Cours annuaires distribués M3104

Jean-marc Pouchoulon Octobre 2014

Définition d'un annuaire

- C'est une collection d'informations. Ces informations sont organisées pour un processus de lecture et pas d'écriture.
- Typiquement on va stocker dans un annuaire des comptes relatifs à des personnes ou à des machines.
- Exemple d'annuaires : Annuaire téléphonique, un fichier .csv avec ses colonnes…

Annuaires électronique

- Même principe que les annuaires papiers, mais avec les avantages du numérique :
  - Puissants : recherches multi-critères complexes et dynamiques : mises à jour plus faciles.
  - Souples : possibilité d'évolution de la structure des données.
  - Sûrs : authentification, contrôles d'accès.
  - Personnalisables : affichage en fonction de l'utilisateur.

Source: http://cesar.resinfo.org/IMG/pdf/formation_ldap_hybride.pdf

Différents types d'annuaires

- Annuaires orientés sécurité (contenant les droits d'accès à une application pour un utilisateur).
- Annuaires orientés « informations » (fiches signalétiques: nom tel bureau, machines…).
- Annuaires techniques (stockage de routes, de scores anti-spam, d'entrées DNS…)

Origine de LDAP:

- Le standard X500 pour la gestion d'annuaires téléphoniques à l'échelle d'un pays.
- Les annuaires suivant les spécifications X500 étaient accessibles par le protocole DAP (**D**IRECTORY **A**CCESS **P**ROTOCOL). X500 a amené la notion de DIT (**D**IRECTORY **T**REE **I**NFORMATION) unique.
- X500 s'est révélé trop coûteux à implémenter. Vinrent ensuite les serveurs LDAP (**L**IGHT **D**IRECTORY **A**CCESS **P**ROTOCOL) qui ont répondus à une partie des spécifications.
- Le protocole est en version 3 actuellement.

Cas d'emplois des annuaires

- Un annuaire est fait afin de supporter un grand nombre d'entrée. Dans le métier on considère qu'un annuaire de taille intéressante est composé d'un million d'entrées à minima. Certains préfèrent une base de données mais les performances d'un annuaire sont supérieures en lecture.

- Un annuaire est fait pour avoir un ratio lecture/écriture élevé. On écrit peu dans un annuaire , en général en batch la nuit afin de mettre à jour de façon massive, ou au cours de la journée (ex: Modification de mots de passes)

Annuaire versus Bases de données

LDAP est

- Un protocole standardisé.
- Proche du fonctionnement de l'entreprise (son schéma peut être celui de l'organigramme)
- Hiérarchique.
- Très performant en lecture.
- Simple à utiliser
- Un annuaire ne s'interroge pas en SQL. Il n'y a pas de jointure comme avec SQL. Il n'est pas dédié à une application mais à toutes.
- Les schémas (structures) sont prêt à l'emploi ou peuvent être facilement aménagé.

Qui l'utilise ?

Tout le monde!!! C'est un élément essentiel des architectures systèmes et réseaux.

On se sert de LDAP pour:

- Fédérer l'authentification des applications (Samba , Web serveurs…)
- Décrire les droits de chaque utilisateur sur des applications.
- Stocker des informations relatives à la messagerie. (routes SMTP, alias, score antispam …)
- Servir de backend au DNS pour stocker les zones.
- Et comme annuaire bien évidemment afin de stocker des informations utilisateurs.

Quels sont les principaux annuaires utilisés:

Opensource

- OpenLdap (OpenSource le plus utilisé)
- 389 Directory Server (RedHat)

Commerciaux:

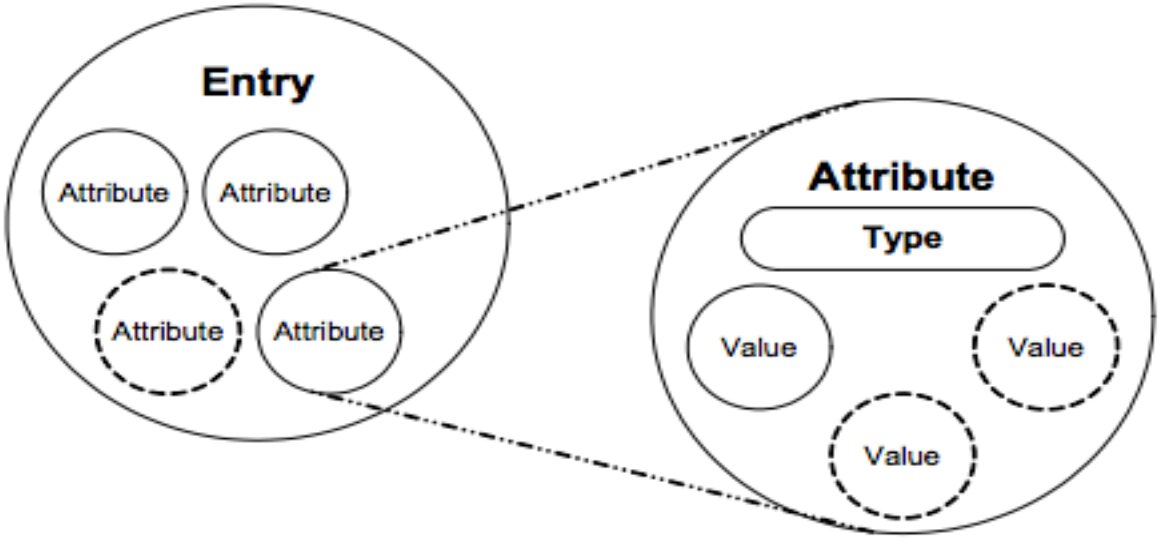
- Active Directory
- Oracle Directory Server (ex sun)

Les challengers: Open DJ, Apache Directory Server…

Vue d'ensemble de LDAP

- Quatre modèles prédéfinis:
  - Le modèle d'information (nature des données).
  - Le modèle de désignation (structure hiérarchique).
  - Le modèle des services (fonctions disponibles).
  - Le modèle de sécurité (droits d'accès).
- Des classes d'objets et des attributs normalisés.
- Des fonctions de recherche évoluées.
- Répartition des données sur plusieurs référentiels , accessibles de façon transparente pour le client LDAP (appel récursif ou itératif)
- TCP/IP of course.

Le modèle d'informations



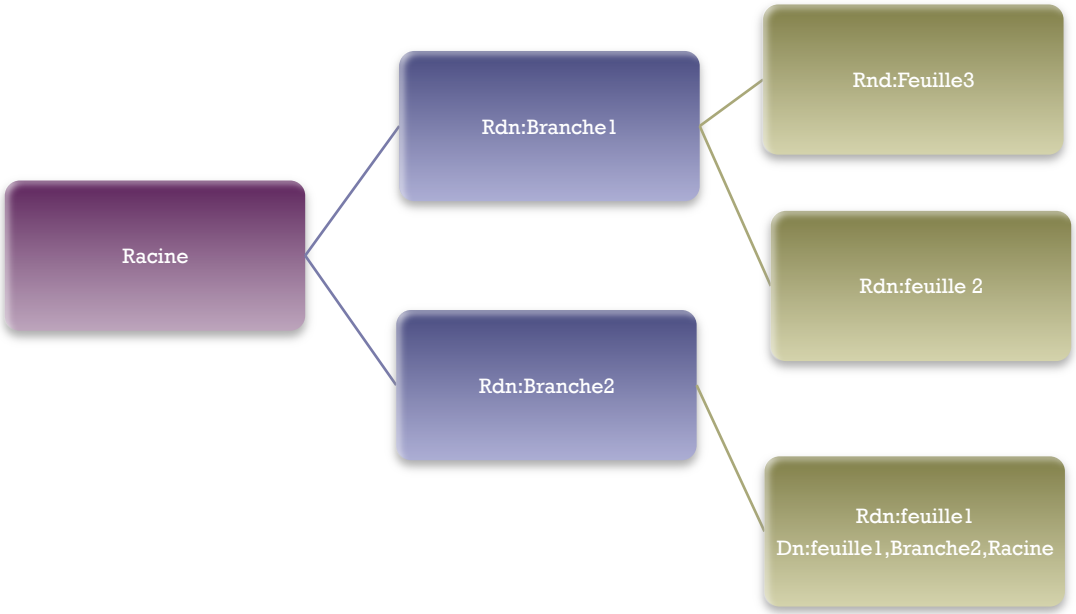
Source: IBM RedBook

Une entrée est l'unité de base d'un annuaire , elle a des attributs (obligatoires ou non), remplis par des Valeurs.

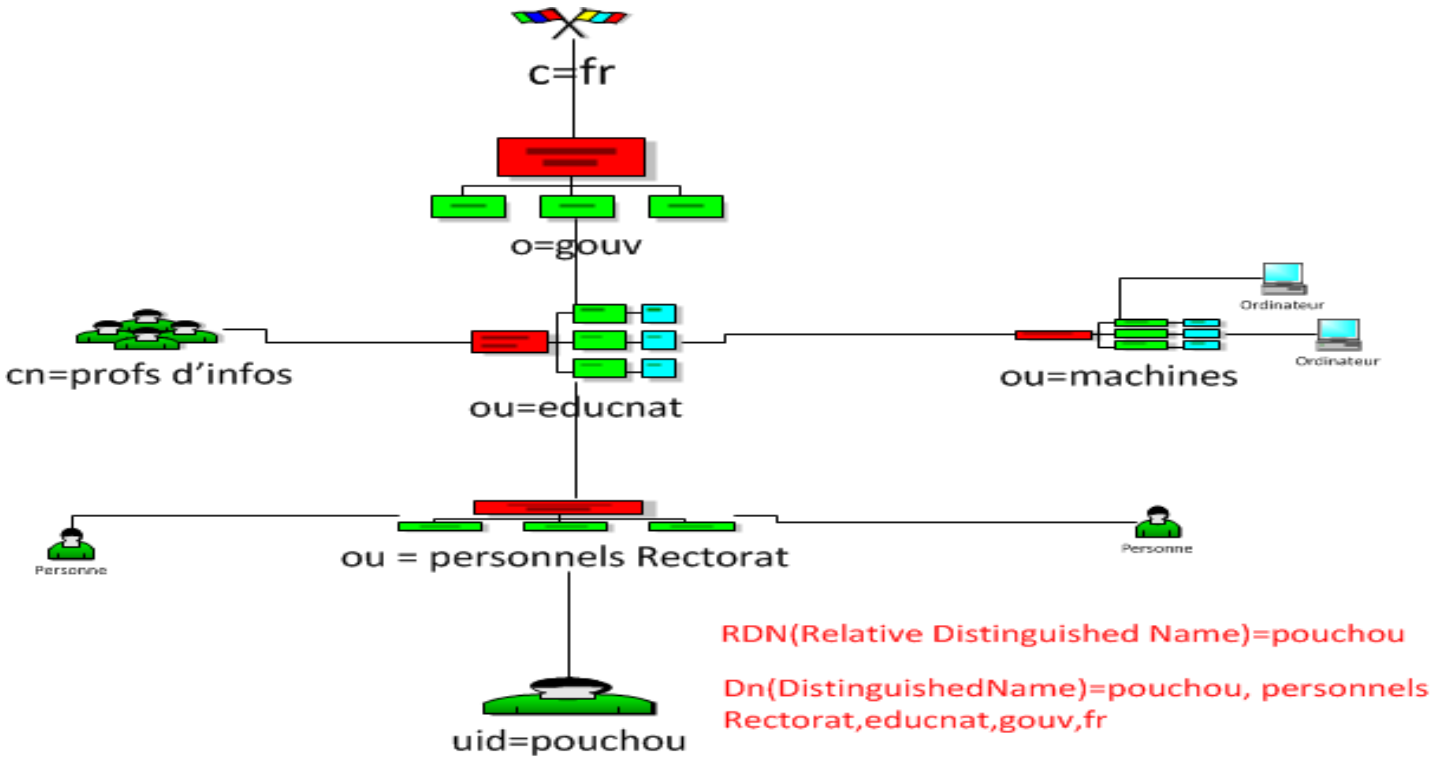
Attribut commonName=cn qui est cis (Case Information Sensitive) et qui a pour valeur jean-marc Pouchoulon

Arbre LDAP (Modèle de désignation)

- Racine
- RDN (Relative Distinguiched Name) Clef d'entité dans une branche. (exemple Branche1)
- DN (Distinguiched Name). Un DN caractérise une entrée de façon unique dans l'annuaire. Il est composé des RDN de l'entité vers la racine (exemple dnRacine,Branche1,Feuille3)



Arbre LDAP ou D.I.T.



Le D.I.T. (Directory Information Tree) est la représentation des entrées et des relations entre elles.

LDAP // avec la prog orientée objet:

LDAP a été conçu avec le modèle objet en tête.

Un objet est une entité qui va être constitué de variables (nom, prénom, âge…) et de méthodes (getNom,getPrenom, setNom,setPrenom)…

De même un annuaire va avoir des ObjectClass qui sont en liens avec des attributs. Les entrées peuvent agréger des objets entre eux et des attributs peuvent être partagés entre plusieurs objets. Un attribut ou une classe peuvent hériter d'autres attributs ou classes.

Conception d'un annuaire

- **C'est l'étape importante du processus de mise en œuvre de l'annuaire.**
- **Quelles informations mettre dans l'annuaire ?(définition des attributs et des classes d'objets).**
- **Quelle est la provenance des données et quels en sont les propriétaires ? (bases de données , fichiers…)**
- **Comment les organiser dans un modèle commun à toutes les applications? (schéma de l'annuaire, organisation hiérarchique).**

Voir http://cesar.resinfo.org/IMG/pdf/formation_ldap_hybride.pdf

Le schéma lien entre Object Class et attributs.

Le schéma définit les relations entre les classes d'objet et les attributs.

Depuis la version 3 du protocole LDAP il est possible de le stocker dans l'annuaire LDAP lui-même ce qui permet de le consulter et de le modifier.(cn=subschema) Il est constitué de l'ensemble :

- des attributs.
- de leurs syntaxes.
- des règles de comparaison.
- des classes d'objets.

Il permet de garantir la validité et l'intégrité des données (la contrainte de validation des données par le schéma est rude mais elle est indispensable)

OID

Un OID est un identifiant unique associé à :

- chaque classe d'objet.
- chaque type d'attribut.

Il est composé de plusieurs numéros séparés par un point.

Chaque numéro représente une branche du DIT.

Tous les attributs du standard commencent par 2.5.4.

Toutes les classes d'objet commencent par 2.5.6.

On a aussi un OID numérique qui correspond aux types de donnée (entier , texte , binaire)

Tout est OID dans LDAP …

OID numérique décrivant le type de l'attribut.

OID	Nom	Description
1.3.6.1.4.1.1466.115.121.1.7	Boolean	Valeur logique
1.3.6.1.4.1.1466.115.121.1.15	DirectoryString	Chaîne de caractères UTF-8
1.3.6.1.4.1.1466.115.121.1.12	DN	Nom complet
1.3.6.1.4.1.1466.115.121.1.26	IA5String	Chaîne de caractères ASCII
1.3.6.1.4.1.1466.115.121.1.27	INTEGER	Entier
1.3.6.1.4.1.1466.115.121.1.28	JPEG	Image au format JPEG
1.3.6.1.4.1.1466.115.121.1.54	LDAP SyntaxDescription	Définition d'une syntaxe LDAP
1.3.6.1.4.1.1466.115.121.1.50	TelephoneNumber	Numéro de téléphone

OID de comparaison

OID	Nom	Description
2.5.13.0	objectIdentifierMatch	Égalité entre deux OID
2.5.13.1	distinguishedNameMatch	Égalité entre deux DN
2.5.13.2	caseIgnoreMatch	Égalité entre deux chaînes, insensible à la casse
2.5.13.8	numericStringMatch	Idem caseIgnoreMatch, insensible aux espaces
2.5.13.3	caseIgnoreOrderingMatch	Comparaison de deux chaînes, insensible à la casse
2.5.13.4	caseIgnoreSubstringsMatch	Inclusion, insensible à la casse
2.5.13.10	numericStringSubstringsMatch	Idem caseIgnoreSubstrings Match, insensible aux espaces

Définition d'un attribut

- Chaque attribut est défini par :

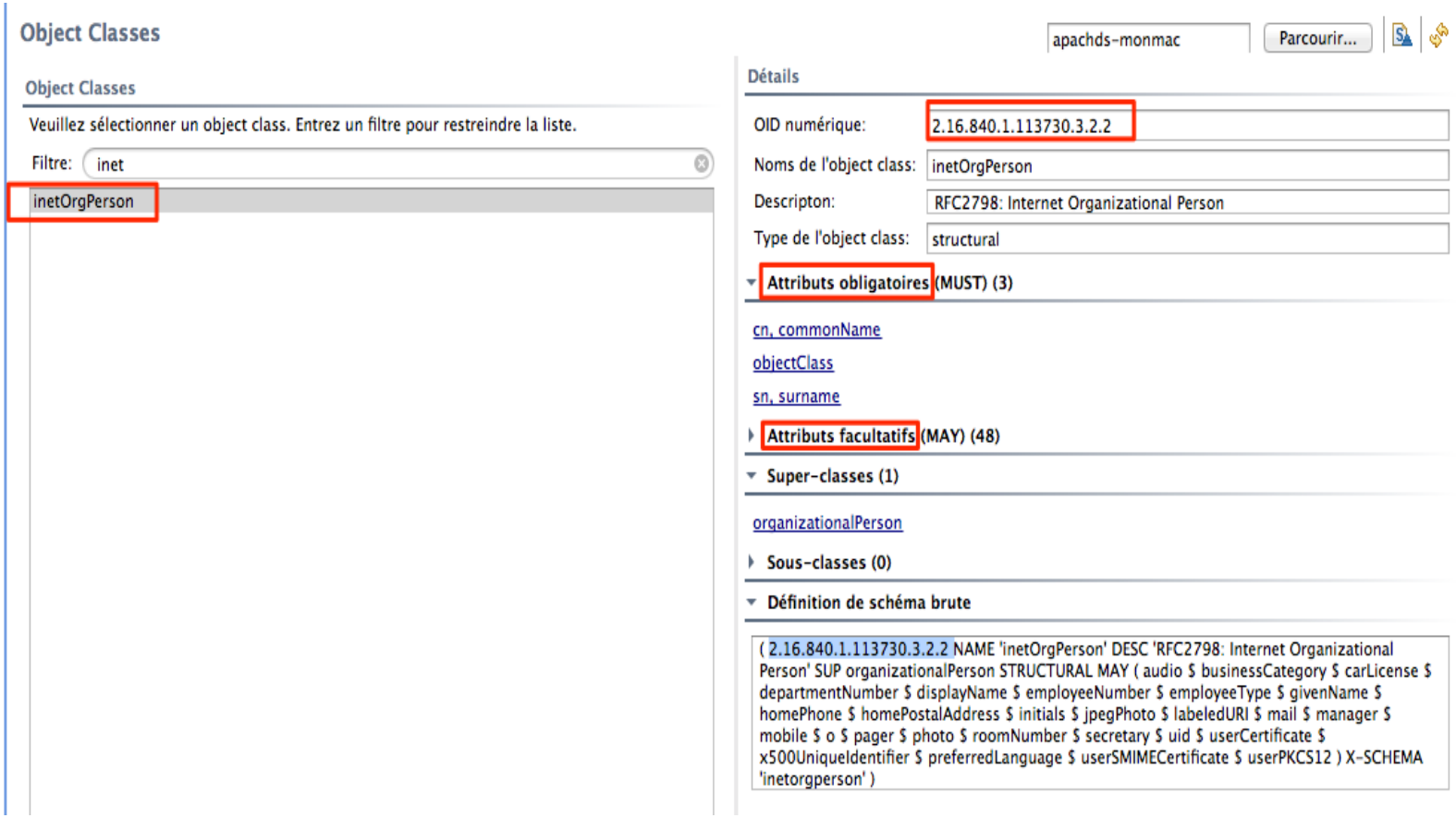
- une description (DESC).
- son nom (NAME).
- les règles d'égalité (EQUALITY).
- une valeur numérique, appelée OID (Object Identifier).
- ce que peut contenir l'attribut (SYNTAX).
- les règles d'égalité de sous-chaînes (SUBSTR).

Définition d'une classe d'objets

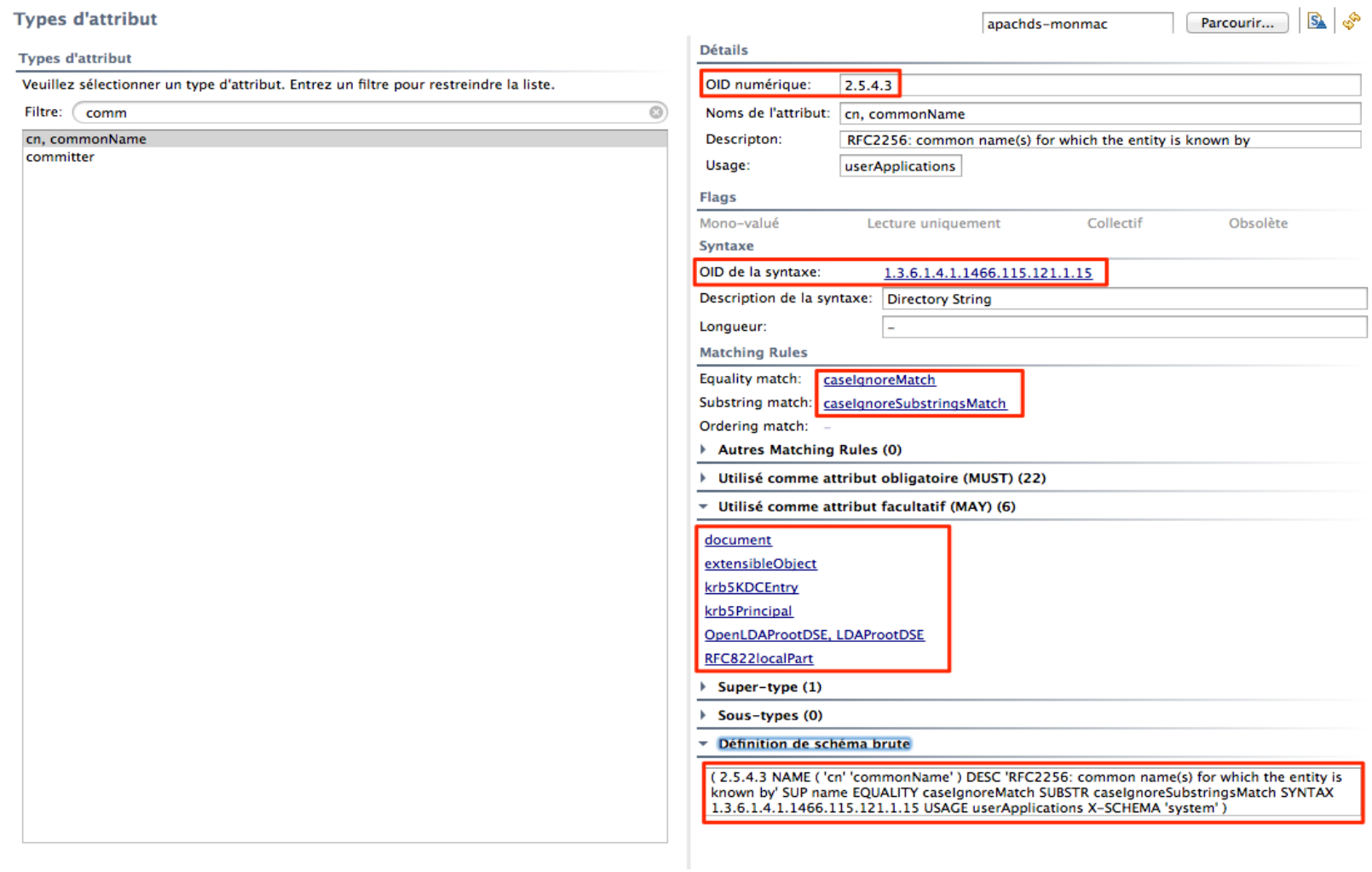
Une classe d'objets est définie par :

- Son OID.
- Son nom.
- Une courte description de la classe.
- La classe dont elle dérive.
- Son type (ABSTRACT, STRUCTURAL, AUXILIARY).
- la liste des attributs obligatoires (MUST).
- la liste des attributs facultatifs (MAY).

Object Class



Attribut cn de l'object InetOrgPersonn



Définition d'un attribut

- Son OID.
- Son nom.
- Une courte description
- Les critères de comparaison utilisés lors d'une Recherche.
- Une syntaxe décrivant le type de données.
- Un attribut peut être multi-valué. Par exemple il peut y avoir plusieurs adresses de mail stockée dans l'attribut mail.

Quelques attributs courants

RFC 2256

- uid User id
- cn Common Name
- sn Surname
- l Location
- ou Organisational Unit
- o Organisation
- dc Domain Component
- st State
- c Country

Règles de comparaison des attributs

- caseIgnoreMatch : ignorer la casse lors de la comparaison de deux chaînes de caractères.
- caseExactMatch : tenir compte de la casse.
- telephoneNumberMatch : ignorer la casse et supprimer les espaces, virgules, points, etc.
- integerMatch : comparer deux entiers.
- booleanMatch : comparer deux attributs booléens.
- distinguishedNameMatch : comparer des DN.
- octetStringMatch : comparer des binaires octet par octet.

Deux types d'attributs :

- Les attributs utilisateurs :

mail, cn, telephoneNumber, uid.

- Les attributs opérationnels :ils sont liés au fonctionnement de l'annuaire et ne sont pas accessibles aux utilisateurs (exemple : modifytimestamp).

Entrée RootDSE

Lorsqu'un client se connecte à un annuaire, il n'a pas d'informations sur les caractéristiques de cet annuaire, la façon dont les connexions sont sécurisées, ou le schéma utilisé. Un serveur LDAP v3 publie une entrée particulière, appelée RootDSE, pour fournir ce type d'information.

Cette entrée RootDSE a comme caractéristique d'avoir un DN vide. Elle contient des attributs opérationnels, qui indiquent au client les caractéristiques que sait gérer le serveur, et un attribut de type subschemaSubentry dont la valeur est cn=Subschema.
Pour l'obtenir: ldapsearch -x -s base -b "" "(objectclass=*)"

Le format LDIF

- LDIF = LDAP Directory Interchange Format

C'est un format de fichier standardisé et indépendant de l'annuaire.

Il permet de dumper un annuaire existant à des fins de sauvegardes, pour initialiser un réplicat ou un autre maitre.

Il permet aussi de modifier des entrées existantes

LDIF de données

```
dn: dc=mathrice,dc=prive
dc: mathrice
objectClass: dcObject
objectClass: organization
o: Mathrice

dn: ou=people,dc=mathrice,dc=prive
ou: people
objectClass: organizationalUnit

dn: ou=groups,dc=mathrice,dc=prive
ou: groups
objectClass: organizationalUnit

dn: uid=anomusu,ou=people,dc=mathrice,dc=prive
objectClass: inetOrgPerson
objectClass: posixAccount
objectClass: shadowAccount
cn: Agtpre.Nomusu
sn: Agthom
givenName: Agtpre
uid: anomusu
uidNumber: 1001
gidNumber: 300
homeDirectory: /home/anomusu
loginShell: /bin/bash
shadowExpire: 0
userPassword: {SSHA}6gHTLCHTwZjOv/HjT//pTq94ba1HyQq1
mail: Agtpre.Nomusu@domain.tld
mail: anomusu
```

LDIF de modification

```
dn: uid=Pouchou,ou=people,dc=mathrice,dc=prive
changetype: add
loginShell: /bin/bash
initials:: QS7CrOKAoEYu
sn: Fonction
gidNumber: 300
objectClass: organizationalPerson
objectClass: person
objectClass: posixAccount
objectClass: shadowAccount
objectClass: inetOrgPerson
objectClass: top
mail: Arcive.Fonction@domain.tld
mail: jean-marc.pouchoulon@iutbeziers.fr
givenName: Arcive
uid: Pouchou
uidNumber: 1017
shadowExpire: 0
cn: Arcive.Fonction
homeDirectory: /home/pouchou
userPassword:: e3NzaGF9SHdBbzlsNmtvNllrVGtuMU15SEhMWTBVMUlDS3pBS0cxdjBLZUE9PQ=
 =

#!RESULT OK
#!CONNECTION ldap://localhost:10389
#!DATE 2014-09-20T14:1 :53.614
dn: uid=Pouchou,ou=people,dc=mathrice,dc=prive
changetype: modify
replace: sn
sn: ChefOuiChef
```

type de modification

Attribut à modifier

Les modifications LDIF (Modèle Fonctionnel).

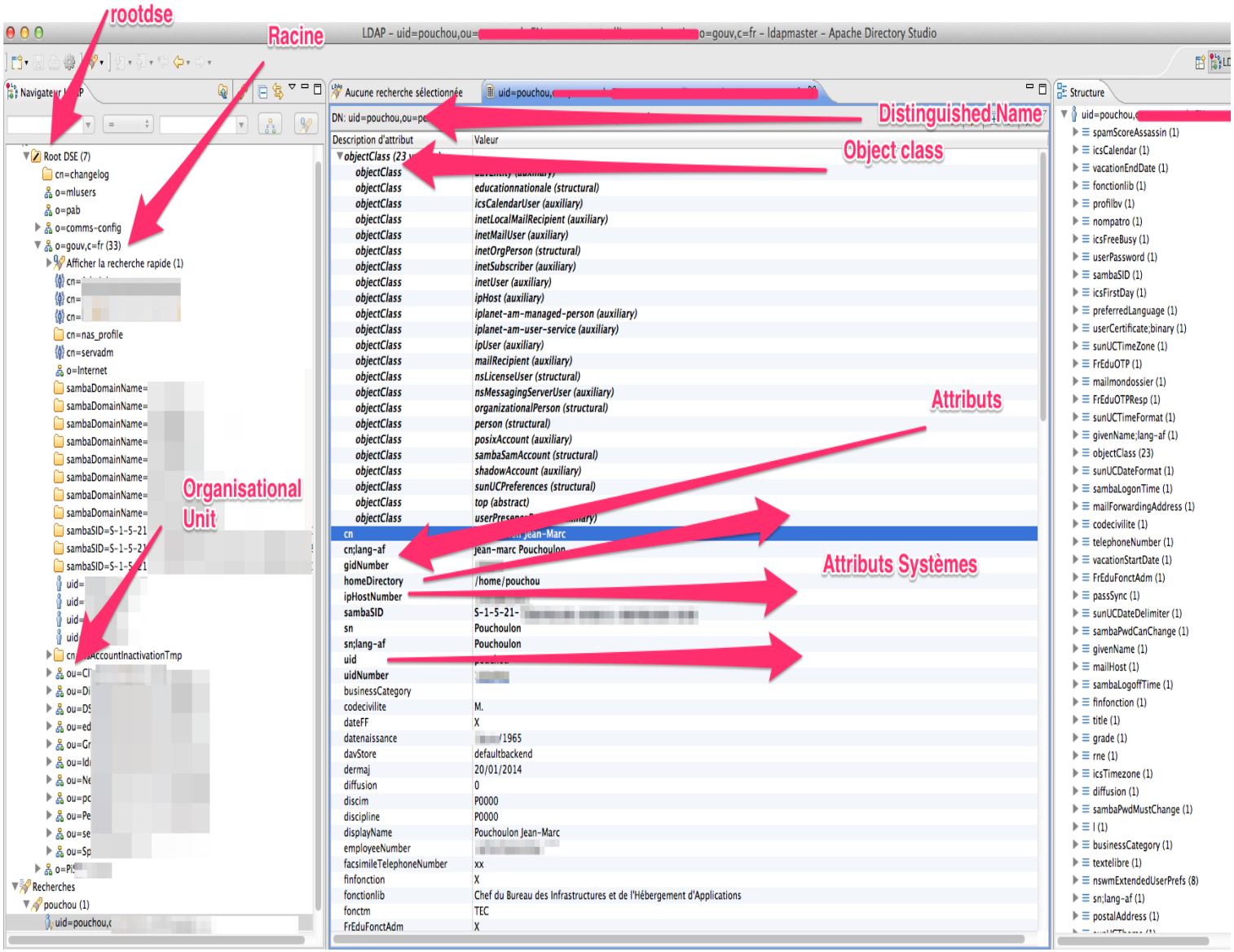
Les opérations possibles sont :

- add : ajout d'une entrée
- delete : suppression d'une entrée
- modify : modification d'une entrée
- modrdn : modification d'un nom relatif
- moddn : modification d'un nom dn.

SuffixeLDAP

- Le choix du suffixe n'est pas contrôlé et est libre. Cela peut poser problème lors de l'intégration de l'entité dans une autre entité. En général la forme est de type o=entité c=pays (o=gouv, c-fr) ou (dc=ent ,dc=fr)

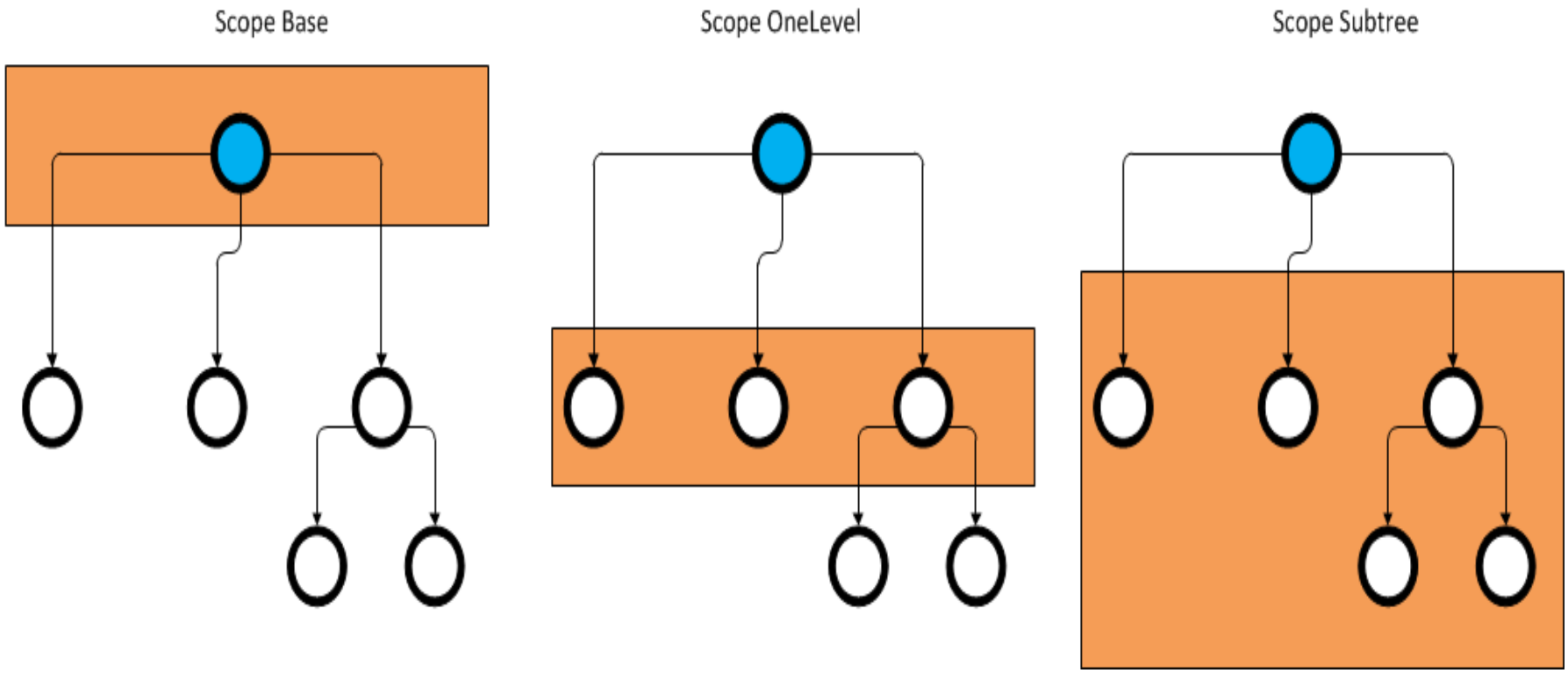
Les éléments d'un annuaire



Recherche LDAP:

- **BASE** est par exemple o=gouv,c=fr ; c'est l'endroit à partir duquel on effectue la recherche ;
- **ATTRIBUT** peut être vide pour indiquer que l'on affiche tous les attributs de chaque entrée, ou bien par exemple peut contenir uid,mail si l'on ne veut afficher que les attributs uid et mail.
- SCOPE est généralement égal à *sub* si l'on veut pouvoir effectuer une recherche complète dans le sous-arbre ; les autres valeurs possibles sont *one* et *base*. « base » permet de ne retourner que l'objet situé au point de recherche, et « one » retourne tous les objets situés immédiatement au niveau inférieur. (voir diapo suivante)
- **FILTER** décrit les critères de recherche qui seront appliqués sur les entrées de la base. Par exemple : **(&(uid=pouchou*)(objectclass=person))** permet de ne récupérer que les entrées ayant l'attribut *uid* commençant par pouchou et l'attribut *objectclass* égal à "person".
  Un filtre simple permettant d'afficher le contenu complet de l'annuaire est : **(objectclass=*)**

Scope LDAP lors des recherches dans l'annuaire.



Filtre LDAP lors des recherches dans l'annuaire.

- Le champ Filtre permet de spécifier des critères de recherche sur des attributs. La forme d'un filtre est (attribut opérateur = valeur). Par exemple, (sn=pouchoulon) permet de trouver tous les objets dont l'attribut sn (surname) vaut pouchoulon.

- Les opérateurs & (AND) , |(OR), ! (NOT) permettent de poser des contraintes croisées sur le résultat renvoyé par le serveur.

Exemples de filtres

(|(uid=pouchou)(uid=mfacerias))

(& (|(sn=pouchou)(sn=mfacerias)) (telephoneNumber =(33)*))

(!(cn=Jean-marc Pouchoulon))

(&((ou=personnels)(|(manager=cn=Jean-marc Pouchoulon,ou=rt2,dc=iutbeziers,dc=fr) (manager=cn=Eric Dubreuil)))

(!(objectClass=InetOrgPerson))

Exemples de filtres

- *Approximation* (uid~=pouchou) *# L'uid ressemble à pouchou*
- *Egalité stricte* (uid=pouchou) *# recherche exactement*
  # pouchou
- *Comparaison* (cn>pouchoulon) , <= , >= , < Présence (sn=*)
  ##retourne les entrées ayant un attribut sn présent
- *Sous-chaîne* (uid=po*), (sn=*ou), (sn=p*p*)

  ## aka « Bash » expansion

Options ldapsearch

ldapsearch [options] [search_filter] [list_of_attributes]`

- -H ldap://nom_d_hôte:389
- -x authentification simple
- -D (nom de l'utilisateur qui va se « binder »)
- -b (l'endroit ou on lance la recherche dans le DIT)
- -s sub/one/base (recherche dans les sous arbres, sur un seul niveau ou uniquemement la base)

Exemple:

ldapsearch -x -D 'iutbeziers\utilisateur' -w 'pass' -H ldap://server-rt.iutbeziers.fr -b 'OU=Comptes,dc=iutbeziers,dc=fr' uid mail

LDAP au niveau protocolaire

- Connexion TCP
- Bind: phase d'authentification ou le client envoie sur le réseau son compte et son mot de passe. Un bind anonyme est possible.
- Suite d'opérations (add , modify , delete, moddn….)
- Unbind
- Déconnexion TCP

L'authentification

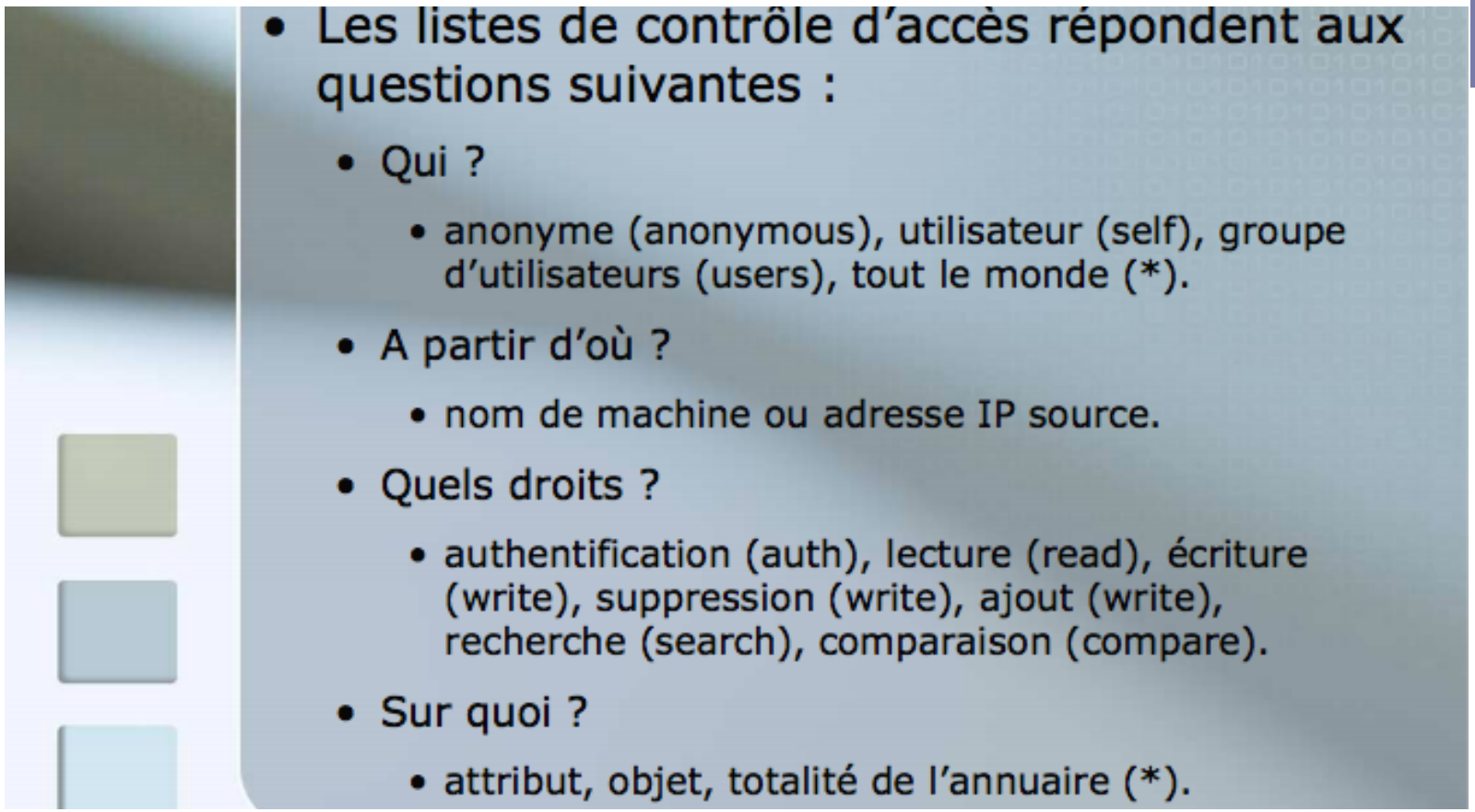
- Se fait en mode simple (envoie de l'identifiant utilisateur et de son mots de passe option –x du ldapsearch)
- En SASL (protocoles qui permet au client et au serveur de s'entendre sur la façon de dialoguer options –ZZ de ldapsearch) . SASL permet de n'utiliser que le port 389 y compris pour les usages chiffrés.

Modèle de sécurité

- Les droits d'accès :
  - permettent de gérer les autorisations sur la totalité des entrées de l'annuaire.
  - s'appliquent sur les objets et sur leurs attributs.
  - consistent à décrire les droits de certains objets de l'annuaire sur d'autres entrées.
- Cette description s'effectue à l'aide de règles (ACL).
- Chaque ACL comprend plusieurs règles (ACI).
- La syntaxe d'une ACI n'est pas normalisée.

Source: http://cesar.resinfo.org/IMG/pdf/formation_ldap_hybride.pdf

Modèle de sécurité



Source: http://cesar.resinfo.org/IMG/pdf/formation_ldap_hybride.pdf

Example sous open LDAP

access to *

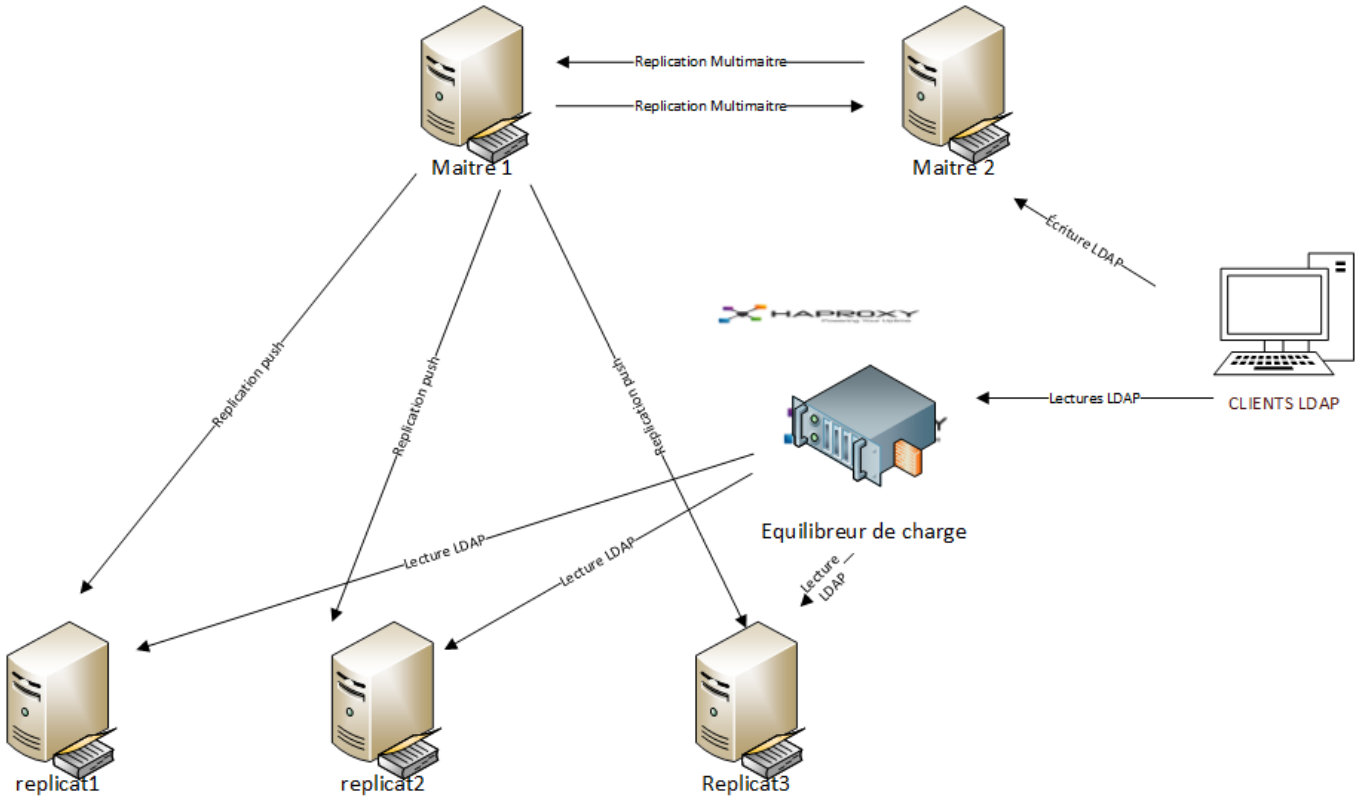
by dn="cn=admin,dc=iutbeziers,dc=fr" write

by anonymous auth

by users read

La réplication

- Les annuaires LDAP sont toujours répliqués car sans eux la production informatique s'arrête…



Réplication SYNCRPL (LDAP Sync Replication Engine)

Il existe un mécanisme de réplication standard (donc rendant interopérable la réplication entre annuaires différents):

The Lightweight Directory Access Protocol (LDAP) Content Synchronization Operation. (RFC 4533)

Il est implémenté entre autre par OpenLDAP et ApacheDS.

SYNCREPL garde une trace du statut de réplication en maintenant et en échangeant des cookies de synchronisation.

L'absence de Cookie correspond à une demande d'initialisation alors que la présence d'un cookie indique une demande de mises à jour.

SYNCRPL

Deux modes :

- refreshAndPersist : l'esclave (consumer) récupère en continu les modifications apportées au maitre (provider) suivant l'état maintenu par le serveur.
- refreshOnly : l'esclave demande périodiquement les mises à jour au maitre

Syncrepl est compatible avec la configuration multimaster et peut n'envoyer que les différences entre les entrées (granularité au niveau des attributs) avec le mode delta-syncrepl.

La réplication

- Réplication « push-based » ou SIR : Server Initiated Replication : c'est l'annuaire maître qui réplique vers les esclaves.
- Réplication « pull-based » ou CIR : Consumer Initiated Replication : ce sont les esclaves qui établissent les requêtes vers l'annuaire maître. Mode Disponible dans OpenLDAP avec Syncrepl (depuis la version 2.2)
- Réplication multimaîtres (multimaster) :La réplication n’est pas dirigée et fonctionne dans les deux sens.

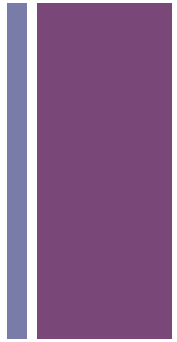
Syncrepl: Réplication différentielle « delta-sync »

- Cette méthode permet de conserver les modifications d'une instance sur un suffixe dédié pendant plusieurs jours. Les autres instances se connecteront sur ce suffixe pour récupérer uniquement les attributs modifiés.
- Si une entrée LDAP a beaucoup d'attributs et qu'un seul est modifié on évite ainsi de renvoyer l'entrée complète.

Referrals

- C’est le « faire suivre à » des annuaires.
- En demandant une entrée le client LDAP reçoit un lien à suivre.
- Utile pour déléguer une branche.

OpenLDAP



Bibliographie

- http://anf2012.mathrice.fr/lib/exe/fetch.php?media=rappels-ldap.pdf
- http://cesar.resinfo.org/IMG/pdf/formation_ldap_hybride.pdf