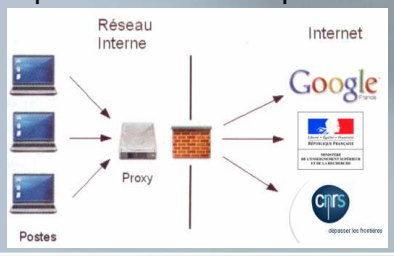
ADF 2009
Reverse Proxy

Thierry DOSTES
tdostes@ifr88.cnrs-mrs.fr

Définition d'un serveur mandataire

- Un proxy (ou serveur mandataire) :
  - agit comme une passerelle et un filtre pour accéder à l'Internet.
  - retransmet les requêtes envoyées par les machines locales.
  - permet une diminution de la bande passante utilisée en offrant un mécanisme de cache.
  - est visible pour les clients qui l'utilisent.

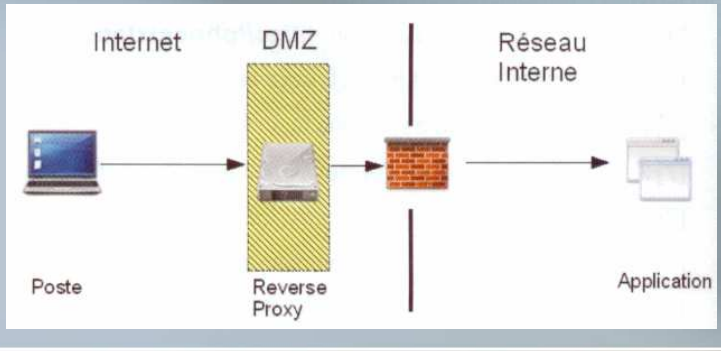


Journées Thématiques SIARS - ADF - 26 & 27 Janvier 2009

Le proxy permet le filtrage des ressources rapatriées (analyse virale du contenu des pages, analyse des scripts dans les pages, etc.). Utilisé en corrélation avec une journalisation de l'activité, les traces d'utilisation d'un serveur proxy permettent de lister les tentatives d'accès à des sites interdits et d'obtenir des statistiques sur l'utilisation de la bande passante (sites les plus accédés, utilisateurs les plus « gourmands »).

Définition d'un Reverse Proxy

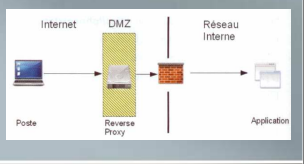
- Un reverse proxy est un relais inversé.
- Il donne l'accès depuis l'Internet à des services Web situés sur un réseau local.
- Il dialogue avec le client en se substituant au serveur vers lequel il relaie les requêtes.



Avantages d'un Reverse Proxy (1)

- Les clients n'ont pas connaissance du système mis en place.
- Le service peut optimiser les performances en assurant des fonctions de cache.
- Il offre aussi des fonctions de sécurité.

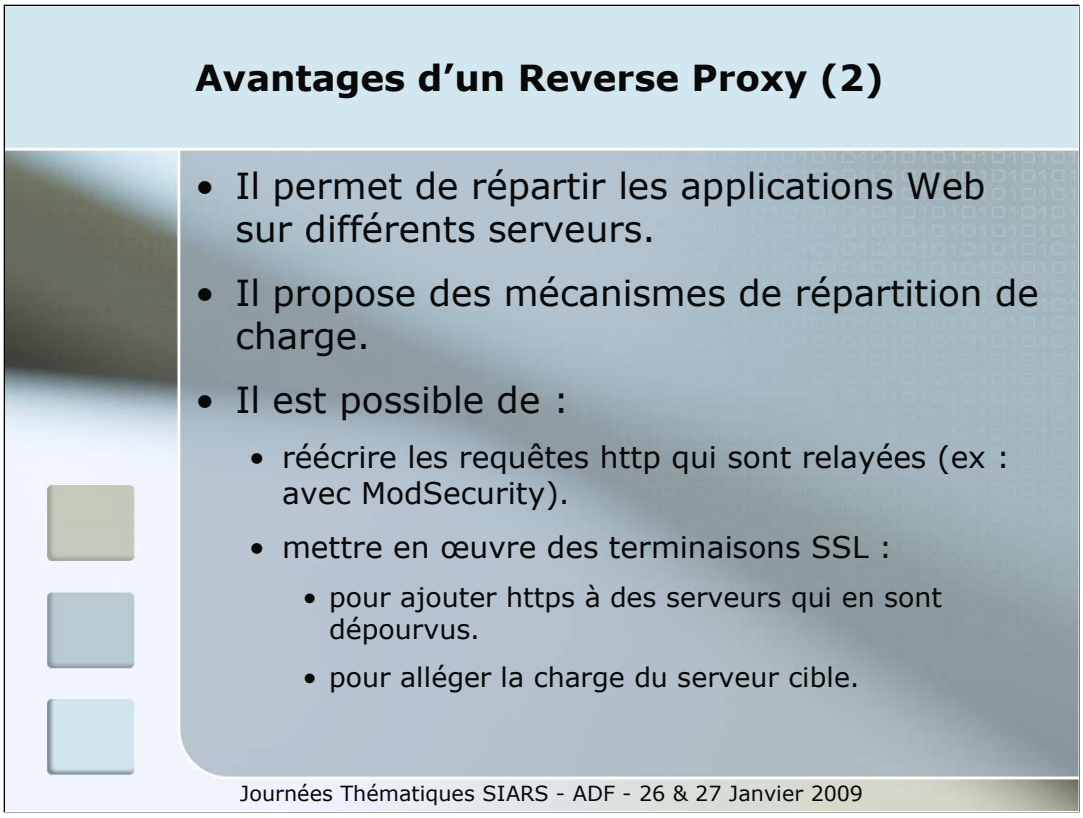
  ex : contrôle d'accès par adresse IP ou par authentification auprès d'un annuaire LDAP.
- Il devient le point d'entrée unique aux services Web :
  - filtrage.
  - centralisation des traces et de la gestion des erreurs 404.



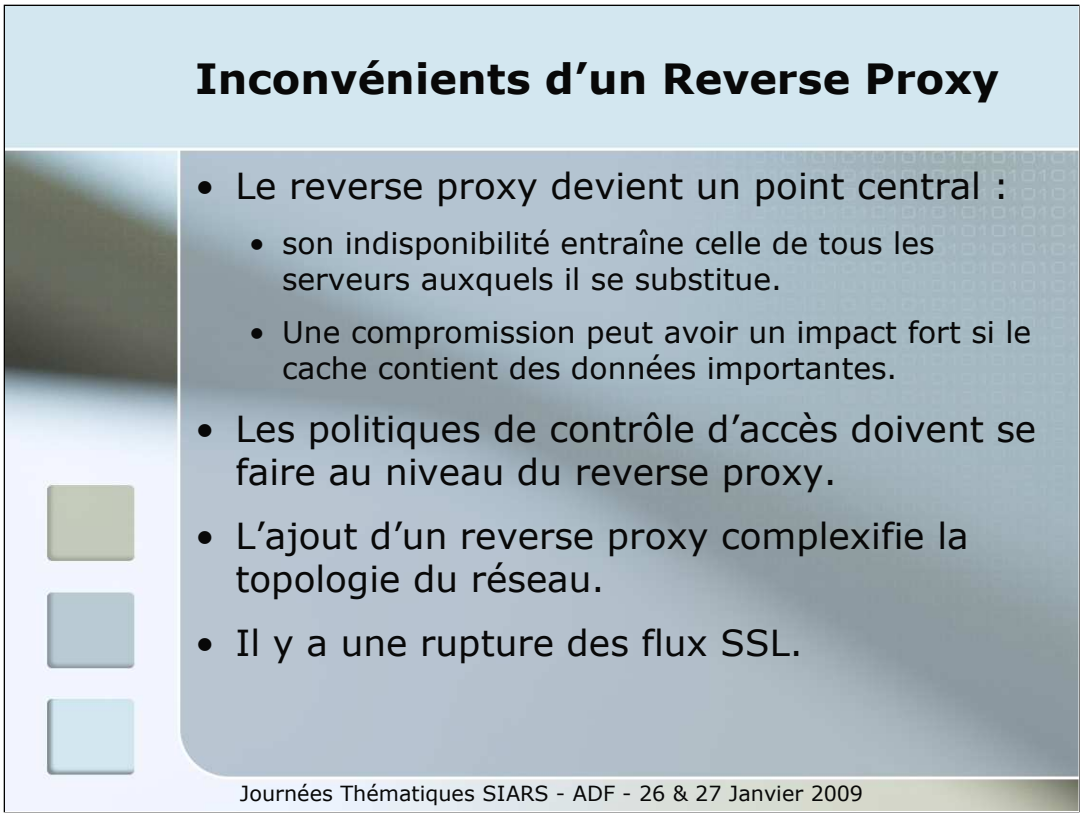
Journées Thématiques SIARS - ADF - 26 & 27 Janvier 2009

Par exemple, un serveur Reverse Proxy peut être utilisé en frontal de serveur d'applications tels que Zope.

Les logs d'un reverse proxy permettent d'inventorier les clients sollicitant le serveur applicatif, de lister les adresses accédées et de connaître le contenu des requêtes (partie DATA des requêtes HTTP).



L'approche idéale est qu'une seule application Web soit hébergée sur un serveur, pour limiter les risques de sécurité et augmenter la disponibilité des services en cas de problème. Grâce aux technologies de virtualisation, il est possible de mettre en œuvre ce type d'architecture tout en conservant un budget matériel raisonnable.



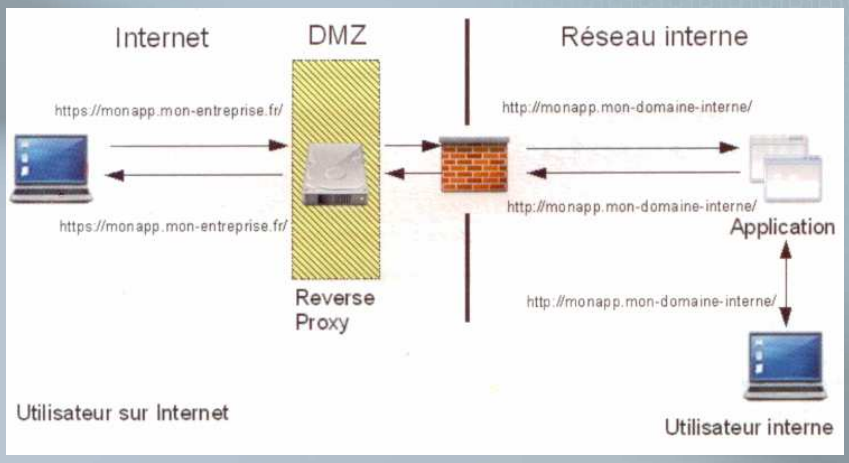
Les serveurs situés derrière le reverse proxy ne voient que son adresse IP. Il est donc indispensable de mettre en œuvre les politiques de contrôle d'accès à son niveau (ex : filtrage IP, authentification LDAP).

Reverse Proxy : pourquoi et comment ?

- Mettre à disposition depuis l'extérieur certaines applications :
  - fournir une communication chiffrée entre l'application et le client Web.
  - ajouter une authentification supplémentaire si nécessaire.
  - filtrer les données envoyées par le navigateur.
- Simplifier/contourner l'établissement de règles sur un pare-feu (ex : hébergeur).
- Déléguer la gestion de sites Web.

Reverse Proxy : un exemple

- Mettre à disposition depuis l'extérieur une application propriétaire :



Journées Thématiques SIARS - ADF - 26 & 27 Janvier 2009

Nous possédons une application interne qui est accédée au sein du réseau interne via l'URL http://monapp.mon-domaine-interne/.

Nous souhaitons y accéder depuis l'internet par l'intermédiaire de l'adresse : https://monapp.mon-entreprise.fr.

Analysons d'abord quels sont les éléments concernés par cette mise à disposition de l'application :

-au sein du réseau interne, lors des accès à l'application via l'adresse interne, toutes les requêtes HTTP et tous les entêtes de ces requêtes (par exemple les champs Location: ou SetCookie:) font référence au nom DNS interne de l'application.

-le contenu des pages générées par le serveur peut utiliser des références au nom DNS interne de l'application.

Le reverse proxy nous fournit des mécanismes qui nous permettent de réécrire les requêtes arrivant sur l'adresse https://monapp.mon-entreprise.fr en requêtes sollicitant l'adresse http://monapp.mon-domaine-interne/.

Quelques logiciels (1)

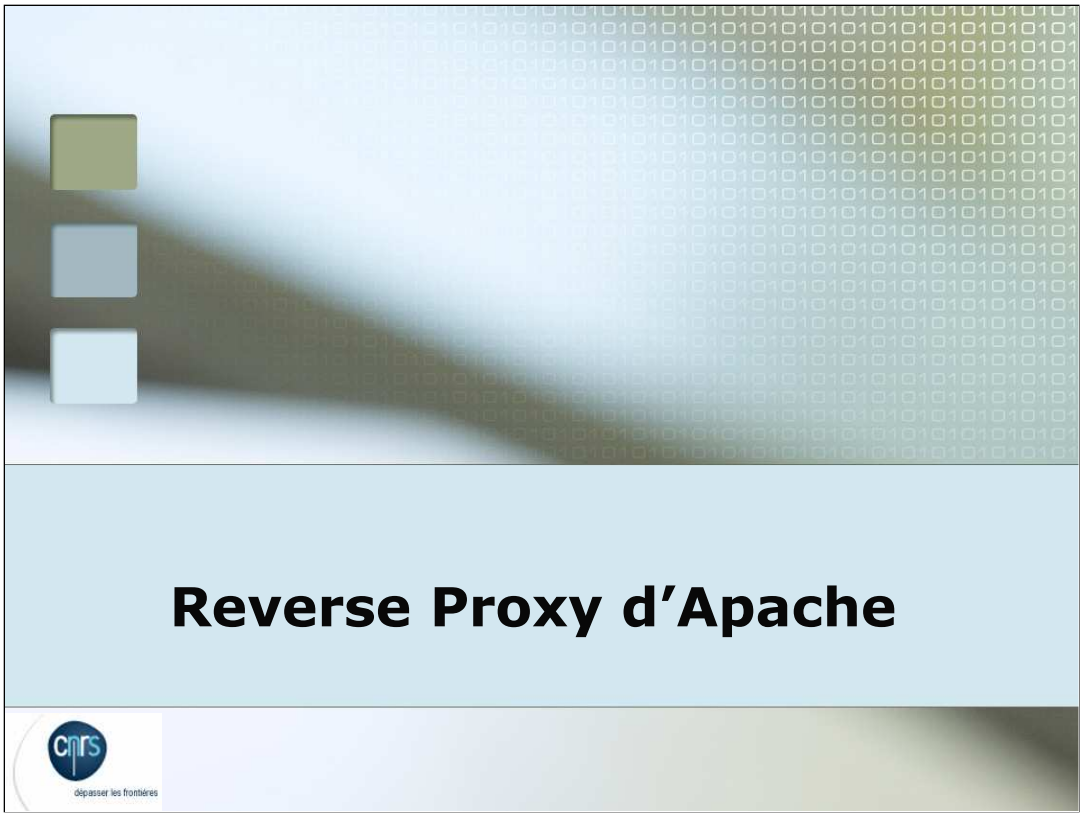
- Le choix d'une solution reverse proxy se fait suivant les fonctionnalités du serveur.
- Voici quelques critères essentiels :
  - gérer les hôtes virtuels.
  - offrir des mécanismes SSL.
  - disposer d'un mécanisme de cache.
  - permettre la réécriture des adresses (URLs).

Quelques logiciels (2)

- Squid :
  - Mécanisme de cache.
  - Gestion des hôtes virtuels (depuis les versions 3.x).
  - Rapide et stable.
- Pound :
  - Développé pour servir de solution frontale aux serveurs d'applications Zope.
  - Mécanisme SSL.
  - Equilibrage de charge.
  - Détection de panne de l'un des serveurs.

Quelques logiciels (3)

- Apache :
  - Mécanisme de cache.
  - Gestion des hôtes virtuels.
  - Terminaisons SSL.
  - Modules de réécriture des adresses.
  - Protection des applications Web à l'aide de modules dédiés.
    (ex : mod_evasive, mod_security).
- Varnish :
  - Accélérateur HTTP.
  - Equilibrage de charge et détection de panne.
  - Réécriture des adresses.

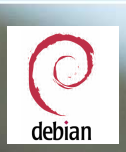


Reverse Proxy d'Apache

Présentation des modules

- Apache fournit plusieurs modules pour gérer les fonctions de reverse proxy :
  - mod_proxy : activation de la passerelle.
  - mod_proxy_http : gestion des requêtes HTTP effectuées auprès de la passerelle.
  - mod_proxy_ftp : gestion les requêtes FTP.
  - mod_cache : mise en place d'un système de cache.
  - mod_proxy_balancer : répartition de charge entre plusieurs serveurs.

Configuration du reverse proxy (1)



- Installons le serveur Apache2 :

```
apt-get install apache2 apache2-doc
```

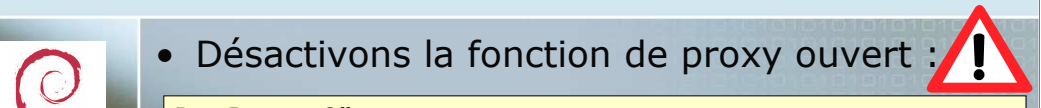
- Chargeons les modules permettant d'activer le reverse proxy pour les requêtes HTTP :

```
a2enmod proxy proxy_http
```

- La configuration du reverse proxy se fait dans le fichier :

/etc/apache2/mods-enabled/proxy.conf

Configuration du reverse proxy (2)



- Désactivons la fonction de proxy ouvert :

```
ProxyRequests Off
```

- Une redirection de type reverse proxy peut être activée de deux façons différentes :
  - en utilisant la directive ProxyPass.
  - En invoquant la directive RewriteRule avec le flag **[P]**.

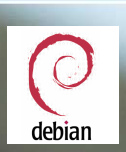
```
ProxyPass /intranet/          http://192.168.1.40/
ProxyPass /service_info/      http://192.168.1.20/spip/
```

Configuration du reverse proxy (3)

- Mettons en place une politique de filtrage pour l'accès à l'Intranet :
  - Les utilisateurs du réseau local peuvent accéder.
  - Les utilisateurs du réseau externe doivent s'authentifier auprès de l'annuaire (requiert le module **mod_authnz_ldap**).

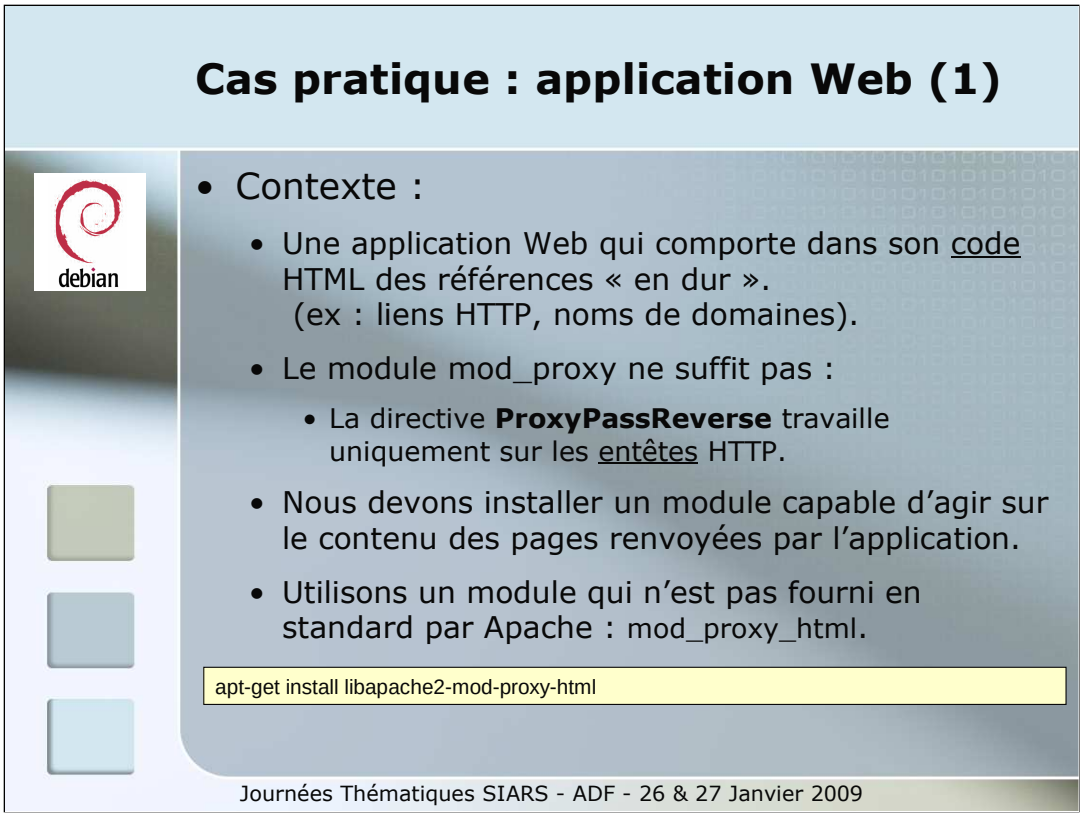
```
<Proxy http://192.168.1.40/>
        AuthType Basic
        AuthName "Zone Intranet IFR 88"
        AuthLDAPURL "ldap://ldap.ifr88.glm:389/ou=accounts,dc=ifr88,dc=cnrs,dc=fr"
        AuthBasicProvider ldap
        AuthzLDAPAuthoritative off
        require valid-user
        Order deny,allow
        Deny from all
        Allow from 192.168.1.0/255.255.255.0
        Satisfy Any
</Proxy>
```

Configuration du reverse proxy (4)



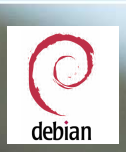
- Exemple de répartition de charge (après activation du module mod_proxy_balancer) :

```
<Proxy balancer://webmail>
        BalancerMember http://192.168.1.50:80/
        BalancerMember http://192.168.1.51:80/
</Proxy>
ProxyPass /webmail balancer://webmail/
```



La documentation de ce module est contenu dans le répertoire **/usr/share/doc/libapache2-mod-proxy-html/**.

Cas pratique : application Web (2)

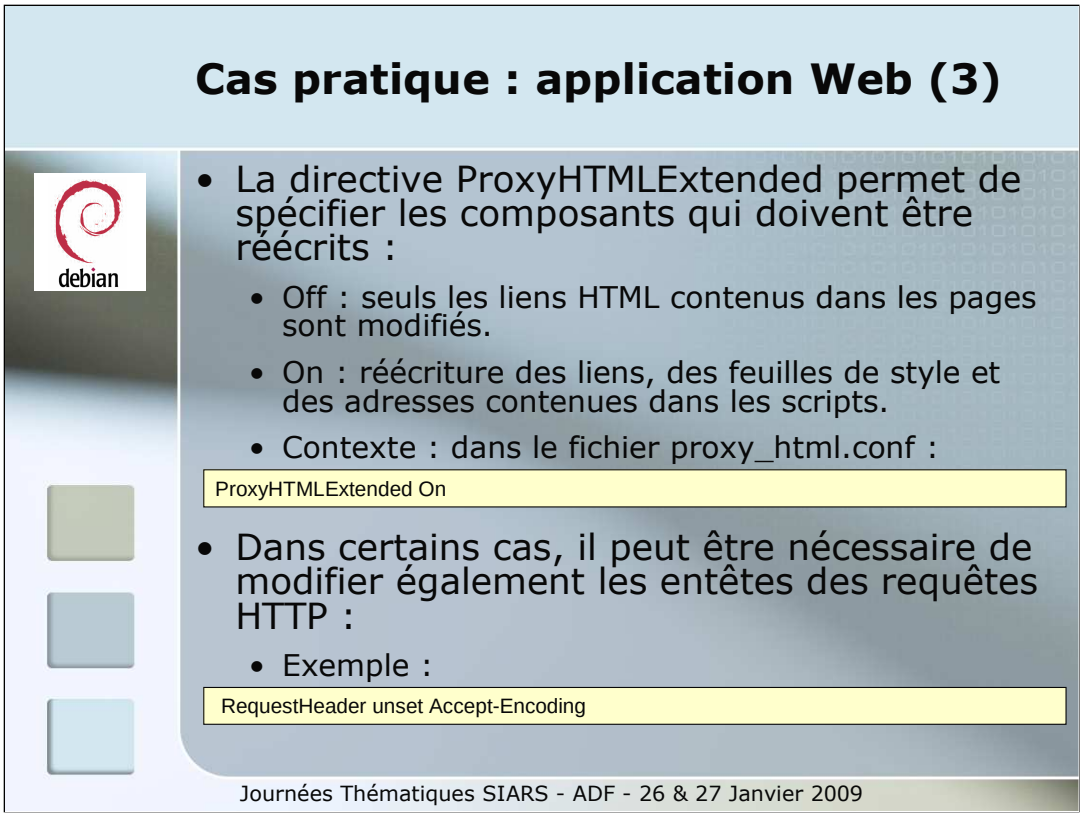


- Le module mod_proxy_html :
  - scrute le contenu des pages et permet les substitutions au sein de la partie *DATA*.
  - permet ainsi de réécrire les liens HTML dans une situation de proxy.
- Nous utilisons la directive **ProxyHTMLURLMap** pour définir les règles de réécriture des liens.
  - Syntaxe :

```
ProxyHTMLURLMap    from-pattern  to-pattern  [flags]  [cond]
```

  - Lorsque un lien correspond au modèle, la partie correspondante est remplacée.
  - Exemple :

```
ProxyHTMLURLMap    http://monapp.domaine.interne https://monapp.domaineentreprise.fr
```



En résumé, dans le cas d'applications contenant de nombreuses références codées en dur, l'utilisation conjointe des directives **ProxyPass** et **ProxyPassReverse** ne suffit pas car elles travaillent uniquement sur les entêtes HTTP.

En conséquence, nous utilisons le module **mod_proxy_html** qui permet la réécriture des adresses et autres références contenues à l'intérieur des pages.

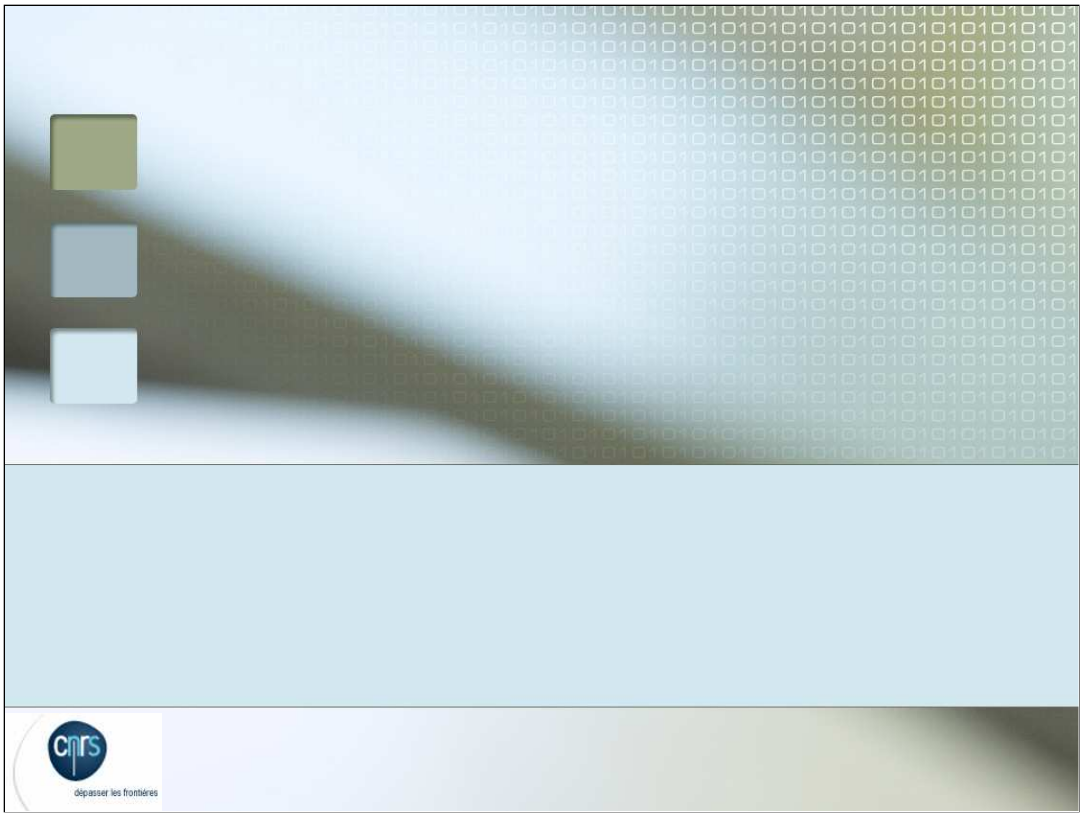
Enfin, la directive **RequestHeader** permet de configurer l'entête des requêtes HTTP. L'option **unset** supprime l'entête de requête correspondant au nom indiqué. Dans l'exemple ci-dessus, le champ **Accept-Encoding** permet au navigateur de spécifier ses préférences en matière d'encodage.

Référence : http://www.apachetutor.org/admin/reverseproxies

Merci à Eric pour ses multiples essais et ses retours d'expérience.

Conclusions

- Une architecture à base de reverse proxy :
  - donne une grande souplesse pour mettre à disposition des applications sur Internet.
  - fournit de nombreux services à valeur ajoutée de manière transparente.
  - offre la possibilité d'intégrer un pare-feu applicatif en amont de l'application.
  - permet de répartir une application par serveur (en corrélation avec la virtualisation).


dépasser les frontières