Services d'intranet

Christian Bulfone
christian.bulfone@gipsa-lab.fr
www.gipsa-lab.fr/~christian.bulfone/MIASHS-L3

Master MIASHS L3
Année 2024/2025

Plan du cours

- La messagerie électronique
- Le service d'annuaire
- Autres services : HTTP, NTP

La messagerie électronique

Notion de service de messagerie

- A la base un service d'échange de textes courts (un transfert électronique de fichiers caractères ASCII)
- Extension à des transferts de fichiers quelconques (en structure et en contenu)
  - avec une limitation sur la taille du fichier
  - par un encodage en **format caractère**
- Transmission **asynchrone**
  - l'émetteur et le récepteur n'ont pas à être connectés en même temps
- Terminologie
  - courrier électronique, courriel, mèl (mail, email)

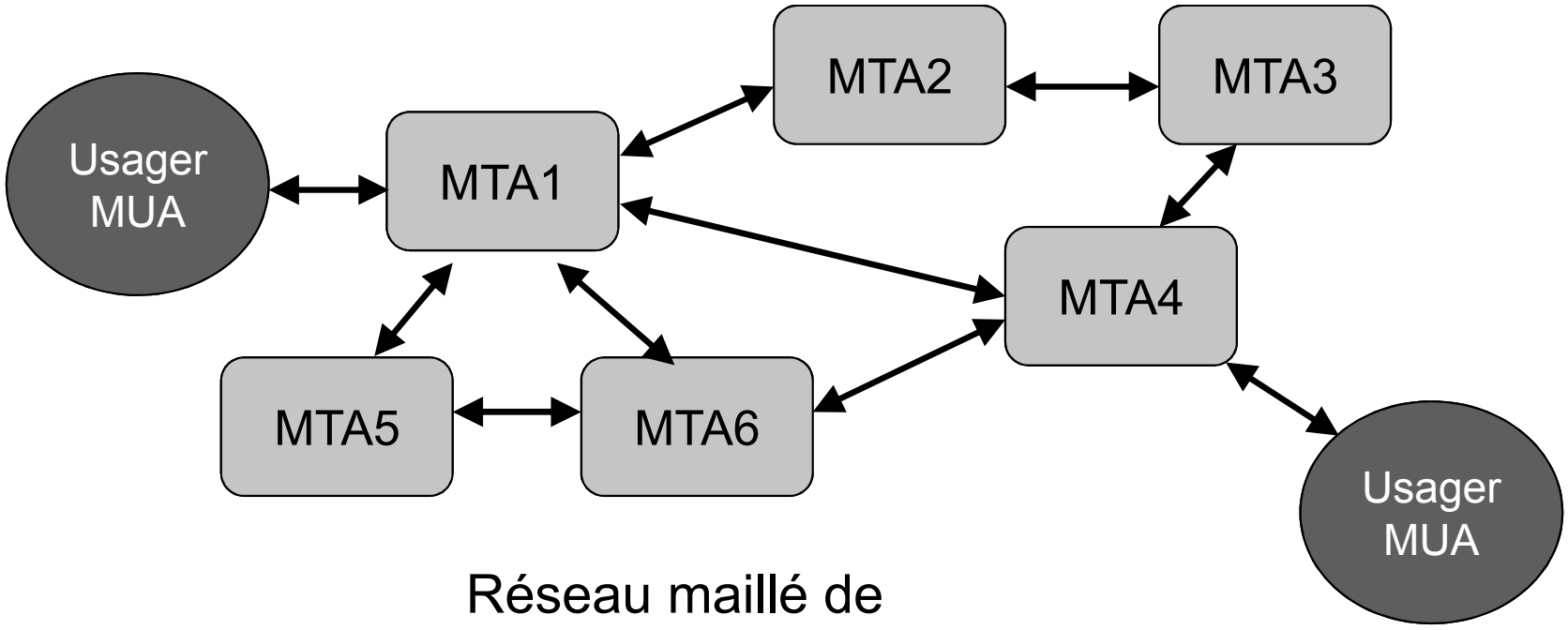
Fonctions d'un service de messagerie

- Utilisation d'un **système d'adressage** permettant l'envoi à **un destinataire** ou à **un groupe de destinataires**
- Composition de message, attachement de pièces jointes
- Emission du courrier
- Lecture du courrier
  - notion de **files d'attente** de courriers ou **boîtes à lettres**
- Gestion des **archives de courriers**
  - classement des courriers selon différents critères dans des boîtes à lettres différentes

Distinction messageries / transfert fichiers

	Messagerie électronique	**Transfert de fichiers**	**Messagerie instantanée**
Mode de transmission	Asynchrone	Synchrone	Synchrone
Type de données	Texte uniquement	Texte et binaire	Texte
Taille des fichiers	Limités	Non limités	Généralement limités
Désignation	Message dans une boîte	Fichier dans une arborescence	Message dans une file de discussion

Architecture : stockage et retransmission



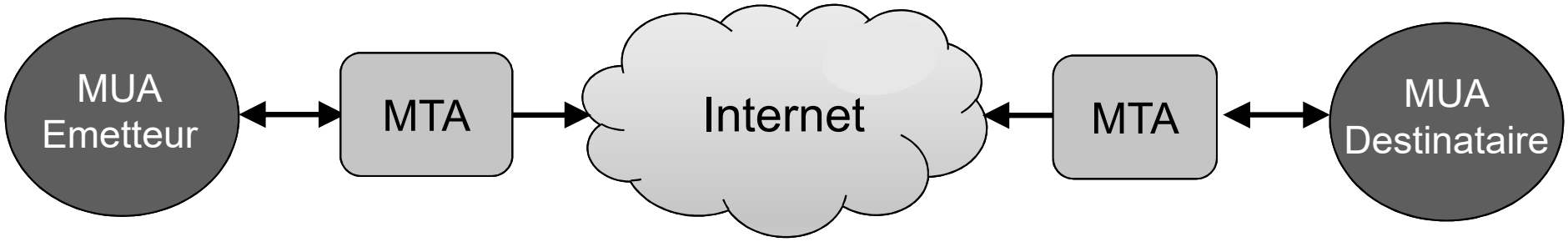
Réseau maillé de
MTA

La messagerie réalise l'acheminement des courriers comme le fait un réseau à **commutation de paquets**

Architecture : stockage et retransmission

- Notion de serveur de messagerie et commutateur de courriers
  - Agent de **transfert de messages** MTA (*Mail Transfer Agent*)
- Notion de client de messagerie
  - Agent **utilisateur de messagerie** ou MUA (*Mail User Agent*)
- MTA et MUA sont des dénominations héritées de cette architecture de messagerie normalisée OSI MHS (*Message Handling System*) ou norme ITU X400

Architecture : acheminement de bout en bout

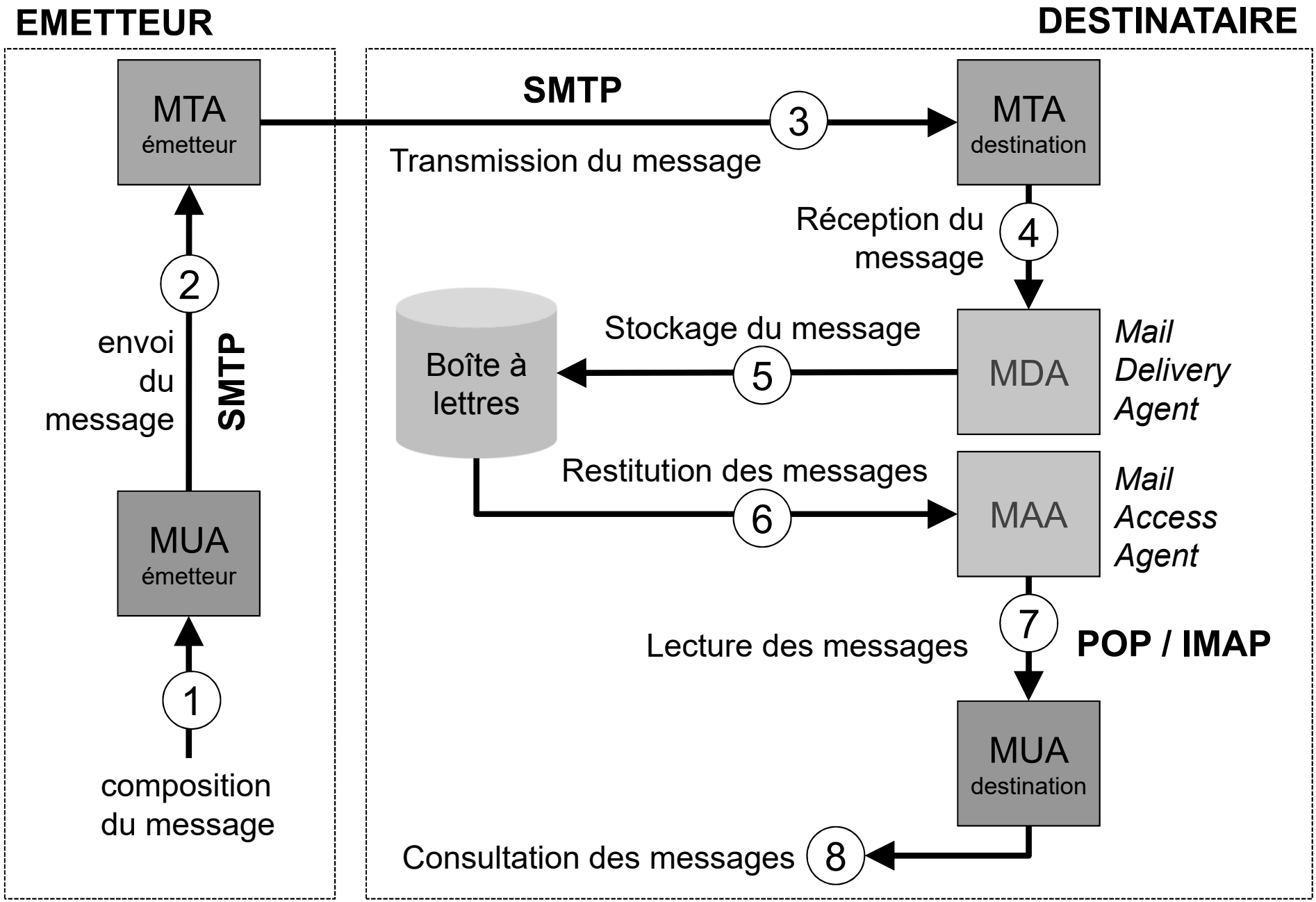


- Le serveur de messagerie MTA achemine directement un message entre un émetteur et un destinataire
- Pour cela il utilise un service de transport existant (typiquement TCP)
  - Un courrier est acheminé comme segments TCP
- Exemple de la messagerie Internet SMTP

Comparaison des deux approches

- Acheminement de bout en bout
  - Un serveur de courrier utilise une connexion de bout en bout (transport) pour remettre ses messages
  - ➔ fiable et simple (si l'on dispose d'une couche transport fiable comme TCP)
- Stockage et retransmission
  - Problème de routage des messages et de contrôle d'erreur (pertes possibles de messages si pannes des MTA intermédiaires)
  - Atout : facilite l'interconnexion avec d'autres systèmes de messagerie
- Aujourd'hui triomphe de la solution de bout en bout (avec la messagerie Internet)

Architecture modulaire



Cheminement des messages

❶ Un usager compose, avec l'aide de son client de messagerie (MUA) un message

- Le message peut être stocké dans la file d'attente du MUA si son expédition n'est pas immédiatement possible

❷ Le message est transmis au MTA de l'usager en SMTP

❸ Le message est transmis au MTA du destinataire (en SMTP)

- Stockage en file d'attente permettant la réémission en cas d'échec

❹ Le serveur transmet le message à un agent le MDA (*Mail Delivery Agent*) chargé de le stocker

Cheminement des messages

❺ Le MDA stocke le courrier dans la boîte à lettres du destinataire
- Deux formats généralement utilisés mbox et maildir

❻ Sur requête du destinataire dans le cadre d'un protocole de relève (POP ou IMAP) les messages sont extraits de la boîte à lettres par le MAA (*Mail Access Agent*)

❼ Les messages sont transmis au client de messagerie utilisateur (protocoles POP ou IMAP) et stockés dans des boîtes à lettres client

❽ Le destinataire consulte ses messages en utilisant son client de messagerie (MUA)

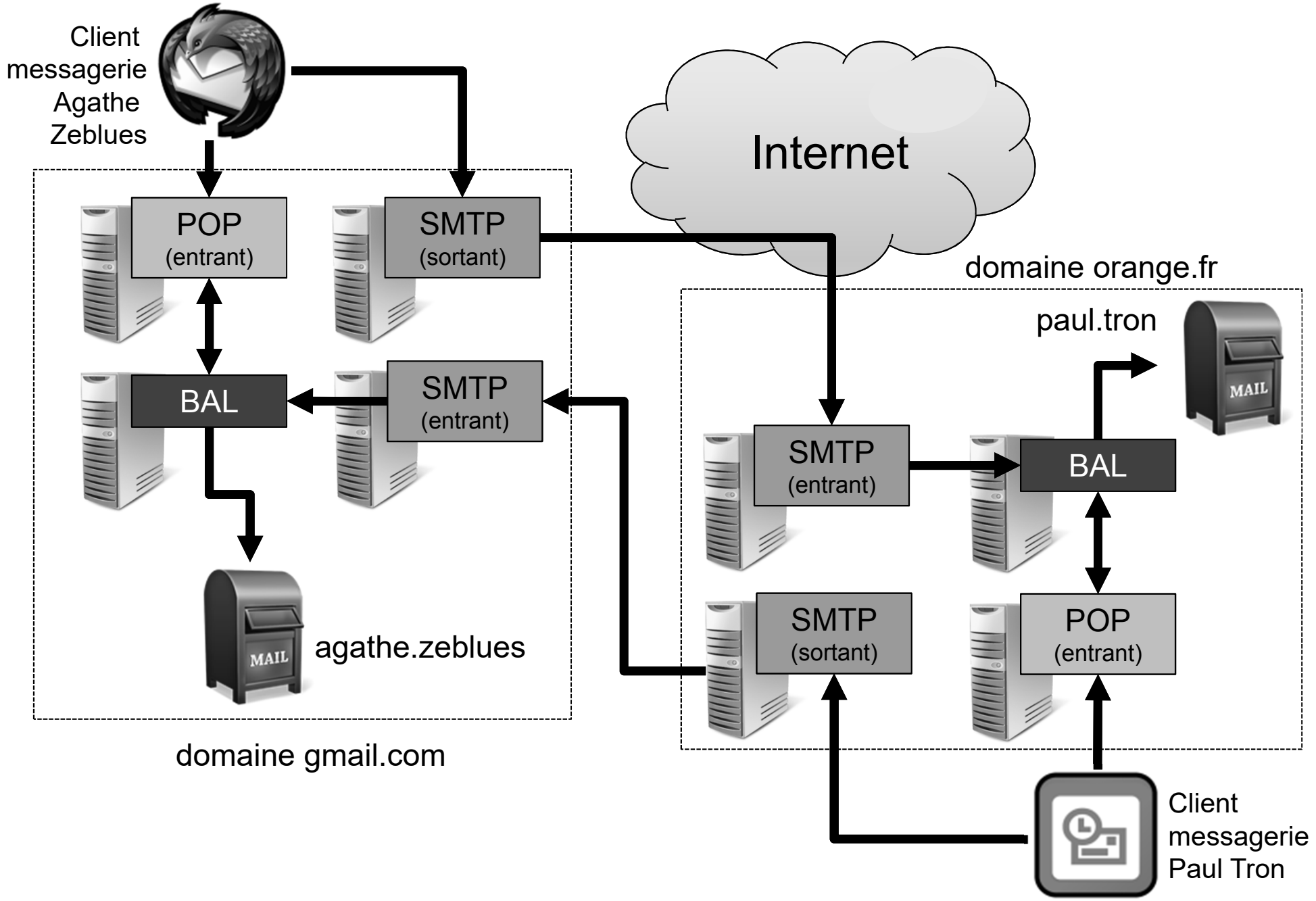
Les protocoles de messagerie de l'Internet

- Emission / transfert
  - *Simple Mail Transfer Protocol* (SMTP)
    - Défini par la RFC 821
    - Protocole basé sur des messages de format textes qui définit les échanges entre serveurs de messagerie
  - *Extended Simple Mail Transfer Protocol* (ESMTP)
    - Défini par la RFC 1869
    - Evolution de SMTP qui définit des commandes supplémentaires

Les protocoles de messagerie de l'Internet

- Relève
  - *Post Office Protocol* (POP)
    - Protocole de base de relève de courrier pour le dialogue entre un client de messagerie MUA et un serveur de messagerie dans sa partie MAA
  - *Internet Message Access Protocol* (IMAP)
    - Autre protocole de relève offrant des possibilités plus larges que POP (gestion des archives de courrier, limitation des volumes de données échangées …)

Mise en œuvre pratique



Adresses globales de courrier

- Syntaxe résumée dans la RFC 3696 (basée sur les RFC 2821 et 2822)
- Codées dans un nombre limité de caractères, sous-ensemble de l'ASCII
- Pour l'émetteur et le destinataire, toujours de la forme

- Le DNS permet de déterminer les serveurs de courrier d'un domaine (enregistrement de type MX)

Syntaxe des adresses globales

- bob@gmail.com
  - Forme la plus simple
- "bob"@gmail.com
  - Guillemets délimitant la chaîne de caractères boîte à lettres
- bob (Bob Morane) @gmail.com
  - Chaîne entre parenthèses = un commentaire ➔ignorée
- Bob Morane <bob@gmail.com>
  - Seul compte la partie entre < et > le reste étant ignoré
- bob@136.173.24.11
  - Forme d'adresse dite littérale avec directement codée l'adresse IP du serveur

Le protocole SMTP

- Protocole fonctionnant en client/serveur entre 2 MTA
- Basé sur des commandes textuelles envoyées en TCP par défaut sur le port 25
- Chacune des commandes envoyées par le client (terminée par la chaîne de caractères ASCII CR LF) est suivie d'une réponse du serveur SMTP composée d'un numéro et d'un message descriptif
- Les MTA gèrent des files d’attente de messages en émission et réception

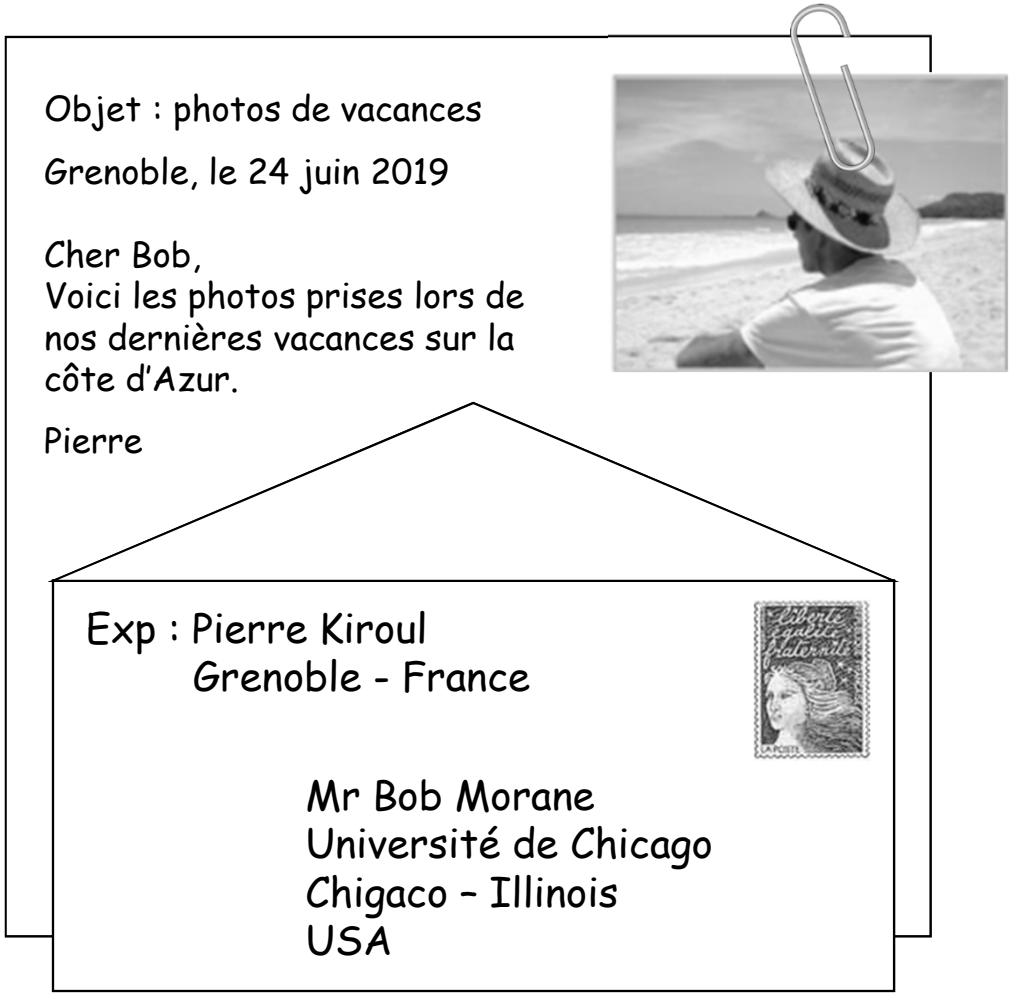
Le format de base (RFC 822)

- Un courrier est composé de lignes de caractères US-ASCII sur 7 bits (8e bit mis à 0)
- Chaque ligne fait au maximum 1000 caractères et est terminée par la séquence CR LF
- Structuré en deux parties
  - une **entête** et un **corps** séparés par une ligne vide
- L'entête est une liste de lignes précisant les caractéristiques du message sous la forme

```
Nom_de_zone: Valeur_de_zone
```

- Le corps contient les données effectivement échangées

Structure d'un message



entête

```
Received: from smtp.orange.fr (smtp.orange.fr
[193.252.22.83) by smtp.gmail.com (8.12.8/8.12.8) for
bob@gmail.com
Date: Thu, 24 Jun 2019 11:56:17 +0200
To: Bob Morane <bob@gmail.com>
From: Pierre Kiroul <pierre@orange.fr>
Subject: photos de vacances
```

Cher Bob,
Voici les photos prises lors de nos dernières vacances sur la côte d'Azur.

Pierre

```
----------=_050001030504020801010208
Content-Type: image/png; name= "plage.png"
Content-Transfer-Encoding: base64
Content-Disposition: inline;
filename= "plage.png"
iVBORw0KGgoAAAANSUhEUgAAAHsAAADyCAYAAABt
FlzAAAuGAAALhgBKqonIAAAACB0RVh0U29mdHdhcmUTVi7k
SokAAAEB3ByVld4nO1bbVLbMBB1MVQJRAbjaJt+TlpLf/9qxfy
715Hl+kBOrpLb9BU61AGqFfuDNk1U/xmgGB7vY+vd3VWvrx6/
vP7Gv2bYcIu53f7TL8G8Iu+F3Idj7D//D/D8
```

encodage de l'image sous forme de caractères

corps

L'entête

- Au moins trois lignes obligatoires
  - `From:` adresse de l'émetteur
  - `To:` adresse du destinataire
  - `Date:` date de création du message
- Nombreuses autres possibilités normalisées mais facultatives
- Possibilité de créer des entêtes propriétaires à condition de les préfixer par `X-`

Quelques champs de l'entête

- Entêtes normalisées facultatives
  - `Received` : une information sur le chemin suivi
  - `Reply-To` : une adresse pour la réponse
  - `Subject` : le sujet du message
  - `Message-ID`: un identifiant unique du message
- Entêtes privées (commençant par X-)
  - `X-Phone` : un numéro d'appel
  - `X-Mailer` : l'identifiant du logiciel de gestion de courrier (Mozilla Thunderbird …)

Le format MIME

- *Multipurpose Internet Mail Extensions*
- Apparition du format MIME (RFC 1341 et 1342 Juin 1992) pour combler les lacunes du format de base
- MIME introduit quelques nouvelles entêtes ayant surtout pour objet de décrire le format des corps de courriers
- Améliorations successives des spécifications de MIME : RFC 1521, 1522 … RFC 2045 à 2049 (Novembre 1996) RFC 2822 (Avril 2001)

Les objectifs de MIME

- A la base transmettre des messages textuels qui utilisent des jeux de caractères autres que l'US-ASCII (par exemple ISO-Latin)
- Permettre la définition d'un système très général de typage pour des documents multimédia (textes, images, sons, tableurs, …)
- Permettre de transmettre des corps de message comportant plusieurs parties (message avec plusieurs attachements)

Principaux types de données MIME

- Cinq types de données
  - **Type texte** : données lisibles
    - `text/rfc822 ; text/plain [RFC2646] ; text/html [RFC2854]`
  - **Type image** : différents codages image
    - `image/jpeg ; image/gif`
  - **Type son** : différents codages audio
    - `audio/basic` (MIC mu 8000 Hz 8 bits)
  - **Type vidéo** : images animées
    - `video/mpeg`
  - **Type application** : données restantes
    - `application/octet-stream ; application/PostScript`

Principaux types de données MIME

- Données **composites** ou **assemblées** (« multipart »)
  - plusieurs types de données combinés en un seul corps
- Cinq principaux types de syntaxe identique mais de sémantique différente
  - `multipart/mixed` : les données assemblées sont indépendantes
  - `multipart/alternative` : les données sont des alternatives d'une même information (au format texte et html par exemple)
  - `multipart/digest` : la forme par défaut `text/plain` est la forme textuelle la plus simple d'un message soit `text/rfc822` ➔ permet de transférer une suite de messages ou une boîte à lettres
  - `multipart/parallel` : les données sont présentées en parallèle
  - `multipart/related` : les données sont reliées (comme un document HTML qui comprendrait des images incluses)

Principaux types de données MIME

- Le **type message** est défini pour transporter dans un corps de courrier électronique un autre courrier électronique
  - permet d'encapsuler un courrier avec toutes ses informations d'entête, de corps dans un autre courrier
- Par exemple une erreur dans un courrier nécessitant le renvoi de ce courrier dans un courrier de diagnostic

Formats de codage MIME

- Pour transférer des données quelconques (des suites d'octets), MIME définit **cinq formats de codage**
  - texte 7 bits
  - Quoted-Printable
  - Base 64
  - 8 bits
  - binaire
- Existence d'autres formats mais non normalisés MIME : binhex (apple), uuencode, xxencode, (Unix) ...

Format texte 7 bits, US-ASCII

- Encodage par défaut si rien d'autre n'est spécifié
- Standard initial de la messagerie Internet (RFC 822)
- Chaque caractère est codé en US-ASCII 7 bits
- Jeu de caractère du Network Virtual Terminal Telnet
- Uniquement pour les textes non accentués

Format Quoted-Printable

- Codage d'un texte d'un alphabet de caractères 8 bits (ex ASCII ISO Latin) en US-ASCII 7 bits
- Les caractères standards (code 33 à 127 sauf le 61 caractère =) sont codés en US-ASCII 7 bits
- Les caractères spéciaux 8 bits (é, è, ç , à ...) sont codés par une séquence `=NM` où N et M représente les deux chiffres hexadécimaux des 8 bits du code ASCII à représenter (espace ⇨code ASCII 32 ⇨ `=20`)
- Utilisation de différents alphabets nationaux possible à condition de définir lequel est utilisé
- A utiliser s'il y a peu de différences avec l'US ASCII

Format Base 64

- Coder tout type de données 8 bits avec des caractères US-ASCII
- Découpage en groupes de 3 octets ; les groupes de 24 bits obtenus sont codés par 4 caractères US-ASCII, un caractère codant 6 bits
- On choisit dans l'US-ASCII 64 symboles différents
  - ABCDEFGHIJKLMNOPQRSTUVWXYZabcdefghijkl mnopqrstuvwxyz123456789+/
- Le codage Base 64 augmente la taille du message d'environ 30%

Format caractères 8 bits

- Pour transporter des données en format caractères 8 bits dans les corps de messages sans les transcoder en US-ASCII
- Définir dans l'entête l'alphabet utilisé
  - Type de données (*content-type*) et jeu de caractères utilisés (*charset*)
- Nécessite des modifications au format standard des courriers
- Ces modifications sont définies dans le cadre du protocole ESMTP (option 8BITSMIME)

Format binaire

- Possible en MIME mais pose des problèmes avec le format standard (utilisation du type `application/octet-stream`)
- Problèmes
  - Longueur des lignes (RFC 822, une ligne doit < 1000 caractères)
  - Déterminer la fin du message par un délimiteur approprié
- Nécessite d'utiliser les extensions ESMTP (Binary)

Codage des champs d'entête

- Existence de données d'autres alphabets que l'US ASCII dans les entêtes
- Règle d'encodage
  `=?charset?encodage?valeur?=`
  - charset : le jeu de caractère utilisé,
  - encodage : Q pour Quoted-Printable et B pour Base 64
  - valeur : résultat de l'encodage
- Exemple

  `Subject: Bonjour =?iso-8859-1?Q?G=E9g=E9?=`

Directives d'entête spécifiques de MIME

- MIME utilise des directives d'entête décrivant le corps d'un message pour permettre son interprétation à l'arrivée
- Champs spécifiques de MIME
  - `Mime-Version` : la version utilisée actuellement 1.0
  - `Content-Type` : le type et les sous-type des données
  - `Charset` : le jeu de caractères utilisé
  - `Content-Tranfer-Encoding` : l'encodage utilisé (Quoted-Printable, Base 64)
  - `Content-ID` : Identificateur unique de partie de message
  - `Content-Description` : Informations complémentaires sur le contenu

Exemples de format MIME

- Courrier en français encodé en format base 64

```
MIME-Version: 1.0
Content-Type: text/plain; charset=ISO-8859-1
Content-Transfer-Encoding: base64
```

- Commentaires :
  - Les données transportées sont des caractères ISO-8859-1 (latin)
  - Le codage du corps est effectué en Base 64

Exemples de format MIME

- Courrier composite multipart

```
MIME-Version: 1.0
To: gerard.mensoif@orange.fr
Subject: Bonjour =?ISO-8859-1?Q?G=E9g=E9?=
Content-Type: multipart/alternative;
  boundary="------------000703010400050308090303"
This is a multi-part message in MIME format.
--------------000703010400050308090303
Content-Type: text/plain; charset=ISO-8859-1; format=flowed
Content-Transfer-Encoding: 8bit
Bonjour Gégé
--------------000703010400050308090303
Content-Type: text/html; charset=ISO-8859-1
Content-Transfer-Encoding: 7bit
<html>
  <head>
    <meta http-equiv="content-type" content="text/html; charset=ISO-
8859-1">
  </head>
  <body text="#000000" bgcolor="#FFFFFF">
    <font color="#cc0000">Bonjour G&eacute;g&eacute;</font>
  </body>
</html>
--------------000703010400050308090303
```

Les protocoles de relève de courrier

- Protocoles dérivés du protocole SMTP
- Offrent des fonctions spécifiques de relève du courrier dans une boîte à lettres
  - Fonctions de transfert de courrier d'un serveur de messagerie vers un client de messagerie
  - Fonctions de gestion des archives de courrier (liste de messages en attente dans une boîte, destruction de message …)

POP

- *Post Office Protocol* défini dans la RFC 1939
- Protocole de relève le plus simple
- Version 3 du protocole, utilise le port TCP 110
- Les messages une fois transférés sont généralement effacés du serveur
  - Une copie peut être conservée pendant un laps de temps sur le serveur
- Ne gère pas les archives de courrier sur le serveur
- Convient bien à l'utilisation à partir du même poste client de messagerie
- Utilisation en mode sécurisé POPS à privilégier

IMAP

- *Internet Message Access Protocol*
- Protocole le plus complet défini dans la RFC 2060
- Version 4 du protocole, utilise le port TCP 143
- Gère le courrier directement sur le serveur
  - Création de dossier, déplacement des messages entre les dossiers …
- Minimise les échanges de données sur le réseau
- Un protocole adapté à la consultation à partir de différents poste clients
- Utilisation en mode sécurisé IMAPS à privilégier

MTA open source

- Sendmail
  - Existe depuis 1980
  - Auteur principal Eric Allman
- Postfix
  - Existe depuis 2001
  - Auteur principal Vietse Venema
- Exim
  - Existe depuis 1995
  - Auteur Principal Philippe Hazel
- Qmail
  - Existe depuis 1997
  - Auteur Dan Bernstein
- Postfix et Qmail sont considérés comme les meilleurs

MTA propriétaires

- Logiciels de messagerie d'entreprise le plus souvent intégrés dans des suites bureautiques ou serveurs Web
  - Exchange/Internet Information Service (MTA commun Microsoft à la messagerie Exchange et au serveur Web IIS)
  - Zimbra (VMware)
  - Lotus Notes/Domino (IBM)
  - IMAIL (Ipswitch)
  - …

Agents de délivrance de messages (MDA)

- Gestion de boîte à lettres (stockage), filtrage des messages, envoi de message de réponse automatique
  - procmail (logiciel libre le plus répandu), deliver, mailfilter, maildrop …
  - MDA aussi intégrés aux grands logiciels de messagerie intégrés (Exim, Exchange)
- Peuvent aussi incorporer des outils de protection contre les virus et le SPAM (SpamAssassin, Amavis, MIMEDefang …)

Serveurs d'accès aux messages (MAA)

- Support POP et/ou IMAP
  - qpopper (POP)
  - courier IMAP (POP et IMAP)
  - Dovecot
  - UW-POP et UW-IMAP (université de Washington)
  - cyrus IMAP
  - …

Service d’annuaire

Les services d'annuaire

- Stockage hiérarchique d'informations
- Permet de modéliser des objets
  - Utilisateurs (people)
  - Machines (computers)
  - Groupes (groups)
  - Unité Organisationnelle (OU)
  - …
- Et des attributs associés aux objets
  - Texte, données binaires, listes…
- Gestion de droits d'accès à l'annuaire (ACL)
- Sécurisé
  - Utilisation de TLS au niveau transport
  - Authentification

Pourquoi LDAP ?

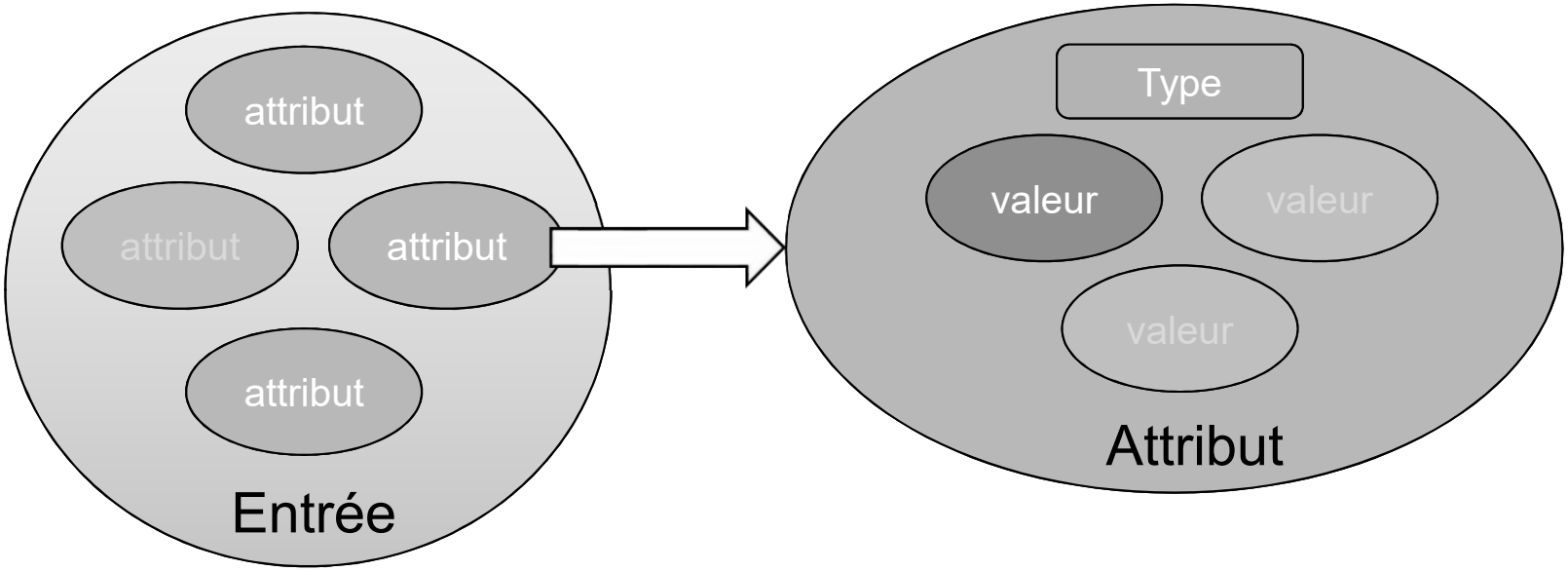
- Besoin d'uniformisation
  - Avant, Un annuaire ⇔ Un protocole d'accès
    - NIS / YellowPages
    - Microsoft SAM (*Security Account Manager*)
    - X.500 DAP jugé trop lourd pour l'implantation
  - Uniformisation protocolaire : LDAP (*Lightweight Directory Access Protocol*)
    - Au dessus de TCP/IP
    - Communique avec tout service d'annuaire
    - Pas d'impact sur l'implantation de l'annuaire
      - Standardisé en 1993 – RFC 1487
      - Révisé en 1995 – RFC 1777
      - LDAPv3 en 2002, révisé en 2006 (dernière version)

Pourquoi LDAP ?

- Pour l'administrateur système
  - Permet de stocker les informations des utilisateurs (login, mot de passe, nom, horaires, home directory, userid…)
  - Authentification uniforme sur le réseau
- Pour le développeur d'applications
  - Authentifier un utilisateur
  - Lire les informations sur une entrée
  - (Eventuellement) associer des informations sur une entrée

LDAP

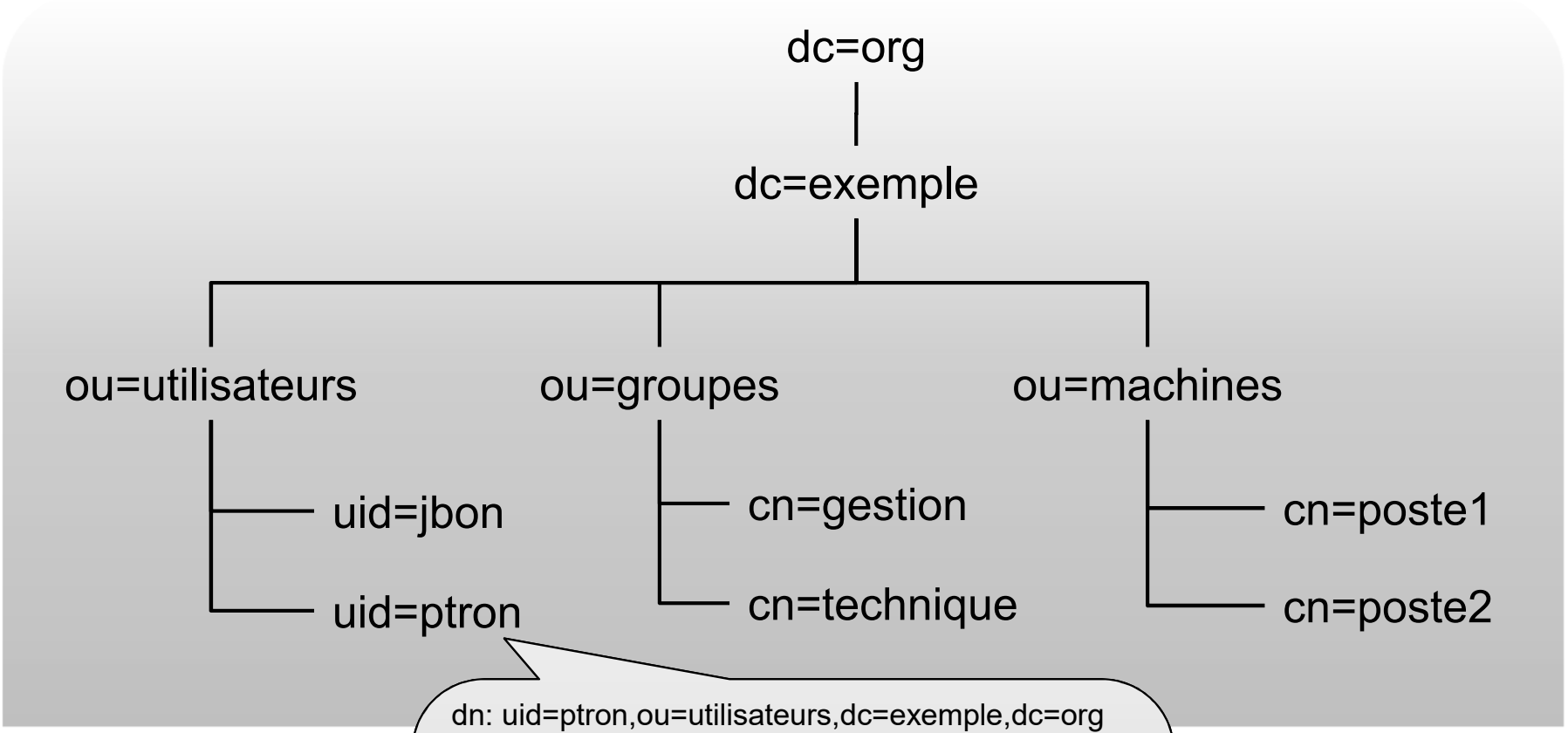
- Un annuaire LDAP est un arbre d'**entrées**
  - l'arbre reflète le modèle organisationnel, politique ou géographique de la structure représentée
- Une entrée est constituée d'un ensemble d'**attributs**
  - Un attribut possède un **nom**, un **type** et **une ou plusieurs valeurs**



LDAP

- Les attributs sont définis dans des **schémas**
  - Caractère multivalué = différence majeure avec les SGBD
  - Un attribut qui n'a pas de valeur est absent de l'entrée
- Chaque entrée a un identifiant unique, le *Distinguished Name* (DN)

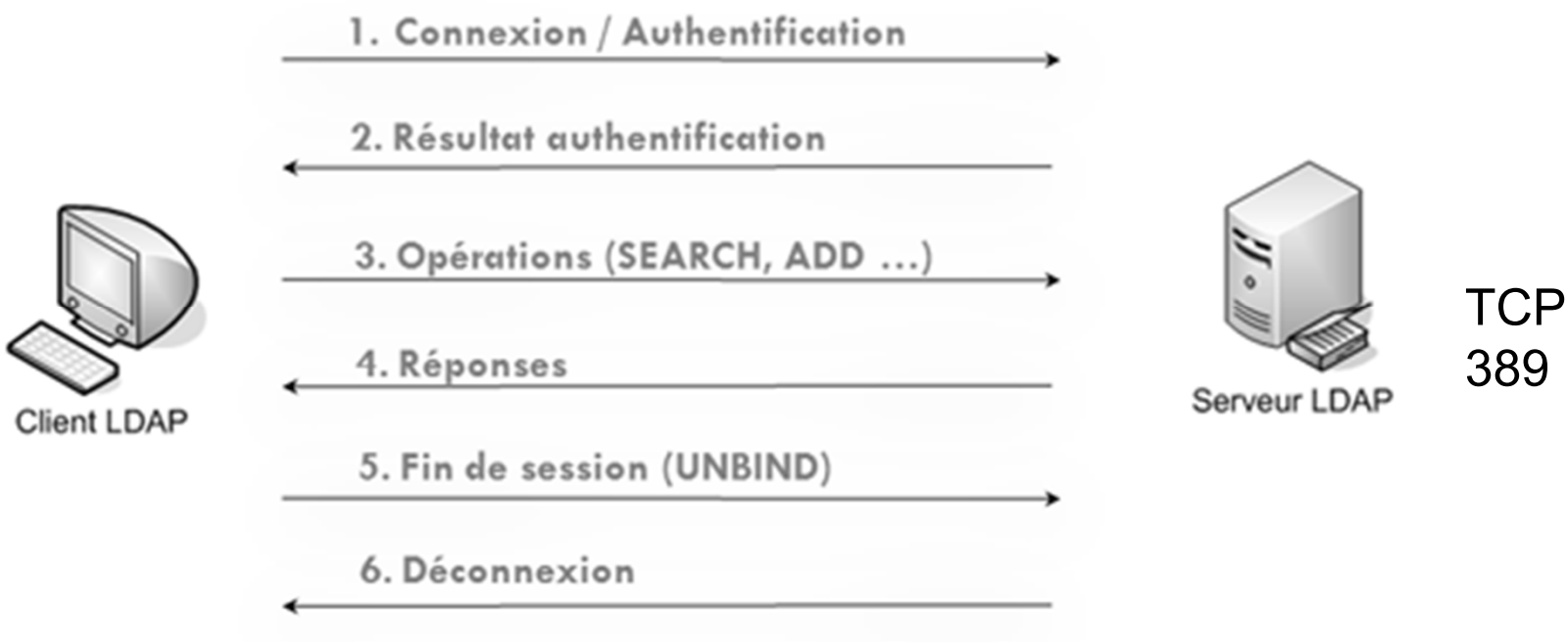
LDAP




```
dn: uid=ptron,ou=utilisateurs,dc=exemple,dc=org
cn: Paul Tron
givenName: Paul
sn: Tron
uid: ptron
telephoneNumber: +33 1 23 45 67 89
mail: paul.tron@exemple.org
objectClass: inetOrgPerson
objectClass: organizationalPerson
objectClass: person
objectClass: top
```

Communication LDAP

- Connexion type client/serveur permettant à une application de se connecter / authentifier auprès d'un serveur LDAP et d'effectuer des opérations de recherche, d'ajout ou de modification de données



Communication LDAP

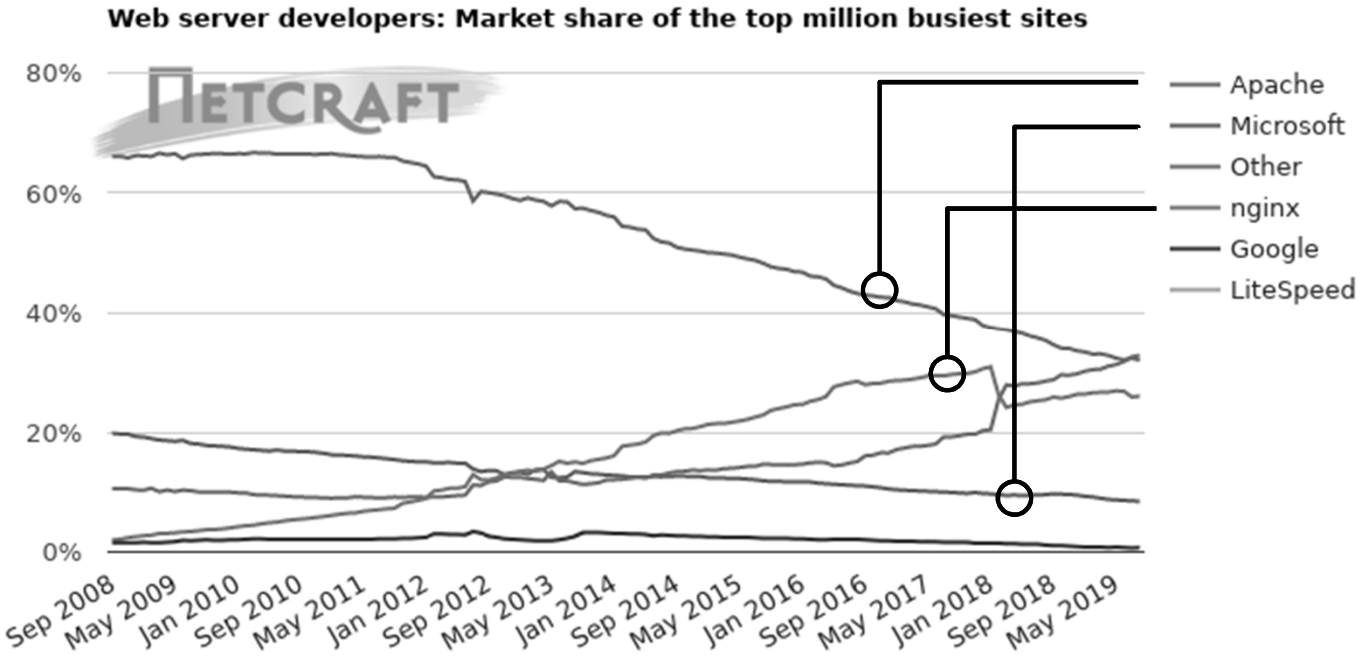
- Autre mode de communication prévu
  - opérations d'import ou d'export
  - communication entre annuaires
- Basé sur le format d'échange LDIF (*LDAP Data Interchange Format*)
  - format d'échange sous forme de fichier ASCII
  - permet de décrire les données d'un annuaire, son schéma ainsi que des opérations (ajout, suppression, ...)
  - particulièrement adapté à l'exportation d'un annuaire dans le but de le sauvegarder ou de l'importer sur un autre serveur (replication)

```
dn: uid=ptron,ou=utilisateurs,dc=exemple,dc=org
changetype: modify
add: mail
mail: ptron@exemple.org
```

Autres services : HTTP, NTP

HTTP (*HyperText Transfert Protocol*)

- Serveur HTTP ou serveur Web
- Sert les requêtes HTTP en TCP sur le port 80, et/ou HTTPS sur le port 443
- Le logiciel libre Apache (fondation Apache) reste encore populaire, même si Nginx connaît une progression importante



NGINX

source : Netcraft

HTTP (*HyperText Transfert Protocol*)

- Apache
  - Apparu en avril 1995, actuellement dans sa version 2
  - Conçu pour prendre en charge de nombreux modules (Perl, PHP, réécriture d'URL …)
  - Possibilité de gérer des sites virtuels
    - www.site1.domain et www.site2.domain par exemple

HTTP (*HyperText Transfert Protocol*)

- Nginx
  - Se prononce « engine-ex »
  - Est distribué sous licence BSD 2
  - Initialement conçu comme réponse au problème de performances liés au traitement de 10 000 connexions simultanées
    - Utilise une architecture asynchrone et événementielle
  - Utilisé comme serveur Web mais aussi comme proxy inverse et équilibreur de charge pour HTTP, TCP et UDP

NTP (*Network Time Protocol*)

- *Network Time Protocol*
- Protocole permettant de synchroniser, via un réseau informatique, l'horloge locale d'ordinateurs sur une référence d'heure (UTC)
  - Dérive des horloge à quartz dans le temps
  - Synchronisation indispensable pour le bon fonctionnement de certains protocoles réseau
- Protocole assez ancien (version 0 en 1985), la version la plus répandue à ce jour reste la 3 (spécifiée dans la RFC 1305)
  - La version 4 de NTP est une révision importante publiée dans la RFC 5905 en juin 2010
- S’appuie sur UDP et utilise le port 123

Architecture NTP

- Basée sur une **structure hiérarchique** composée de
  - récepteurs récupérant l'heure de référence par radios, câbles, satellites ou directement depuis une horloge atomique
  - de serveurs de temps récupérant l'heure de référence auprès des récepteurs ou bien auprès d'autres serveurs de temps
  - de clients récupérant l'heure de référence auprès des serveurs de temps
- Chaque couche ou niveau est appelé une **strate**
- Chaque client NTP est également un serveur et se synchronise avec d'autres serveurs, le plus souvent de la strate supérieure
  - La strate 0 comprend des horloges de référence
  - Jusqu'à 16 strates sont prévues dans la norme, mais la plupart des clients se situent dans les strates 3 ou 4
- La redondance des serveurs et leur organisation permet une répartition de la charge

Architecture NTP