Le multimédia dans l'Internet

—————

Les réseaux de distribution de contenu (CDN)

Christophe Deleuze
ActiVia Networks
Christophe.Deleuze@activia.net

DESS Télécom

29 janvier 2002

Plan

- introduction : les réseaux de contenu
- principe des CDN
- systèmes
  - ✓ *delivery*
  - ✓ distribution
  - ✓ redirection
  - ✓ comptabilité *(accounting)*
- gestion
- interconnexion de CDN
- *peer-to-peer*

Introduction

- applications multimédia
  - ✓ qualité de service *(streaming)*
  - ✓ multicast
- difficulté de déployer ces services
  - ✓ pb techniques (échelle)
  - ✓ pb politiques
- solution à court terme pour une application spécifique
  - ✓ la distribution de contenus

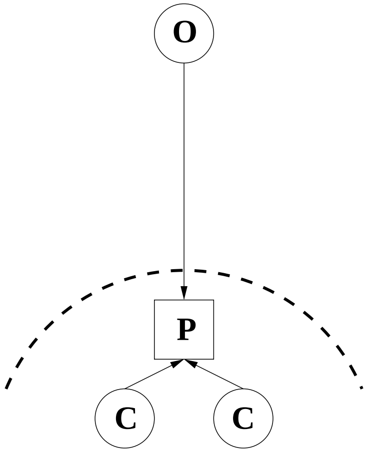
Introduction aux réseaux de contenu

Les réseaux de contenu *(Content Networks)*

- **contenu**
  - ✓ pages web (statiques, dynamiques) (HTTP)
  - ✓ *streaming* (RTSP)
  - ✓ à la demande
  - ✓ en direct
- **le problème de la distribution**
  - ✓ fournisseur du contenu = serveur origine
  - ✓ consommateur = client
  - ✓ point de rendez-vous
    - ☞ URL ➠ URI

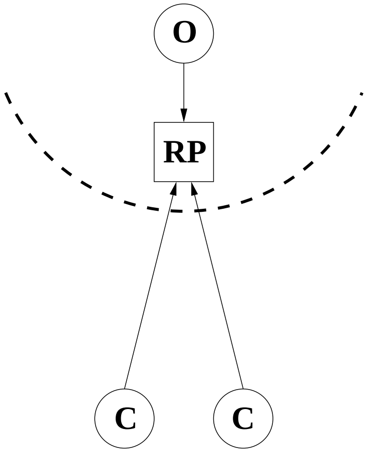
Foward Proxy Caching

- un *proxy* sert le contenu au client
- explicite/forcé/transparent
- géré par le FAI
  - ✓ + débit
  - ✓ + délai
  - ✓ + charge origine
  - ✓ – stats origine
  - ✓ – fraicheur



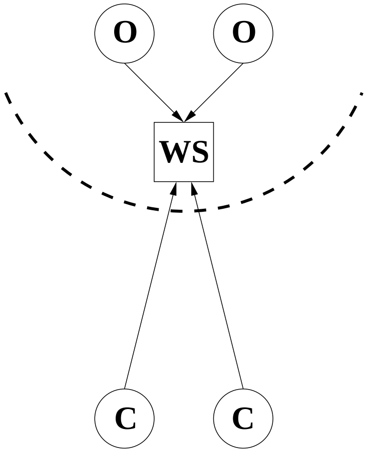
Accélérateur web *(reverse proxy)*

- origine caché derrière un proxy
- géré par le fournisseur de contenu
  - ✓ + charge origine
- variante : *web switch/server farm*



web switch/server farm

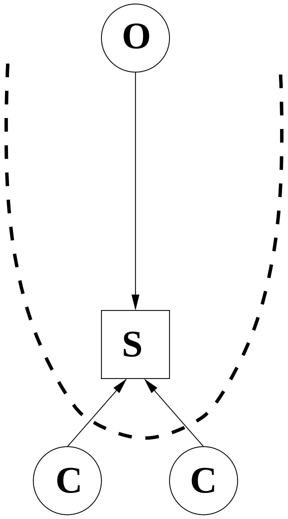
- *server farm* = 1 seule adresse IP
- *web switch* dispatche les requêtes sur les différents serveurs
  - ✓ charge
  - ✓ redondance



Content Delivery Network

surrogate
placé près des clients
géré par le fournisseur de contenu

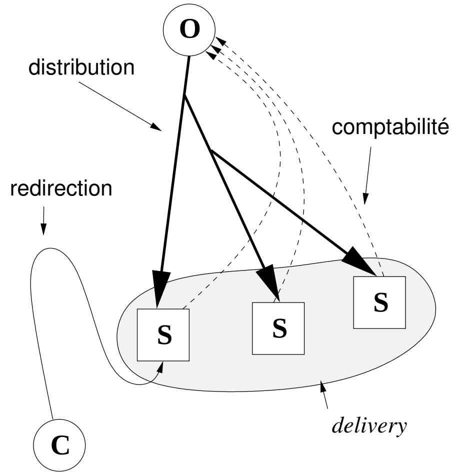
- + charge origine
- + débit réseau
- + délai client
- + stats origine
- + fraicheur



Les réseaux de distribution de contenu

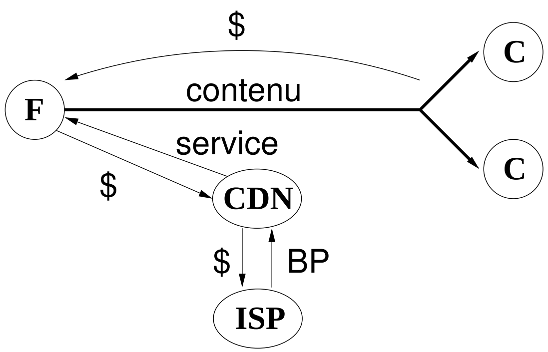
Qu'est-ce qu'un CDN ?

- rapprocher le contenu des clients
- quatre systèmes
  - ✓ *delivery*
    - ☞ serveurs *surrogates*
  - ✓ redirection
  - ✓ distribution
    - ☞ réplication
    - ☞ synchronisation
  - ✓ comptabilité



Qu'est-ce qu'un CDN ?

- fournisseur de contenu
- consommateurs
- opérateur réseau
- opérateur CDN

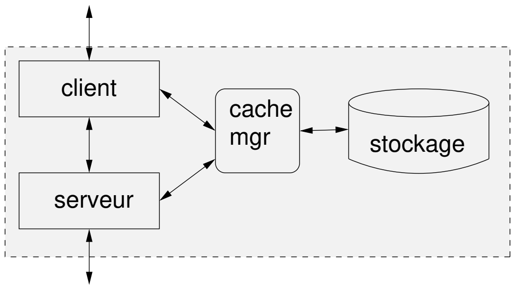


Delivery system

- proxy caches
  - ✓ web (statique, dynamique)
  - ✓ *stream*
    - ☞ real networks
    - ☞ windows media
- machines spécialisées (hardware, OS)
  - ✓ Network Appliances (NetApp) : *NetCache*
  - ✓ CacheFlow
  - ✓ Inktomi
  - ✓ Cisco : *CacheEngine*
  - ✓ Intel : *Web Traffic Express*

Architecture d'un *proxy cache*

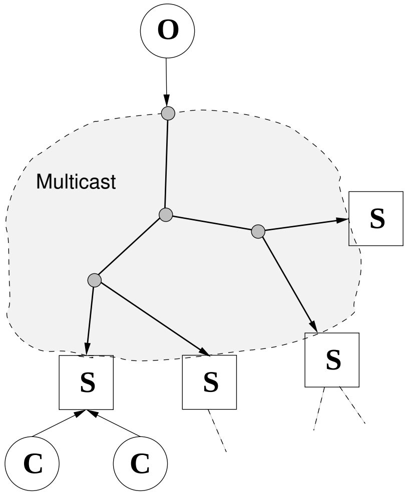
- performances (ex.Squid)
  - ✓ cache DNS
  - ✓ métadata en mémoire pour chaque objet caché
  - ✓ objets en transit non cachés n'accèdent pas au disque
  - ✓ organisation espace disque (nb. fichiers par répertoire)



Distribution : direct *(live)*

CDN ➠ réseau *overlay*

- multicast
  - ✓ soulage l'origine
  - ✓ soulage le réseau
- routage ad-hoc
  - ✓ ex min bandwidth
  - ✓ ...
- protection des flots
  - ✓ FEC
  - ✓ flots redondants



Distribution : réplication

(contenu à la demande)

- amener le contenu dans les *surrogates*
- ➠ localisation du trafic

push

- multicast
- problème de la fiabilité
- contrôle de flux
- ALM *Application Layer Multicast*

Distribution : réplication

pull

- *demand driven surrogates*
  - ✓ où chercher le contenu ?
  - ✓ hiérarchie de caches
  - ✓ voisins
- forme de routage

Distribution : synchronisation

maintenir la cohérence des *surrogates* avec l'origine

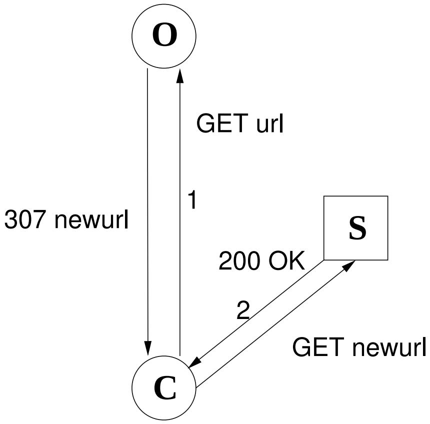
- *surrogate* = serveur authoritatif
  - ✓ invalidation de contenu
  - ✓ mises à jour
  - ✓ multicast
- *content signaling*
- tentatives actuelles
  - ✓ WCIP *Web Cache Invalidation Protocol* (IETF/Cisco)
  - ✓ RUP *Resource Update Protocol* (IETF)

Redirection *(request routing)*

- trouver le "meilleur" surrogate (équilibrage de charge global et local)
- surrogate = f(consommateur, contenu, état réseau, états surrogates)
- transparent pour le client...
- routage de requête
  - ✓ relayage *(forwarding)*
  - ✓ routage (construction des tables)
- différentes méthodes pour le relayage

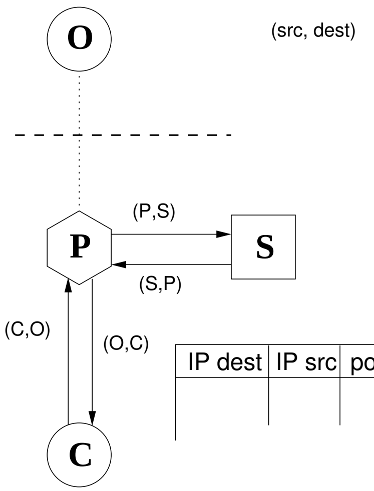
Redirection : *application redirect*

- ex HTTP : 307 temporary redirect
- \+ précis
  - ✓ adresse client
  - ✓ URL contenu
- – connexion vers l'origine
- – à implémenter pour chaque type de contenu



Interception niveau transport

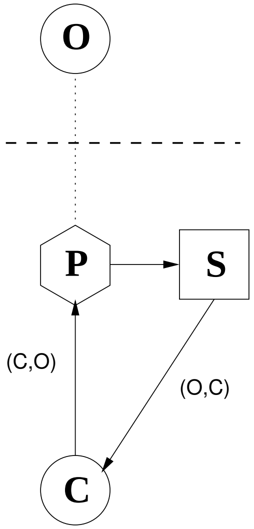
- switch L4 = proxy TCP transparent
- examen du premier paquet (`SYN`)
  - ✓ adresses IP (C,O)
  - ✓ port TCP



IP dest	IP src	port	S

Interception niveau transport avec triangulation

- canal descendant
- réseau d'accès
- réseau du serveur origine
- *spoofing*
  - ✓ P et S directement connectés
  - ✓ configuration spéciale de S

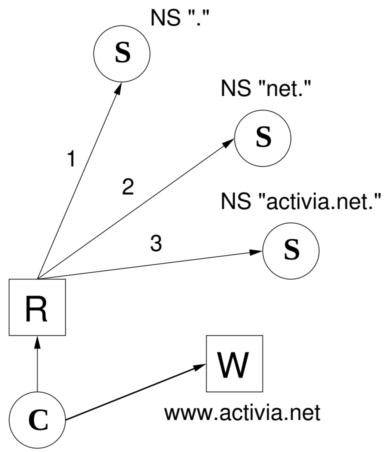


Interception niveau application

- URL
- en-têtes MIME
- proxy transparent *in-path*
  - ✓ utilisé dans les *Access Content Networks*
  - ✓ + rien à faire (f = cste = proxy)
  - ✓ − point de faiblesse pour **tout** le trafic
- proxy filtrant redirecteur
  - ✓ triangulation possible (mais délicat)

Rappels sur le DNS

- BD distribuée et redondante
- associations (nom, adr. IP)
- espace de nommage hiérarchique
- serveur
- résolveur *(cache server)*

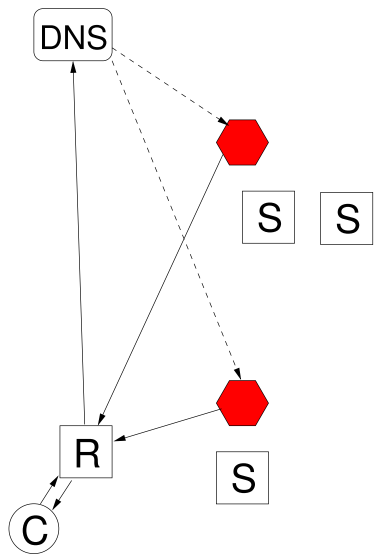


Redirection : DNS

- répondre selon le client
- + élégant
- + scalable, robuste (bientôt sécurisé)
- – précision
  - ✓ adresse du *resolver* du client
  - ✓ contenu : nom de domaine
    - ☞ `http://www.toto.com/toto.gif`
    - ☞ possibilité de conventions (www.* = HTTP port 80)
    - ☞ reécriture des URL (ex. Akamaizer)
- – *DNS accelerator*

DNS boomerang

- Boomerang/Flash DNS/DNS flooding
- les clusters répondent
- proximité réseau
- si charge ➠ retarder la réponse

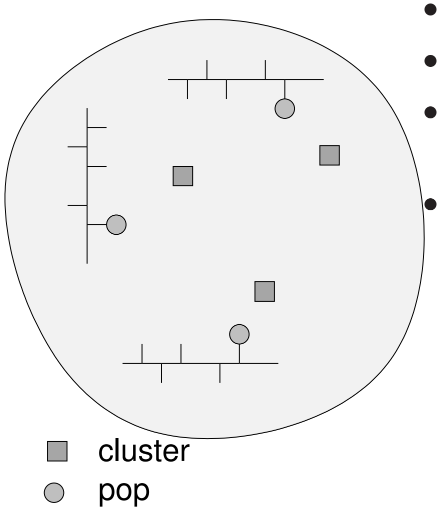


Redirection : routage

- un ensemble de PoP
- un ensemble de clusters de surrogates
- trouver pour chaque PoP le meilleur surrogate (pour un contenu donné)
- mesure métriques (PoP, cluster)
- monitorer surrogates
- RIPE, OREGON
- BGP, IGP

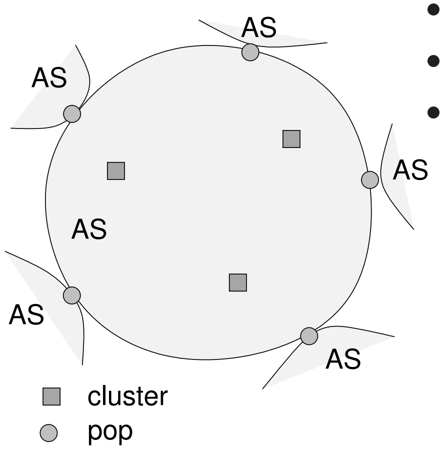
- Akamai ➠ *secret sauce*
- optimal ➠ marketing
- trouver un surrogate raisonnable

Redirection : routage : intra AS



- clusters de *surrogates*
- pop (agrégats de clients)
- associations { cidr } ➠ pop (ex. IGP)
- association pop ➠ cluster

Redirection : routage : extra AS



- point de *peering*
- = pop
- liste des CIDR fournie par BGP (route “descendante”)

Comptabilité

- les logs sont répartis dans tous les *surrogates*
- besoin de générer une vision globale
- scalabilité ?
- structure hiérachique (multicast inverse)

Gestion du CDN

- deux points de vue
  - ✓ fournisseur de contenu
  - ✓ opérateur du CDN
- fournisseur de contenu
  - ✓ publication/invalidation
  - ✓ monitoring
  - ✓ reporting
- opérateur du CDN
  - ✓ configuration infrastructure
  - ✓ monitoring

Le besoin d’interconnecter les CDN

- qqes CDN mondiaux (Akamai...)
- intérêt des ISP
  - ✓ service ➠ revenu
  - ✓ infrastructure
- mais pb de couverture
- ➠ interconnecter les CDN

- Content Alliance
- Content Bridge
- IETF CDI *Content Distribution Internetworking*

Interconnecter les systèmes

- *delivery* = union des *delivery*

Trois systèmes à faire *peerer*

- comptabilité ➠ simple à priori
- redirection
  - ✓ qui délivre le contenu au client ?
- distribution
  - ✓ qui participe à l'acheminement du contenu ?

CDN de transit ?

Perspectives

- URN : uniform resource name
  - ✓ identificateur de ressource persistent et indépendant de l'emplacement
  - ✓ ex. : `urn:foo:a123,456`
- *Application Distribution Networks*
- fin du bout en bout
  - ✓ *middleware* d'interception
  - ✓ proxies en tout genre
- adresses *anycast*
- *peering* de CDN
- peer-to-peer ...

Perspectives : *peer-to-peer*

- chaque client est aussi un serveur
- Napster
  - ✓ catalogue sur serveur centralisé
- Gnutella
  - ✓ connection avec qqes sites
  - ✓ requêtes diffusées par inondation

peer-to-peer : Freenet

- réseau anti-censure
  - ✓ publier des documents anonymement
  - ✓ récupérer des documents anonymement
  - ✓ suppression de document pratiquement impossible
  - ✓ décentralisé
- fonctionnement idem Gnutella, mais
  - ✓ document caché sur tout le retour de la requête
  - ✓ gestion de cache LRU *(least recently used)*
  - ✓ réplication des documents populaires
  - ✓ déplacement des documents près de la demande
- uprizer.com

peer-to-peer : Swarmcast

- swarm = essaim, fourmillement
- système de distribution de fichiers (gros)
- principe
  - ✓ encodage : découpe le fichier en n morceaux *(chunks)* avec FEC
  - ✓ rappatriement : depuis tous les autres utilisateurs courants
  - ✓ re-assemblage
    - ☞ n morceaux quelconques
    - ☞ FEC recalculée par chaque nœud
  - ✓ re-distribution
- opencola.com

Références

- WREC *Web Replication and Caching* www.wrec.org (maintenant WEBI)
- WEBI *Web Intermediaries* (IETF)
- Content Alliance www.content-alliance.org et CDI *Content Distribution Internetworking* (bientôt IETF)
- Content Bridge www.content-bridge.org
- OPES *Open Pluggable Edge Services* www.ietf-opes.org (bientôt IETF)
- openp2p.com O'Reilly network
- google.fr ...

...

Fin

Document réalisé avec LaTeX.
Généré le 29 janvier 2002.
Classe de document foils.
Dessins réalisés avec xfig.

Christophe Deleuze, Christophe.Deleuze@activia.net
http://www.activia.net

Ce document est disponible en postscript compressé avec gzip à
http://rp.lip6.fr/~deleuze/activia.html