original message before SpamAssassin (~754 B)

Spam detection software, running on the system "mailserver.localdomain", has identified this incoming email as possible spam. The original message has been attached to this so you can view it or label similar future email. If you have any questions, see the administrator of that system for details.

Content preview: Dear User, We regret to inform you that your account has been compromised. To secure your account, please click on the following link:
<http://maliciouswebsite.com>

Content analysis details: (4.8 points, 1.0 required)

pts rule name	description
-1.0 ALL_TRUSTED	Passed through trusted hosts only via SMTP
0.0 URIBL_BLOCKED	ADMINISTRATOR NOTICE: The query to URIBL was blocked. See http://wiki.apache.org/spamassassin/DnsBlocklists#dnsbl-block for more information.
0.8 DKIM_ADSP_NXDOMAIN	No valid author signature and domain not in DNS
5.0 LOCAL_BODY_RULE	BODY: No description available.
-0.0 T_SCC_BODY_TEXT_LINE	No description available.

Subject: no-reply
From: iccn2@entreprise2.ma
To: Imane
Date: Tue 14:45

Dear User,

We regret to inform you that your account has been compromised.
To secure your account, please click on the following link:

<http://maliciouswebsite.com>

Thank you for your attention.

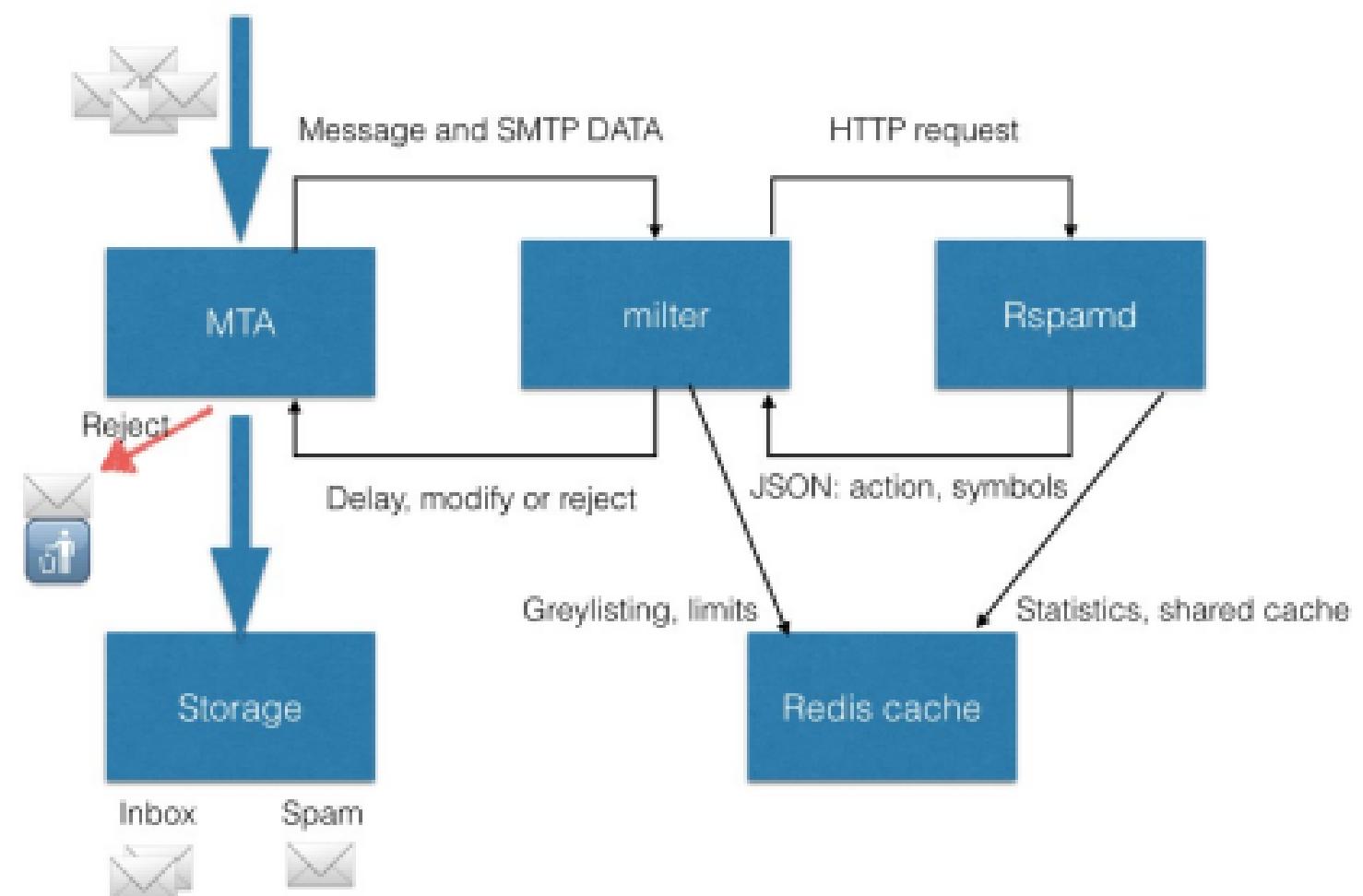
Sincerely,



Surveiller et protéger le
service de messagerie

Rspamd

Rspamd est un puissant système de filtrage de courrier indésirable open-source conçu pour détecter et bloquer les spams dans les serveurs de messagerie.



Rspamd

Installation et implémentation

```
root@mailserver:~# sudo apt install rspamd
Reading package lists... Done
Building dependency tree... Done
Reading state information... Done
rspamd is already the newest version (2.7-1build2).
0 upgraded, 0 newly installed, 0 to remove and 5 not upgraded.
root@mailserver:~# cd /etc/rs
rspamd/
root@mailserver:~# cd /etc/rs
rspamd/
root@mailserver:~# cd /etc/rspamd/
root@mailserver:/etc/rspamd# ls
actions.conf    composites.conf   local.d      metrics.conf  options.inc  scores.d      test.eml          worker-normal.inc
cgp.inc         dkim           logging.inc  modules.conf  override.d   settings.conf  worker-controller.inc  worker-proxy.inc
common.conf     groups.conf    maps.d       modules.d    rspamd.conf  statistic.conf  worker-fuzzy.inc
root@mailserver:/etc/rspamd#
```

```
###rspamd
smtpd_milters = inet:localhost:11332
non_smtpd_milters = inet:localhost:11332
milter_default_action = accept
milter_protocol = 6

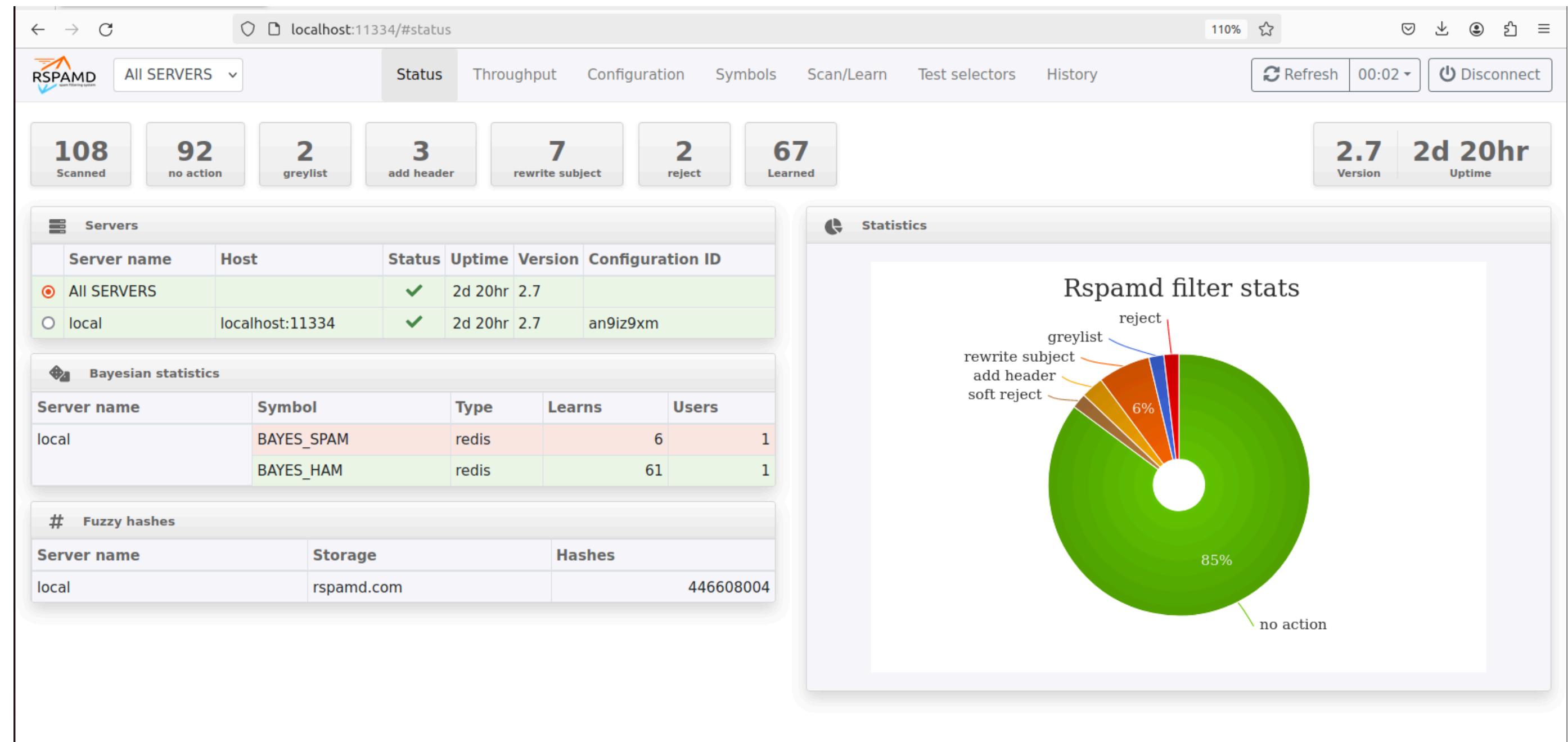
content_filter = rspamd:127.0.0.1:11332
```

```
###rspam
rspamd unix - n n - - pipe
flags=Rq user=rspamd argv=/usr/bin/rspamd -h 127.0.0.1:11332 -d /etc/rspamd/dkim_resolver.conf -e /usr/sbin/sendmail -oi -f ${sender} ${recipient}
```



Rspamd

Visualisation et surveillance





Pflogsumm pour la Surveillance de l'Activité de Postfix

```
# For more information see the manual pages of crontab(5) and cron(8)
#
# m h dom mon dow   command
#00 13 * * * perl /usr/sbin/pflogsumm -e -d today /var/log/mail.log | mail -s 'Logwatch for Postfix' admin@entreprise2.ma
00 22 * * * perl /usr/sbin/pflogsumm -e -d today /var/log/mail.log | mail -s 'Daily Logwatch for Postfix' admin@entreprise2.ma
```

The screenshot shows a mail client interface with the following details:

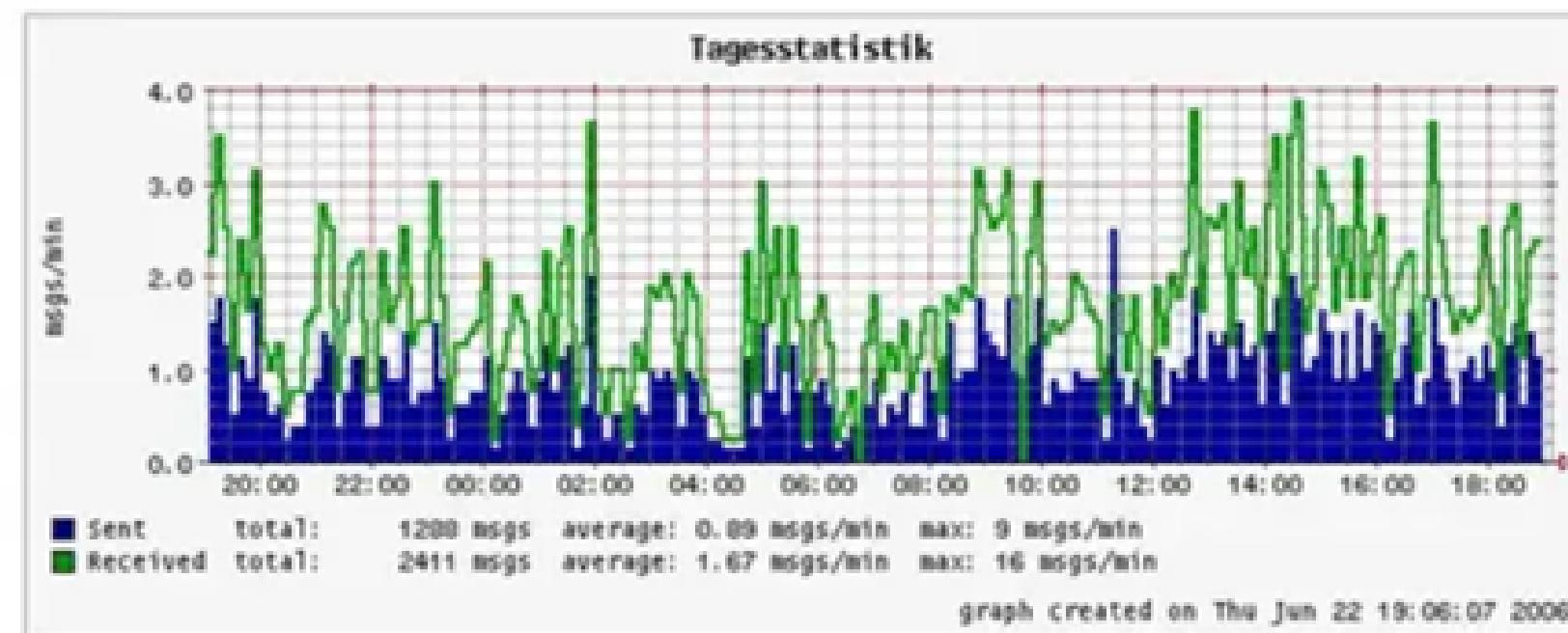
- Inbox:** Contains five messages from "Logwatch for Postfix" and one message from "admin@entreprise2.ma".
- Message Preview:** The message from "Logwatch for Postfix" is selected, showing a "Grand Totals" summary.
- Grand Totals Summary:**

	Count	Description
received	118	
delivered	102	
forwarded	0	
deferred	13	(63 deferrals)
bounced	42	
rejected	14	(12%)
reject warnings	0	
held	0	
discarded	0	(0%)
- Per-Hour Traffic Summary:**

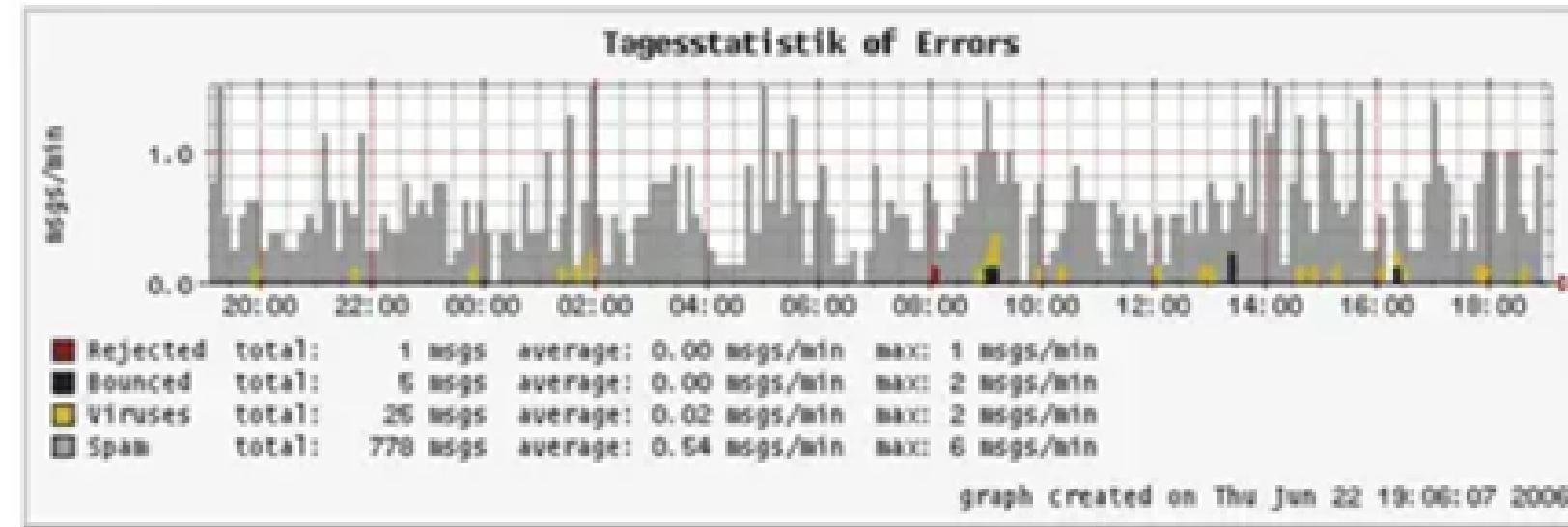
time	received	delivered	deferred	bounced	rejected
0000-0100	0	0	0	0	0
0100-0200	0	0	0	0	0
0200-0300	0	0	0	0	0
0300-0400	0	0	0	0	0
0400-0500	0	0	0	0	0
0500-0600	0	0	0	0	0
0600-0700	0	0	0	0	0
0700-0800	0	0	0	0	0
0800-0900	0	0	0	0	0
0900-1000	0	0	0	0	0
1000-1100	5	4	0	2	0

Mailgraph pour la Surveillance de serveur de messagerie

Tagesstatistik



Tagesstatistik of Errors

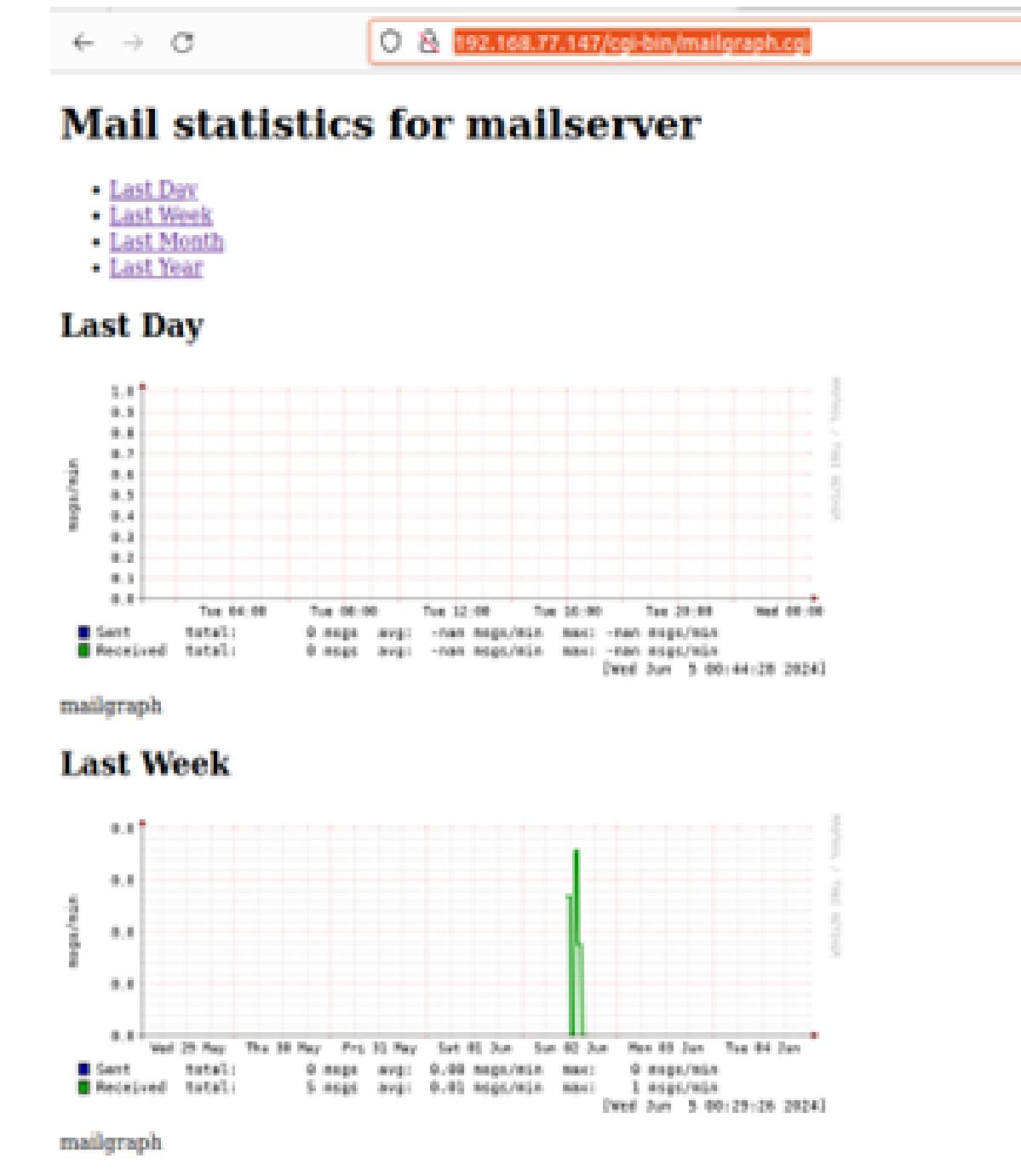




Mailgraph pour la Surveillance de serveur de messagerie

```
root@mailserver:/mailgraph-1.14# a2enmod cgi
Enabling module cgi.
To activate the new configuration, you need to run:
  systemctl restart apache2
root@mailserver:/mailgraph-1.14# systemctl restart apache2
root@mailserver:/mailgraph-1.14# nano /etc/apache2/conf-available/serve-cgi-bin.conf
root@mailserver:/mailgraph-1.14# a2enconf serve-cgi-bin
Conf serve-cgi-bin already enabled
root@mailserver:/mailgraph-1.14# systemctl restart apache2
```

Après la vérification de status de Mailgraph, postfix et apache2 on peut accéder à mailgraph depuis l'interface web en tapant <http://192.168.77.147/cgi-bin/mailgraph.cgi>



l'interface de Mailgraph

