



Extrait du PI-GROUPE

<http://www.pi-groupe.net>

Etude économique et sociale du Peer-to-Peer

- FRANCAIS - Ressources & Publications - Publications -

Date de mise en ligne : samedi 5 septembre 2009

Description :

Une étude en trois volets par Patrice Lamarque & Mamadou Dian Baldé. Février 2002, DESS Communication Réseau Image Université de Marne-la-Vallée.

PI-GROUPE

Avec l'arrivée de l'Internet et son rôle croissant dans l'économie et la société, l'informatique réseau n'a eu de cesse de trouver des innovations pour exploiter les ressources qu'un réseau de cette ampleur contient.

Le monde de l'informatique est en effervescence autour d'un phénomène portant le nom barbare de peer-to-peer. Mal identifiée, mal comprise et mal considérée à ses débuts, l'idée a beaucoup mûri au cours des deux dernières années. Aujourd'hui, on parle sérieusement du peer-to-peer comme un modèle de communication capable de changer radicalement certaines approches de l'informatique en réseau.

L'objectif de cette étude n'est pas de faire une présentation exhaustive des tenants et aboutissants du peer-to-peer, mais plutôt d'observer le phénomène avec un regard technique, économique et social en présentant quelques exemples révélateurs.

La première partie couvre les aspects techniques. Nous tentons d'y situer l'apparition de l'informatique peer-to-peer dans le contexte de l'Internet. Dans une présentation générale du concept nous décrivons quelques idées qui naviguent autour de l'informatique peer-to-peer avant d'exposer sous quelles formes d'architectures il est correct d'envisager des systèmes peer-to-peer.

Le second volet de cette étude analyse les enjeux économiques. Après une revue des types d'applications possibles, nous essayerons de déterminer la viabilité du concept dans le contexte économique actuel.

Une approche sociale vient enfin clôturer notre présentation. Nous observerons l'impact sur les comportements de l'internaute avant de soulever les problèmes d'ordre moral et éthique que certaines applications peer-to-peer connaissent.

Chapitre 1 - Présentation technique

Ce chapitre s'attache à présenter les aspects techniques de notre étude sur le peer-to-peer. Nous livrons tout d'abord un bref historique de l'informatique pour y situer l'apparition du peer-to-peer. Une étude technique introduit le concept tel que nous le connaissons aujourd'hui et l'analyse. Nous décrivons ensuite les principales architectures possibles, pour finalement considérer les possibilités d'optimisation dont peuvent bénéficier les protocoles peer-to-peer.

Historique

L'informatique s'est d'abord développée autour du concept de l'ordinateur central tout puissant. C'est ce qui a donné naissance aux grands systèmes centraux comme l'IBM360. Les terminaux ne communiquent alors entre eux qu'au travers de la machine centrale. Ce concept a ensuite évolué vers un concept de client/serveur quasi-identique, dans ce modèle les clients sont attachés au serveur par le réseau au travers d'un logiciel propriétaire.

L'hégémonie du client/serveur

Le monde universitaire, fonctionne selon ses propres logiques internes, mais il subit également des influences issues des autres acteurs sociaux. Ainsi, chaque secteur cherche avant tout à conquérir le plus d'autonomie possible par rapport aux autres. Et, c'est dans ce contexte de conquête de l'autonomie qu'il faut comprendre l'origine d'Arpanet, l'ancêtre d'Internet.

De l'Arpanet à l'Internet

Dans les années 60, le monde de la recherche informatique universitaire subit les conséquences de la logique de concurrence du milieu économique. Il faut savoir que, déjà à l'époque, le monde de l'informatique est dominé par l'économie de marché. Les grandes entreprises d'informatique font en sorte que leurs logiciels ne soient pas compatibles avec ceux de la concurrence, et cela pour s'assurer le contrôle de la production d'ordinateurs et de logiciels. Dès lors, il est quasi impossible de partager des ressources et des données entre ordinateurs construits par des sociétés différentes. Les firmes espèrent ainsi obliger les centres de recherche et les entreprises, qui sont les seuls à utiliser l'informatique vu le coût et la taille volumineuse des ordinateurs de l'époque, à se tourner exclusivement vers leurs produits pour assurer la compatibilité entre les machines.

Mais cette logique pose un problème par rapport à la dynamique interne du milieu de la recherche. Le fonctionnement de ce secteur repose sur une logique de progression du savoir scientifique qui nécessite pour les chercheurs de partager les ressources et les informations qu'ils possèdent. Dès lors, pour les ingénieurs qui travaillent au développement de systèmes informatiques, l'incompatibilité qui existe entre les ordinateurs des différents centres de recherche représente un frein à ce partage du savoir lorsqu'il doivent échanger des ressources avec des collègues d'autres centres de recherche qui n'utilisent pas le même matériel.

Pour résoudre ce problème, un ingénieur de l'agence ARPA (Advanced Research Project Agency) du département de la défense de l'armée américaine, R. Taylor, va trouver une solution tout en permettant au monde de la recherche informatique de contrecarrer la stratégie des fabricants d'ordinateurs. Si les ordinateurs ne peuvent permettre l'échange direct de ressources, alors, il faut mettre ceux-ci en réseau et ce de manière à ce qu'ils puissent communiquer sans problème, même s'ils sont produits par des sociétés différentes. Pour Taylor, cela représente la seule solution pour que l'informatique serve réellement à l'avancée du savoir et non à la seule recherche de profit économique.

Pour atteindre cet objectif, ce réseau doit être mis au point sous l'égide des chercheurs eux-mêmes et non par une firme commerciale qui, en se l'appropriant, risquerait de préserver le pouvoir des acteurs économiques sur le monde de la recherche informatique. Ce projet de réseau défini par Taylor représente véritablement une possibilité, pour la recherche, de s'approprier la fonction de communication des ordinateurs, en libérant cette fonction du pouvoir économique et donc d'autonomiser ainsi, en partie, l'univers de la recherche par rapport à l'activité économique.

Ainsi, en 1969, l'ARPA a mis au point le premier réseau mondial ARPANET qui ne dépend pas d'un centre névralgique, et qui ne pourrait donc pas être détruit en cas d'attaque nucléaire par une puissance étrangère. Le réseau a subi diverses améliorations au cours des années 70 et 80, notamment l'introduction de plusieurs protocoles qui font que les serveurs se comprennent en tenant le même langage ; SMTP pour l'échange de courrier électronique ou FTP pour les échanges de fichiers par exemple.

En 1990, Arpanet cède définitivement la place à l'Internet et, avec l'introduction des liens hypertextes en 1992, il ne cesse de devenir plus populaire et compte aujourd'hui plus de 30 millions de serveurs dans le monde pour plus de 100 Millions d'utilisateurs. Arpanet fut créé en reliant quatre ordinateurs situés chacun dans des centres universitaires différents. Au niveau technologique, ceci représente une grande originalité qui va permettre l'avènement d'un espace de communication et de partage de ressources totalement libre.

Naissance du client/serveur

L'Internet est né comme un vrai réseau peer-to-peer. Toutes les communications entre les hôtes de ce réseau

étaient égales. Il n'y avait pas de concept de client ni de serveur comme aujourd'hui. Les peers de l'Internet original étaient des participants actifs, fournissant des services.

L'Internet a alors grandi, et avec sa croissance la centralisation dut être ajoutée pour supporter la montée en charge. Des machines plus petites, comme les PC qui manquaient de ressources et de bande passante, devinrent clientes de ce réseau. Elles ne pouvaient pas être des fournisseurs actifs du système. Ces facteurs créèrent la nécessité de l'architecture client/serveur.

Même si le protocole de communication de l'Internet (IP) a été conçu dès le début comme un système symétrique, les premières « Killer applications » de l'Internet, comme FTP pour l'échange de fichiers, ou Telnet pour utiliser une machine à distance, étaient en fait des applications client/serveur.

La particularité du protocole TCP/IP est d'être parfaitement symétrique. Les premières applications qui ont réellement utilisé cette symétrie ont été les serveurs Usenet et les serveurs de noms de domaines (DNS). Dans les deux cas le problème fut celui du partage de données.

Il est intéressant de noter que dans les deux cas, c'est une structure hiérarchique qui a permis de résoudre ce partage d'espace. Par exemple, talk.philosophy.human est un exemple de nommage hiérarchique pour un newsgroup. De la même façon, le DNS adopte un nommage hiérarchique pour les noms de domaine, www.finger.com, par exemple.

Le rôle des interfaces graphiques

Le second événement clé de l'histoire de l'informatique est l'apparition des interfaces graphiques dans les années 1970, depuis les laboratoires XEROX à Palo-Alto, intégrées dans le Macintosh d'Apple en 1978. A une époque où les micro-ordinateurs sont d'austères consoles texte réservées à un faible groupe d'initiés que peu de monde comprend, les interfaces graphiques offrent la possibilité aux utilisateurs novices d'appréhender l'informatique. C'est le début de la popularisation de l'informatique.

Ces deux concepts, que sont les réseaux et les interfaces graphiques, convergent avec l'apparition des navigateurs HTTP. En 1994, la société Netscape crée Mosaic, le premier navigateur web conçu pour présenter de manière agréable les documents, et surtout de naviguer aisément dans l'espace du réseau internet.

L'usage d'Internet explose et, encore une fois, c'est le modèle simpliste du client/serveur qui est privilégié pour les serveurs web. Le réseau Internet est rapidement perçu comme un simple éventail de pages web.

Le modèle client/serveur en vigueur repousse les internautes en bordure du réseau en tant que feuilles de l'arbre alors que les serveurs forment le cœur du réseau.

L'apparition du peer-to-peer

Garantissant un bon contrôle de l'information le client/serveur s'est montré capable de répondre parfaitement aux attentes des acteurs de l'Internet. Sa suprématie s'est longtemps montrée la seule solution acceptable pour intégrer les clients légers au réseau. Mais le modèle client/serveur s'éloigne de la philosophie égalitaire qui avait donné naissance à l'Arpanet.

Un concept redécouvert

Fin 1998, un étudiant de l'université de Boston, Shawn Fanning, alors âgé de 19 ans, vient bouleverser le monde bien établi du client/serveur. Il écrit, pour lui et ses amis un logiciel gratuit qui permet d'échanger des fichiers audio au format mp3 à travers internet.

La raison d'être de ce logiciel repose sur le constat suivant : rechercher des mp3 sur les moteurs de recherche habituels conduisait à une perte de temps énorme tant les réponses étaient inappropriées. Fanning, dont le pseudo sur les forums était napster, donnera ce nom à son application.

Ce fameux Napster est constitué de l'assemblage d'un navigateur web, d'un serveur de fichiers, d'un module de messagerie instantanée et d'un lecteur de fichiers mp3. Le serveur permet de mettre à jour en permanence la liste des fichiers audios partagés par les utilisateurs connectés.

L'originalité réside dans le fait que les fichiers transférés ne le sont pas au travers d'un serveur mais directement d'un utilisateur vers un autre utilisateur, définissant ainsi le principe d'échange du « peer-to-peer » ; c'est à dire d'usager à usager.

En fait ce système redécouvre l'Internet dans son essence fondamentale, lorsque l'Internet offrait une structure symétrique pour faciliter le partage de l'information à l'intérieur de la communauté scientifique des premières universités connectées.

Raffinement du concept

Devant les menaces de la justice pour violation des droits d'auteurs, puis l'interruption du service sur décision judiciaire, de nouveaux logiciels similaires apparaissent. Mais, au lieu de faire appel à un serveur centralisé indexant les ressources pour mettre en relation les usagers, ils transfèrent cette fonction d'annuaire à des milliers d'ordinateurs de par le monde : le « pur peer-to-peer » est né. Le modèle semble indestructible, tout comme le réseau Arpanet.

Ce concept connaît bien d'autres prolongements que l'échange de morceaux de musique entre jeunes mélomanes, et les entreprises cherchent désormais à surfer sur la vague et à tirer parti de cette tendance.

Au delà du simple utilitaire de partage de fichiers mp3 qui a fait sa renommée, l'informatique pair à pair (dite également « point à point » ou « poste à poste ») revêt des nombreuses applications commerciales.

Présentation

Si le mot peer-to-peer a pris toute son ampleur très récemment, l'informatique peer-to-peer, elle, ne date pas d'hier. Il y a 30 ans, certaines entreprises travaillaient déjà avec des architectures qui seraient aujourd'hui qualifiées de pair à pair.

Une tentative de définition

Dans son essence, l'informatique pair à pair se définit comme le partage des ressources et des services informatiques par échanges directs entre systèmes. Ces échanges peuvent porter sur les informations, les cycles de traitement, la mémoire cache ou encore le stockage sur disque des fichiers.

Contrairement au modèle client/serveur, le modèle de communication peer-to-peer fait de chacun des nœuds du réseau une entité complète qui remplit à la fois le rôle de client et de serveur.

Aujourd'hui, néanmoins, plusieurs facteurs participent à l'explosion du phénomène : la puissance, la largeur de bande passante et la capacité de stockage, désormais disponibles à bas prix. Le peer-to-peer tire parti de la puissance existante des PC et de la connectivité des réseaux afin de mettre le potentiel collectif des nœuds au service de l'ensemble du réseau.

Notons que le terme peer-to-peer n'aide pas à clarifier les choses. Le jeu en réseau Doom, l'email ou même le téléphone peuvent être vus comme des systèmes peer-to-peer (on échange des informations directement entre pairs) alors que Napster, qui a popularisé le concept, est paradoxalement construit autour d'un serveur central.

Le vrai changement, c'est la nature des éléments qui constituent les réseaux peer-to-peer. Dans le passé, les centaines de millions d'ordinateurs connectés à Internet de manière intermittente ne faisaient pas partie du réseau. Avec le peer-to-peer, les ordinateurs personnels ont le pouvoir de devenir une partie intégrante du réseau.

Le peer-to-peer désigne donc une classe d'applications qui tirent partie des ressources matérielles ou humaines qui sont disponibles en bordure du réseau de l'Internet.

Caractéristiques

Comme ces ressources ont une connectivité instable ou des adresses IP variables, elles fonctionnent de manière autonome, indépendamment de systèmes centraux comme les DNS. Ce qui a rendu Napster et des systèmes similaires populaires, c'est le fait de tirer partie des ressources qui étaient auparavant inutilisées en tolérant d'une connectivité aléatoire.

Un vrai système peer-to-peer peut se reconnaître en se posant la question des deux caractéristiques suivantes :

1. Est-ce que le système permet à chaque pair de se connecter de manière intermittente avec des adresses IP variables ?
2. Est-ce que le système donne à chaque pair une autonomie significative ?

Si la réponse à ces deux questions est oui, le système est un système peer-to-peer. Notons au passage que le modèle client/serveur apporte une réponse négative aux deux questions.

Une autre manière de distinguer un système peer-to-peer est de raisonner en terme de « propriété ». La question à se poser est alors : « Qui possède les ressources qui font tourner le système ? ». Dans un système comme celui de Yahoo ! un grand architecte des services web -, l'essentiel des ressources est possédé par la société et ses partenaires, qui les mettent à la disposition des internautes, tandis que dans un système comme celui de Napster, les ressources sont un bien commun des membres du réseau.

Dans les vrais systèmes peer-to-peer, on ne se contente pas de décentraliser les fonctions mais aussi les coûts et les charges d'administration. Si l'on fait un calcul basique, 100 millions de PC connectés à l'Internet avec 100 Mo d'espace disque représentent une puissance potentielle de 10 000 téraoctets de stockage. C'est de ces ressources que le peer-to-peer permet de tirer parti.

Une caractéristique importante des réseaux de peers décentralisés est que la valeur perçue est directement liée à la quantité et la qualité des données qui y sont disponibles. Les ressources sont ajoutées à mesure que le nombre de peers du réseau augmente. Donc la valeur du réseau augmente avec sa popularité.

Enfin, la pertinence d'un système peer-to-peer réside dans sa capacité à localiser les ressources efficacement quelle que soit la taille du réseau. Aussi, ces systèmes doivent reposer sur des méthodes efficaces de découverte des ressources désirées. En général, la solution employée est l'utilisation de méta-données.

Avantages et inconvénients

Les systèmes peer-to-peer, en plus d'élargir le nombre de ressources disponibles pour les services, permettent d'en faire un usage plus efficace. Ces ressources inexploitées, se trouvant en bordure du réseau.

Le peer-to-peer trouve parfaitement sa place dans le travail collaboratif. Malgré l'utopie de pouvoir éviter l'usage de serveurs centralisés, peu d'applications (ou services) pourront effectivement se passer de points d'entrée centralisés. La combinaison des services web (centralisés) et du peer-to-peer apporte une collaboration totale, en temps réel ou asynchrone. Certaines applications peer-to-peer seront mieux adaptées pour fonctionner en Intranet afin d'exploiter et de partager l'information entre les utilisateurs (partage des fichiers, recherche dans des formats de fichiers différents : mail, tableurs, etc).

Les réseaux peer-to-peer doivent être capables d'identifier uniquement les ressources disponibles. Ce que le DNS tente de réaliser pour les machines, le peer-to-peer doit le faire au niveau des ressources. Par conséquent, les systèmes peer-to-peer ont dû adopter leurs propres schémas de nommage et de localisation indépendamment des adresses IP. Cette caractéristique leur permet de tirer habilement parti de la disponibilité intermittente des ressources se trouvant en bordure de l'Internet (celles des machines personnelles).

Un réseau peer-to-peer présente la particularité de ne fonctionner correctement que lorsqu'un nombre suffisant de nœuds participe au partage de ressources. Un réseau peer-to-peer peu populaire ne sera pas capable de survivre sans participants. A l'opposé, un réseau peer-to-peer mal adapté à la montée en charge risque de souffrir de cette popularité. La nature décentralisée des réseaux peer-to-peer implique un trafic important pour localiser les autres participants, supporter l'intermittence des connexions et acheminer les données. Certaines études [AH1000] montrent qu'un système fortement décentralisé peut générer un trafic plus que néfaste pour la bande passante.

En facilitant le partage et la diffusion d'information, les systèmes peer-to-peer créent une certaine redondance des données les rapprochant physiquement des bordures du réseau. Ce mécanisme de réplication de l'information est profitable aux nœuds disposant d'une faible bande passante.

Le contrôle de l'information dans une architecture décentralisée rend l'administration et la sécurité difficiles à maîtriser. La problématique de la sécurisation peut être résolue par des bibliothèques de fonctions peer-to-peer dédiées faisant intervenir des techniques de cryptage et d'authentification à clé publique. Néanmoins,

comme l'a montré le cas Napster, l'échange des contenus protégés par la propriété intellectuelle (ouvrages, musique, vidéos, jeux, logiciels) est très difficilement contrôlable dans une communauté peer-to-peer native.

Architectures

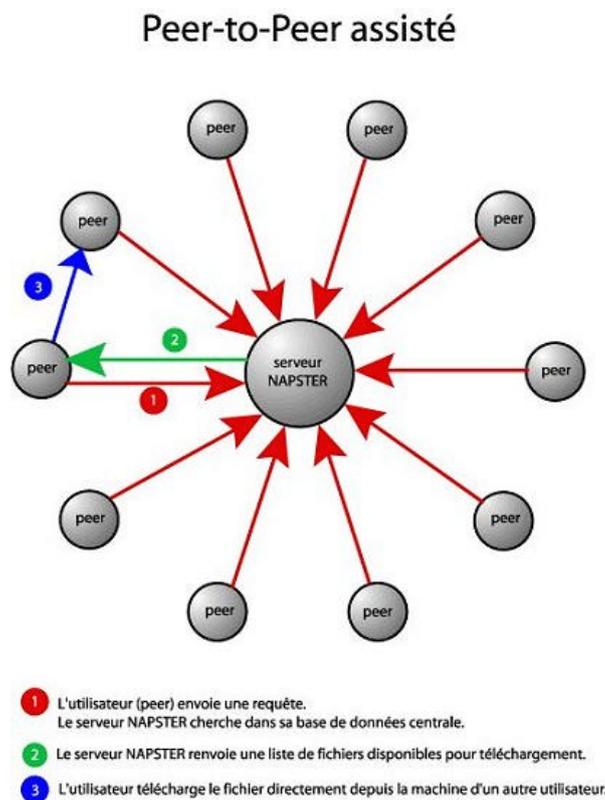
Derrière la définition largement répandue de peer-to-peer coexistent deux modèles d'architecture.

D'un côté, le peer-to-peer assisté, construit autour d'un serveur central. De l'autre, le peer-to-peer décentralisé qui repose sur des nœuds de réseaux.

Sur ces deux modèles se greffent un certain nombre d'optimisations et de variantes topologiques dont les applications peer-to-peer peuvent tirer parti en fonction de leurs besoins.

Le peer-to-peer assisté

Dans ce modèle, on emploie un serveur central qui permet d'indexer les peers connectés au réseau et les ressources disponibles. Cet index est utile pour fournir aux autres peers une cartographie des ressources disponibles sur le réseau. Le serveur assure le service de localisation des ressources et les peers peuvent alors communiquer directement entre eux sans l'assistance du serveur.



Un peer A se connecte au réseau en annonçant à un serveur quelles ressources il met à disposition. Le serveur maintient un index des ressources disponibles sur le réseau à chaque instant. Si le peer A a besoin d'une ressource, il interroge le serveur qui localise la ressource sur un autre peer B connecté au réseau. Le serveur n'intervient plus à ce moment là et le transfert de la ressource s'effectue directement entre A et B.

Certains systèmes reposent complètement sur un serveur, comme c'est le cas de SETI@Home, d'autres utilisent l'assistance d'un serveur pour réaliser une partie de leurs tâches. Le service de recherche de Napster est centralisé alors que le service de téléchargement est décentralisé.

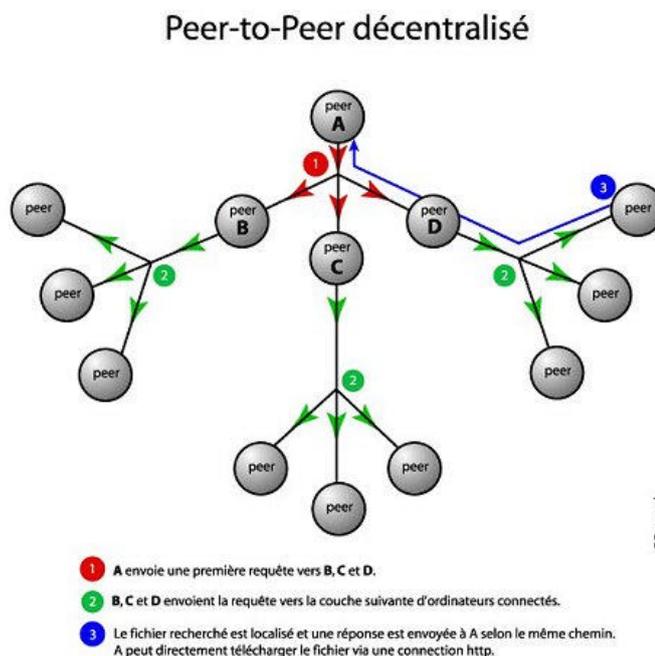
L'un des principaux avantages de ce modèle est l'utilisation d'un index central qui permet de localiser les ressources rapidement et efficacement. Cet index doit être mis à jour en permanence en fonction de l'activité du réseau. Les peers peuvent avoir une connexion intermittente sans que cela ne nuise à l'intégrité du système.

Principal problème, en revanche, ce type de système n'autorise qu'un seul point d'entrée sur le réseau, et n'est pas à l'abri d'une coupure¹ de serveur susceptible de bloquer toute l'application. Des architectures hybrides (voir 1.3.3) permettent de contourner cette faiblesse.

Le peer-to-peer décentralisé

Le deuxième modèle repose entièrement sur les nœuds du réseau plutôt que sur un serveur. Chacun des nœuds met à disposition de l'ensemble du réseau un certain nombre de ressources. Le service de localisation comme le routage des requêtes est pris en charge par l'ensemble du réseau au lieu de faire appel à un serveur.

Ce modèle rend l'accès au réseau plus difficile à utiliser que le premier, car les nœuds qui désirent rejoindre le réseau doivent trouver un nœud de départ pour pouvoir accéder au service.



Le principe est le suivant : un peer A, équipé d'un programme spécifique, se connecte à un peer B, lui aussi équipé de ce programme. A lui annonce ainsi sa présence sur le réseau. B relaie cette information à tous les peers auquel il est connecté. Ceux-ci signifient alors leur présence à A et relaient l'information à leur tour aux peers auxquels ils sont connectés, et ainsi de suite jusqu'à « l'horizon visible³ » du peer A.

Une fois que A a annoncé sa présence aux autres membres du réseau de peers, il peut s'appuyer sur les peers qu'il connaît pour effectuer ses recherches.

Pour obtenir une ressource, A lance une requête vers certains peers du réseau qui la relaient vers les peers auxquels ils sont connectés, qui eux même la transmettent. Si l'un des peers dispose d'une ressource qui pourrait convenir à A, il transmet l'information vers A. Ce dernier pourra ainsi ouvrir une connexion directe vers cet ordinateur et obtenir la ressource.

Ce modèle, en étant décentralisé, est beaucoup plus robuste qu'un modèle centralisé puisqu'il n'est pas dépendant d'un serveur, qui est le point de défaillance potentiel d'un réseau centralisé. Il tire partie de l'intermittence des connexions des nSuds car si l'un des nSuds se déconnecte du réseau, la requête pourra être poursuivie vers les autres ordinateurs connectés.

En revanche, en raison de la façon dont sont transmises les requêtes (broadcast), la bande passante nécessaire pour chaque requête croît exponentiellement quand le nombre de peers croît linéairement.

De plus, ce type de mécanisme est très facilement victime d'activités malicieuses. Des membres malintentionnés peuvent envoyer en grande quantité des requêtes erronées qui produisent une lourde charge sur le réseau, réduisant ainsi son efficacité.

Alternatives et améliorations

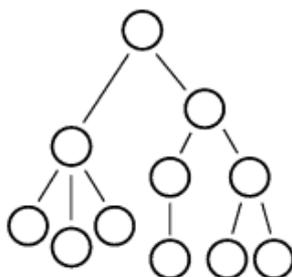
Il existe un certain nombre de techniques et d'architectures alternatives qui tentent d'apporter des solutions aux problèmes que connaissent ces deux modèles.

Architectures hybrides

Le modèle assisté apporte certainement la meilleure des solutions en terme de rapidité de résolution des requêtes. Le modèle décentralisé, lui, apporte la robustesse mais doit consommer beaucoup de ressources pour acheminer les requêtes. Des modèles hybrides font leur apparition afin de tirer parti des points forts des deux modèles de base. La gamme des topologies hybrides imaginable est assez vaste et permet de répondre à de nombreuses problématiques. C'est notamment le cas dans un environnement d'entreprise où des points d'entrée centralisés sont nécessaires pour des raisons de maintenance et de sécurité.

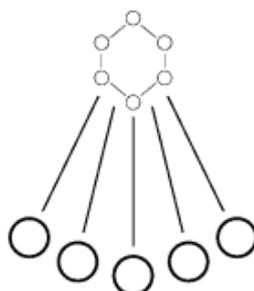
La combinaison de plusieurs services peut nécessiter plusieurs serveurs organisés de façon hiérarchique. Cette topologie peut être nécessaire lorsqu'il existe une chaîne d'actions bien déterminée entre les services. Ces systèmes offrent généralement une grande possibilité d'extension comme l'a démontré le DNS conçu il y a 15 ans et qui supporte aisément les millions de nSuds que comporte l'Internet aujourd'hui. Ce modèle est une alternative aux systèmes centralisés et permet d'intégrer facilement de nombreux services. La racine de la structure hiérarchique représente cependant un point potentiel de défaillance et la sécurisation du système est difficile.

Architecture Hierarchique



L'organisation de grappes de serveurs en anneau est largement utilisée dans le web. L'anneau permet de réaliser la répartition de charge et dilue le risque de défaillance localisée. En connectant le réseau de peers à ces anneaux, on bénéficie de la simplicité d'un système assisté avec la robustesse d'un système décentralisé.

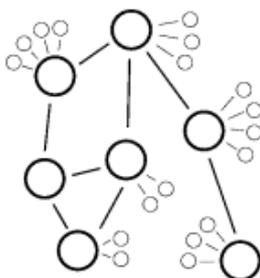
Architecture en Anneau



Une nouvelle vague de systèmes peer-to-peer utilise une architecture combinant des systèmes centralisés au sein de systèmes décentralisés. Ce type de réseau remplace les serveurs dédiés centraux qui réalisent l'indexation du contenu par un grand nombre de « super-nœuds » disposant de bande passante et de puissance de calcul au dessus de la moyenne.

Les super-nœuds peuvent aussi être utilisés pour relayer des requêtes lorsque les ressources trouvées sont insuffisantes. Ce principe est populaire et améliore grandement les performances des réseaux décentralisés. Il a notamment été intégré au réseau Gnutella et dans le système de FastTrack [FSTRK]. Les peers sont reliés à au moins un super-nœud qui n'est pas isolé mais intégré dans le réseau décentralisé.

Architecture avec super-nœuds



C est un modèle qui se rapproche de celui de l email. Les clients mail sont rattachés à un ou plusieurs serveurs mail qui se partagent la tâche de routage des emails. La croissance phénoménale d Internet depuis la conception initiale du système d email n a en rien dégradé ses performances.

Un tel système est donc capable d équilibrer la charge sur l ensemble du réseau, et de limiter le risque d interruption de service puisque si un super-noeud n est pas disponible, les autres réaliseront ses tâches.

Cependant dans les systèmes actuels le choix de fonctionner en tant que super-noeud est une option laissée à la générosité de chaque utilisateur.

Optimisations

Un certain nombre d optimisations du réseau peut être intégré aux systèmes peer-to- peer [DRDLPBN].

Une des pratiques les plus néfastes aux réseaux peer-to-peer d intérêt commun est celle du freeloading. Celle-ci consiste à utiliser les ressources du réseau sans partager les siennes. La charge du réseau s en trouve augmentée sans que sa valeur n augmente.

La lutte contre le freeloading est un point crucial que certains systèmes comme Mojo Nation [MJNTN] combattent en attribuant plus de crédits aux peers qui favorisent le fonctionnement du réseau.

Les systèmes actuels utilisent généralement TCP comme protocole de transport. De par sa nature non connectée, il est difficile de reprendre une connexion perdue entre deux sessions d utilisation du réseau.

Rendre les connexions persistantes permettrait de maintenir un historique des sessions et de construire des profils locaux sur les autres peers. Ces profils seraient alors utilisés pour déterminer la valeur d un peer ainsi que pour monitorer les ressources qu il a utilisées.

Au sein d une communauté, les membres n ont pas tous les mêmes affinités. Etablir des profils de peer donne l avantage de permettre au réseau de se réorganiser en groupes d intérêts différents. Par exemple un amateur de science- fiction partagera des ressources qui ne conviendront pas nécessairement à un amateur de littérature du 18ème siècle.

Favoriser les requêtes vers les peers disposant, d après leurs profils, de ressources plus intéressantes permet de rendre les recherches plus pertinentes. Des groupes de peers ayant des intérêts similaires s organiseront alors spontanément comme ils le feraient dans le monde réel. C est une approche sociale que préconise le réseau ALPINE [ALPN].

Dans un environnement aussi hétérogène que l Internet, les peers disposant d une faible connexion sont souvent des goulets d étranglement. Pour contrer ce problème, des profils de connexions peuvent être intégrés au système. Un peer disposant d une connexion à grande bande passante est favorisé pour fournir la ressource. Les peer disposant d une connexion plus faible sont alors repoussés en bordure du réseau, évitant ainsi sa congestion, tout en bénéficiant d une qualité de service similaire.

Une extension de ce principe a été introduite par clip2 sur le réseau Gnutella. Il s'agit d'utiliser les peers disposant d'une grande bande passante comme proxies pour les peers plus lents.

Certains réseaux comme Freenet [FRNT], utilisent des identifiants obtenus par hashage de chaînes plutôt que des méta-données. L'utilisation de catalogues et de méta-index permet d'augmenter la pertinence et la rapidité des requêtes. Un projet nommé Espra [ESPR] stocke des catalogues à l'intérieur de Freenet. Le catalogue fournit une description pour chaque identifiant (hash key). Mais ces méta-index restent difficiles à maintenir sur l'ensemble du réseau.

D'autres optimisations portent sur la diminution de la taille des paquets internes au protocole, notamment par l'utilisation d'UDP. Ces paquets forment l'essentiel du trafic dans des protocoles de flood broadcast comme Gnutella. Rendre le protocole de communication compact et léger diminue la bande passante gaspillée.

Les systèmes d'échange de fichiers sont particulièrement concernés par la taille des paquets puisque des gigaoctets de données sont transportés en permanence sur le réseau. Aussi, une variante, notamment utilisée par Mojo Nation la technologie SwarmCast de OpenCola [OPCL], consiste à découper les fichiers partagés en petits morceaux. Le système permet alors d'agréger la bande passante de nombreux peers pour transférer ces portions de fichier. Le résultat est une grande rapidité de téléchargement et une fluidification des points de congestion.

Permettre à chaque peer de contrôler les peers avec qui il communique, en terme de bande passante consommée, et d'utilisation du réseau fournit la possibilité de résister aux abus de réseau, d'allouer de la bande passante en fonction des préférences de l'utilisateur, et aussi d'optimiser le « processus de découverte du réseau » (network discovery).

Quel modèle adopter ?

Le peer-to-peer n'est donc pas limité à un seul concept et se prête à un grand nombre d'applications [voir 2.1].

Il n'y a pas de véritable modèle gagnant [MINAR0102]. Chaque modèle offre des caractéristiques, apportant des améliorations significatives selon les besoins spécifiques de chaque application. Les concepteurs de systèmes doivent évaluer les points critiques et adopter les topologies appropriées. Il est important de considérer que nous ne sommes pas limités aux deux modèles de base. On se rend compte du grand potentiel des systèmes peer-to-peer, notamment en les combinant avec les systèmes classiques.

Si la combinaison des deux modèles ouvre de nouveaux horizons, on peut se demander si l'avenir promet au peer-to-peer le même succès économique qu'a connu le client/serveur.

Chapitre 2 - Enjeux économiques

Dans ce chapitre nous nous éloignons des considérations purement techniques pour essayer de déterminer les enjeux économiques du peer-to-peer. Nous présenterons dans un premier temps un ensemble de champs d'application présents et futurs pour le peer-to-peer. Ensuite, nous tenterons de déterminer sous quelles conditions ce concept pourra être adopté par les entreprises, pour finalement évaluer son potentiel pour les réseaux de distribution de contenu multimédia ainsi que sa valeur pour le marketing.

Les applications

Le poids d'une technologie ou d'un paradigme peut se mesurer par rapport au nombre d'applications y faisant appel. Nous donnons un aperçu des idées qui ont cours autour du peer-to-peer. Nous verrons que bien que le concept date, les communautés d'intérêt l'ont redécouvert récemment. Certains spécialistes y voient désormais un grand potentiel qui pourrait bien redéfinir l'usage fonctionnel des réseaux.

Le calcul distribué

Les communautés scientifiques ont toujours nécessité de grandes capacités de calcul pour leurs travaux de recherche. Citons, par exemple, la recherche des nombres premiers de Mersenne.

Les solutions proposées sont généralement d'énormes supercalculateurs tel que l'IBM ASCI White fournissant une capacité de calcul de 12 téraflops. En comparaison, un simple PC à 1 gigahertz fournit un gigaflop, soit mille fois moins. Ces calculateurs sont rares et non destinés aux départements de recherche modestes ou peu populaires.

La naissance de l'idée

Contrairement à ce que l'on croit, les scientifiques ont imaginé une ressource disponible permettant d'assurer largement leurs besoins de calcul. Ce qu'un seul supercalculateur réalise généralement est simplement le découpage d'un gros calcul en portions plus petites, calculées en parallèle. En somme, un supercalculateur n'est qu'un effort collectif de plusieurs petits calculateurs.

Le calcul distribué est donc né de la constatation que la mise en commun des cycles des processeurs peut être réalisée à travers un réseau. Un simple logiciel installé sur des machines clientes, permet de collecter les cycles des processeurs inutilisés pour des tâches de calcul intensives.

Et, à l'évidence, l'Internet en est le terrain le plus propice. Aujourd'hui c'est une manne de 10 milliards de mégahertz et 10 000 téraoctets de stockage qu'offre le monde des PC connectés à l'Internet [GENEBR3].

Le calcul distribué n'est donc pas à proprement parler une technique peer-to-peer puisqu'il fonctionne sur un mode client/serveur plutôt classique, mais certains le justifient en temps que telle au sens où il s'agit de la mise en commun et l'exploitation de ressources (les cycles CPU) d'un groupe.

SETI@Home

L'un des premiers projets scientifiques à bénéficier du calcul distribué par les machines personnelles des internautes a été le projet SETI (Search for Extraterrestrial Intelligence). L'existence de formes de vies intelligentes dans l'univers a alimenté depuis longtemps les fantasmes des auteurs de science-fiction et a fini par devenir une réalité scientifique depuis 1959 avec la publication de l'article « Searching for Interstellar Communication » par les docteurs G. Cocconi et P. Morrison dans la revue Nature.

Depuis, de nombreuses initiatives internationales ont vu le jour, notamment la NASA qui lança en 1992 un programme SETI intensif d'écoute d'émission radio dans l'espace dans l'espoir de trouver des séquences redondantes non aléatoires, signe d'une émission « intelligente ». Ce programme a été abandonné un an plus tard. Mais de nombreux enthousiastes n'ont pas voulu en rester là et en 1997 le projet SETI@Home [SETI] a vu le jour à l'

UC Berkeley. L'initiative consiste à utiliser les capacités de calculs des machines personnelles volontaires pour analyser les énormes quantités de données radio qui sont enregistrées par le télescope Arecibo à Porto Rico.

De part son image de pionnière, l'opération fut très rapidement un succès puisque ce projet a réussi à rassembler 2,7 millions d'utilisateurs dans 226 pays. Grâce à cette popularité, le projet a accumulé 500 000 ans de temps de calcul. La capacité de calcul de la « flotte » de SETI@Home est de 25 téraflops, ce qui représente deux fois la vitesse du plus rapide des supercalculateurs dans le monde (L'ASCI White d'IBM). L'activité a d'ores et déjà permis d'analyser 45 téraoctets de données, sans résultat pour l'instant. Mais peu importe, SETI@home est déjà une réussite car il a confirmé une toute nouvelle façon de faire de la science en utilisant le calcul distribué sur internet.

Des applications commerciales

La recherche technologique a bien compris l'intérêt du calcul distribué. Le leader mondial des microprocesseurs, Intel, utilise les cycles CPU de plus de 10 000 PCs, dans 25 pays, pour conduire des simulations de conception de ses nouvelles puces dans une application propriétaire (NetBatch). Cette même infrastructure distribuée est aussi utilisée pour tester des compilateurs [ROBWIT0801].

En plus de leurs activités philanthropiques, des sociétés spécialisées utilisent leur « capital CPU » pour des projets commerciaux. C'est notamment le cas de sociétés comme Parabon [PARA] ou Entropia [ENTROP] qui proposent leurs plateformes de calcul distribué à leurs clients afin de réaliser leurs gros calculs moyennant finance.

Des applications ont déjà été trouvées dans les biotechnologies, les simulations financières et les tests de performance de serveurs.

La recherche médicale

Le Intel Philanthropic peer-to-peer Program [INTLPP], démontre les capacités du calcul distribué à accélérer la recherche médicale. Ce projet, né de l'expérience d'Intel (voir ci-dessus), sponsorise les activités de recherche médicales portant sur la maladie d'Alzheimer à l'université de Stanford. Intel sponsorise aussi la société United Devices [UD] qui propose sa plateforme de calcul distribué pour la recherche contre le cancer ou, plus récemment, l'anthrax.

Mais il existe aujourd'hui de nombreux projets similaires, notamment pour la recherche contre le SIDA ou dans le cadre du décryptage du génome.

Les communautés

Après avoir partagé des cycles CPU, il ne fallait qu'un pas pour généraliser l'idée au partage de fichiers.

L'épopée Napster

L'année 2000 a connu une aventure dont l'ampleur médiatique a été telle que certains l'associent ; à tort, à la naissance du peer-to-peer : l'épopée Napster.

Plutôt que de mettre en place de larges serveurs qui hébergent l'information, Napster repose sur la communication entre les ordinateurs personnels des membres de la communauté Napster. Cette information est donc distribuée à travers de l'Internet, autorisant une profondeur virtuellement illimitée. [SFT]

En permettant à des particuliers de s'échanger des titres musicaux librement, son programme, Napster, a créé la plus grande collection d'enregistrements pirates dans le monde. En quelques mois le succès est faramineux, plusieurs centaines de milliers de personnes s'enthousiasment pour ce programme générant une activité parfaitement illégale : la distribution publique de musique sans s'acquitter des droits d'auteur. A la fin de l'année 2000, environ 25 millions d'utilisateurs utilisent ce logiciel et le site Napster est dans le Top 50 des sites américains les plus consultés.

La réponse de l'industrie de la musique ne se fait pas attendre et la société est poursuivie en justice par des poids lourds comme Universal, Sony, BMG et même certains artistes. Condamnée par la justice fédérale américaine à supprimer de son système l'ensemble des titres bénéficiant de copyright, Napster se voit contraint d'adopter des partenariats pour rendre les téléchargements payants.

Indépendamment de ses problèmes avec la justice américaine, Napster a tout de même amené un certain nombre d'idées intéressantes. Malgré son échec, la société peut au moins se targuer d'avoir mis à jour aux yeux du grand public et du monde des affaires l'extraordinaire potentiel du concept de peer-to-peer.

L'arrivée des successeurs

Les communautés d'intérêt s'emparent alors du concept d'échange de fichiers et on voit apparaître une myriade de systèmes peer-to-peer avec des philosophies différentes mais un objectif commun : faciliter l'échange de données.

Tirant la leçon des problèmes judiciaires de Napster, les communautés optent pour des architectures totalement décentralisées et plusieurs réseaux voient le jour avec la

ferme intention d'être inarrêtables. Pour cela, les spécifications de protocoles sont librement diffusées et la décentralisation est le mot d'ordre. Il existe une multitude de programmes d'échange de fichiers voués aux communautés. Certains sont généralistes et d'autres spécialisés dans certains types de fichiers. Les présenter exhaustivement serait fastidieux, redondant et hors du cadre de ce document, aussi nous allons étudier quelques applications phare qui couvrent l'essentiel de la problématique.

En juillet 1999, Ian Clarke lance le projet Freenet [FRNT] après la publication d'un mémoire [CLARK] à l'université d'Edimbourg. Il s'agit d'un réseau peer-to-peer dont l'objectif est de garantir l'anonymat des publications sur internet. Il assure la protection des documents contre les altérations en utilisant un système de cryptage fort. L'idée principale est de créer un index décentralisé des documents publiés afin de permettre une recherche rapide sur le réseau. Il fait appel à la réplication pour rapprocher l'information de ses demandeurs. Le réseau se réorganise donc en permanence en fonction de l'activité des membres.

Freenet dénonce les problèmes de liberté et de censure sur Internet et se présente comme un système ouvert et démocratique, qui ne peut être contrôlé par qui que ce soit. Cette philosophie a de quoi séduire les défenseurs de l'anonymat et des libertés d'expression sur Internet, pourtant la recherche d'information sur le réseau est pénalisée par un certain nombre de facteurs :

1. La taille du réseau est limitée et les nSuds mettent longtemps (plusieurs jours) avant d'être pleinement actifs

sur le réseau.

1. Les recherches ne fonctionnent pas sur le principe des méta-données mais avec un système d'identifiants. Il faut connaître l'identifiant pour accéder à un document.
1. Le protocole est en développement intensif et les versions ne sont pas toujours compatibles entre-elles

Gnutella est un protocole de partage de fichiers peer-to-peer initié par Nullsoft [NLSFT], les auteurs du plus populaire des logiciels pour jouer des mp3 : Winamp. Après la première version bêta en mars 1999, AOL, la compagnie qui a racheté Nullsoft en juin 2000, a obligé la fermeture du site officiel et le développement a dû être interrompu.

Le protocole a cependant rapidement été redécouvert (reverse-engineered) et rendu public. Rapidement une grande quantité de clones [GTLM] est apparue. Malgré le soutien de la communauté open source, le projet a échoué à fédérer et coordonner ses développements et le protocole souffre de certaines faiblesses dans sa conception initiale.

Sa technique de diffusion des requêtes repose sur le flooding, c'est à dire que chaque requête est passée à quelques nœuds voisins qui eux-mêmes propagent la requête à leurs voisins jusqu'à expiration d'un délai fixé par l'émetteur de la requête. Cette caractéristique est une véritable plaie pour le réseau comme l'on mis en évidence deux chercheurs du Xerox Palo Alto Research Center dans une étude approfondie [AH1000] du protocole.

De plus, Gnutella n'offre aucune garantie quant au contenu des fichiers échangés, pas plus qu'il ne garantit l'anonymat des utilisateurs. Ces deux caractéristiques sont pourtant des valeurs fortes dans les communautés de partage de fichier.

Beaucoup critiqué pour ses faiblesses en terme de conception, Gnutella a tout de même grandement profité de la chute de Napster, ce qui a contribué à améliorer les performances globales du réseau.

Mais le véritable successeur de Napster, en terme de popularité se nomme Morpheus [MCTY]. Lancé en avril 2001, ce logiciel de partage de fichiers a connu une croissance extrêmement rapide.

De plus, Morpheus partage son réseau avec Kazaa [KZA] et Grokster [GRKSTR] deux autres clones qui utilisent la même technologie propriétaire développée par la société néerlandaise FastTrack [FSTRK].

Le système possède toutes les caractéristiques d'un réseau purement décentralisé très optimisé, mais dans les faits, le service d'accès au réseau est géré de manière centralisée. Le service peut donc facilement être interrompu sur décision judiciaire, rendant l'accès au réseau impossible.

Un autre projet intéressant se nomme Mojo Nation [MJNTN], il ne s'agit pas d'un simple système de partage de fichiers mais plutôt d'un prototype de réseau de distribution. Cette plateforme est basée sur le principe de la « swarm distribution » (voir 1.3.3) pour assurer la rapidité de distribution. Toutes les données transférées sur le système sont cryptées avec le codage RSA à clé publique afin d'assurer leur authenticité.

L'originalité du système est son système de « monnaie numérique » dénommée « Mojo ». Les utilisateurs

gagnent du Mojo quand ils contribuent au réseau en publiant du contenu ou en servant les requêtes des autres. De même, télécharger quelque chose sur le réseau nécessite un paiement en Mojo. Cette approche économique permet de décourager le freeloading et de rémunérer les meilleurs participants en leur donnant un meilleur accès aux ressources.

Le corporate peer-to-peer

Le peer-to-peer ne se limite pas aux applications de partage de fichiers. La possibilité de se dégager des coûts engendrés par les infrastructures centralisées actuelles a de quoi intéresser les entreprises. Une vague de startups s'est donc lancée dans le nouveau créneau du « corporate peer-to-peer ».

Le travail collaboratif

Une des applications les plus en vogue dans le domaine du corporate peer-to-peer est Groove, de la société Groove Networks [GRVNT]. Ce logiciel permet à ses utilisateurs de collaborer en temps réel depuis n'importe où. Groove offre de nombreux services comme la messagerie instantanée, la conférence, le partage de documents, la navigation simultanée, des outils de planification de tâche et de gestion de projet. Mais la liste est longue et d'autres extensions sont en cours de développement afin de couvrir un large éventail d'activités professionnelles. L'objectif est de diminuer les coûts de communication interne, spécialement dans le cas où les collaborateurs sont disséminés à travers le monde.

Groove peut travailler de manière complètement décentralisée, mais généralement ce mode de fonctionnement ne convient qu'aux environnements personnels, ou aux petits réseaux d'entreprise dans lesquels le mécanisme de découverte des nœuds par broadcast reste efficace. Plus généralement, Groove fait appel à une variété de services centralisés transparents pour l'utilisateur.

De plus, Groove est basé sur des technologies XML comme XML-RPC ou SOAP (Simple Object Access Protocol). Cette caractéristique le rend parfaitement compatible avec les web services et Microsoft l'a bien compris puisque la firme de Redmond a récemment acquis près de 20% du capital de Groove Networks. La plateforme .NET de Microsoft supporte aussi le protocole, ce qui rend l'intégration de nouveaux services parfaitement flexible. Groove adopte COM comme modèle de composants ce qui le rend compatible avec la plateforme windows. Les utilisateurs de cette plateforme pourront donc facilement intégrer Groove dans leur environnement de travail sans que cela ne nécessite d'utiliser des outils spécifiques. Groove fournit simplement une plateforme d'échange peer-to-peer destinée à faciliter les échanges entre collaborateurs.

Enfin, Groove intègre un système de sécurité assez évolué, basé sur des mécanismes de clé publique, dans lequel toutes les transmissions sont cryptées. Un point essentiel pour séduire les directions informatiques.

Aujourd'hui seulement disponible pour la plateforme Windows, Groove devrait bientôt être porté sur des plateformes alternatives comme Linux ou MacOS.

Un autre produit peer-to-peer du nom de « Consilient Sitelet Technology » fournit une approche toute nouvelle à la gestion de workflow. Consilient fournit un framework java pour développer des projets qui contrôlent eux-mêmes leur propre workflow. Tout le nécessaire, les documents, l'information de workflow et les outils sont compris dans un paquetage appelé sitelet. Cette sitelet permet à différents collaborateurs de travailler sur un projet simultanément de manière peer-to-peer. Les programmeurs peuvent fixer des contraintes (comme par exemple

dire qu'un responsable financier doit faire calculer un budget à un certain moment ou qu'un cahier des charges doit être transmis au responsable clientèle après validation par le responsable technique) et chacun peut ajouter des tâches.

Le framework est volontairement très général puisqu'il est voué à pouvoir s'adapter au mode de fonctionnement interne de n'importe quelle entreprise. Ceci permet d'obtenir une flexibilité variable adaptable à tout style de gestion de projet.

La sitelet fonctionne comme une sorte d'agent intelligent qui contrôle le workflow. Imaginons une proposition commerciale publiée dans la sitelet. Automatiquement (et éventuellement à la date indiquée), la proposition est rendue accessible au département financier et présente, à la personne qui la reçoit, un document type (spécifique à l'entreprise) qui lui permet d'annoter la proposition. Dans la chronologie du workflow de l'entreprise, le responsable des projets doit ensuite approuver la proposition. La sitelet, se charge alors de présenter la proposition (avec les annotations du département financier) au responsable des projets et permet au manager d'évaluer la proposition puis de la valider.

Le workflow est donc contrôlé automatiquement par la sitelet et ne repose plus sur les individus.

Ces deux exemples démontrent que les technologies peer-to-peer ont le potentiel de grandement améliorer la collaboration au sein d'une entreprise et par là même diminuer les coûts de coordination et de gestion de projets

Autres applications

La collaboration est l'une des applications les plus évidentes du peer-to-peer dans l'entreprise mais il existe de nombreux champs d'application. Voici quelques exemples d'applications dont les entreprises peuvent tirer parti.

Les mécanismes de réplication souvent intégrés aux applications peer-to-peer peuvent aussi diminuer les coûts de maintenance des réseaux locaux internes aux entreprises. Lorsque des mises à jour de sécurité apparaissent, la tâche de mise à jour de l'ensemble des postes peut être prise en charge par une application peer-to-peer sans que les utilisateurs n'aient à s'en soucier.

Une autre problématique pour les entreprises dont l'activité tourne autour de la fourniture de logiciels ou d'applications web sont les tests de performance souvent coûteux et gourmands en termes de ressources. De même, de grandes entreprises comme Boeing ont des besoins en puissance de calcul que même les plus gros calculateurs ne sont pas toujours en mesure de fournir.

Le grid computing (une forme de calcul distribué) représente une solution abordable et de nombreuses entreprises comme Endeavors, Avaki [AVK], Paraborn ou encore Porivo [PRV] fournissent des plateformes de calcul distribué prêtes à l'emploi.

Le peer-to-peer représente un potentiel intéressant pour construire des réseaux de distribution efficaces et peu onéreux. Kontiki [KTK] propose un exemple de réseau peer-to-peer de distribution de contenu (CDN) multimédia au service des entreprises désireuses de délivrer un contenu riche aux utilisateurs finaux.

Si l'idée et la valeur du peer-to-peer a encore un long chemin à parcourir avant de convaincre les responsables informatique, les grands éditeurs proposent d'ores et déjà des plateformes comme JXTA de Sun ou .NET de

Microsoft fournissant les composants et services de base prêts à intégrer la logique d'entreprise.

La libération du potentiel du peer-to-peer augmente le rendement, réduit les coûts d'infrastructure et contribue ainsi à l'accroissement de la rentabilité.

Des applications militaires

Pendant des années, les militaires ont utilisé de grands systèmes client/serveur pour construire des environnements d'entraînement des soldats dans des conditions de simulation. Aujourd'hui les militaires américains envisagent les technologies peer-to-peer pour éviter la vulnérabilité des serveurs centralisés [KOM0801].

L'enjeu est de construire des systèmes de simulation mobiles fournissant aux soldats ce que les militaires nomment « simulation on demand ». Lorsqu'un soldat a besoin d'apprendre, tester, planifier quelque chose, il peut l'avoir instantanément.

Des systèmes d'entraînement tactique au combat rapproché existent déjà, mais il s'agit de simulateurs coûteux qui nécessitent d'énormes bases de données qui modélisent des champs de bataille. Ceci est assez limité et l'objectif futur sera de permettre à chaque soldat de transporter l'environnement de simulation avec lui. En somme, les hommes pourront porter des casques fournissant un environnement de réalité augmentée. Ils auront une vraie interaction physique avec la réalité, et seront en mesure de voir des civils, les forces de police ou des ennemis virtuels.

Tous ces soldats seront reliés ensemble par des réseaux locaux mobile et pourront coordonner leurs opérations tactiques grâce à ce genre de dispositifs.

A terme, ces systèmes intégreront, en surimpression sur le dispositif visuel du soldat, des viseurs haute précision et des informations tactiques sur les objets réels comme la distance des bâtiments environnants. Ces informations seront délivrées à l'aide des systèmes GPS différentiels. Les soldats pourront aussi se laisser des informations comme la position des ennemis éventuels ou de munitions.

Le véritable challenge est donc de pouvoir interconnecter en temps réel toutes ces informations tactiques sans avoir recours à des serveurs centralisés. Le peer-to-peer offre une excellente solution en permettant de résoudre ces problèmes de connexion à la volée.

Réseaux de diffusion de contenu

L'une des révolutions, annoncée pour 2001 par les spécialistes, et qui devait transformer radicalement l'expérience des utilisateurs d'Internet est le streaming.

Si la technologie commence à atteindre une certaine maturité, les infrastructures réseaux ne permettent toujours pas aux internautes de profiter de vidéo ou de contenus multimédia riches dans des conditions acceptables. Les vidéos transportées par streaming sont généralement d'une qualité effroyable et la résolution ridiculement petite.

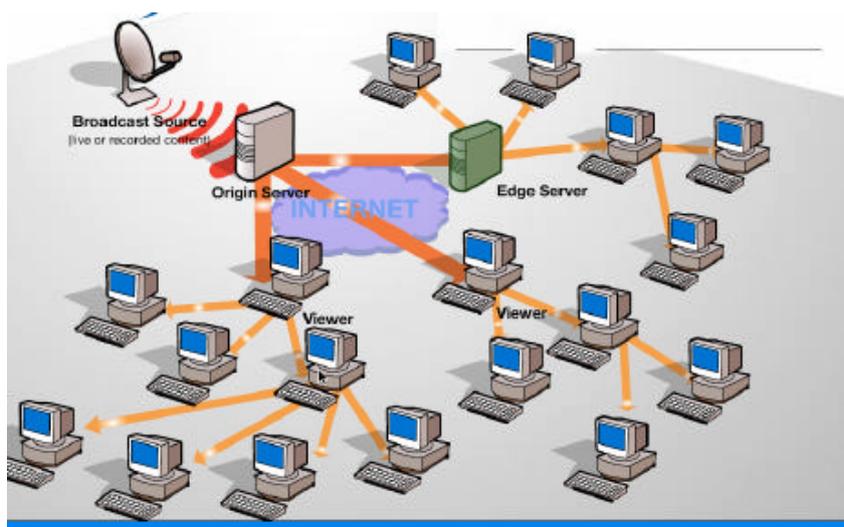
Principe

L'Internet n'a pas été conçu au départ pour garantir de la diffusion en temps réel. De plus les bandes passantes proposées par les fournisseurs d'accès à leurs clients ne permettent pas la diffusion en temps réel de flux vidéo de qualité acceptable.

Les fournisseurs de contenu doivent donc faire appel à des infrastructures spécialisées dans la distribution de contenu (CDN). Ces réseaux utilisent des techniques de cache pour rapprocher physiquement le contenu des utilisateurs.

Le peer-to-peer permet à chaque peer de servir de site miroir pour les autres. Par exemple si l'un des employés d'une entreprise télécharge une vidéo, sa machine peut servir de cache local au sein du réseau local pour les autres peers se trouvant à l'intérieur du LAN. Ceci permet d'éviter de télécharger plusieurs fois la vidéo depuis le serveur distant et économise la bande passante.

Exemple de fonctionnement d'un CDN



Quelques exemples

Il existe plusieurs sociétés, notamment Blue Falcon, travaillant sur des technologies utilisant les techniques de partage de fichier P2P pour diminuer les temps de réponse et les coûts de bande passante.

Kontiki est une jeune société financée en partie par les fondateurs de Netscape qui utilise la technologie Akamai (cache serveur) combinée avec le peer-to-peer (cache client) pour fournir une infrastructure haute performance de distribution de contenu.

EverNet [EVRNT] offre une technologie similaire qui se rapproche du modèle Napster puisqu'elle utilise des

serveurs pour localiser efficacement les contenus demandés sur des peers proches. Pour améliorer les performances, elle fait appel au principe de la swarm distribution.

Les problèmes à résoudre

Cependant l'approche peer-to-peer de distribution comporte quelques problèmes :

La qualité de service peut être difficile à garantir puisque les peers peuvent se déconnecter à tout moment du réseau.

Les firewalls et le NAT représentent souvent une barrière pour les peers se trouvant derrière.

Les fournisseurs d'accès payent cher la bande passante et ce genre d'applications est très gourmand. Ils pourraient vouloir une compensation.

Des questions de sécurité et de droits d'auteur risquent de se poser.

Si le nombre de CDN augmente, et que leurs systèmes ne sont pas capables de travailler ensemble, leur efficacité risque de rester limitée.

Quand peer-to-peer et services web convergent

Deux des thèmes les plus discutés de ce nouveau millénaire ont certainement été les services peer-to-peer et les services web. Même si ces concepts ont des zones de chevauchement, leurs manifestations sont bien différentes. Si l'on étudie les types de problèmes que ces systèmes résolvent et la manière dont ils le font, on peut imaginer que les deux vont converger vers une solution hybride [SCHN0701].

Le modèle web services

Classiquement, le modèle du web service est celui où plusieurs entités distinctes interagissent. Le consommateur de service fait appel à un fournisseur de service qui met à disposition un référentiel d'annuaires permettant au consommateur de localiser les services qu'il désire utiliser. Les web services englobent un grand nombre de champs d'activité comme l'ASP, l'ERP, le SCM, le CRM, &

Topologiquement, les web services utilisent le modèle client/serveur pour la communication entre les entités.

Auparavant très propriétaires, les architectures orientées services s'ouvrent désormais et permettent aux trois entités de communiquer via un ensemble de standards basés sur XML comme, SOAP, UDDI ou WSDL.

Cette volonté permet notamment à des systèmes hérités du passé (legacy systems) de communiquer avec de nouvelles applications et ainsi de participer facilement à une dynamique économique.

Le modèle peer services

Dans le cas du peer-to-peer, les entités décrites ci-dessus coexistent sur chaque poste (peer). Au-delà des simples applications de partage de fichiers, les grands éditeurs (Sun, Microsoft) ont travaillé à fournir des solutions qui correspondent à de réels besoins pour les entreprises.

De plus, de par leur nature décentralisée, les peer services sont capables de contourner les problèmes de non disponibilité dont peuvent souffrir les web services, par des mécanismes de réplication.

Les peer services tendent à utiliser les mêmes standards et sont donc aussi capables d'intégrer les systèmes de définition de règles de gestion comme ebXML ou BizTalk.

La convergence

Peer services et web services opèrent donc sur des bases similaires mais présentent des approches différentes.

La volonté des éditeurs de ne pas manquer le virage technologique actuel, les a conduits à fournir des plateformes ouvertes, autorisant ainsi les deux modèles de coexister et même d'interagir. Le point d'ancrage pour l'interopérabilité est XML qui offre l'énorme avantage d'être utilisable indépendamment du protocole de transport des données.

Les deux approches divergent essentiellement sur le mécanisme de localisation de service. Alors que les peer services utilisent des mécanismes décentralisés, les web services font appel à des mécanismes largement centralisés comme UDDI.

De plus, les processus qui régissent l'activité des entreprises recoupent ceux qui régissent l'activité de leurs employés, ce qui tend à annoncer la convergence. En d'autres termes, la convergence des peer services et des web services permettra aux employés d'utiliser des outils efficaces qui s'intégreront complètement au système d'information de l'entreprise assurant un gain d'efficacité et réduction des coûts.

Convergence entre peer et web services



Le concept est-il économiquement viable ?

Le peer-to-peer se prête donc à de nombreuses activités dans des domaines variés. Pourtant, certains arguent qu'il ne s'agit que d'un jouet technologique porté par une de ces nombreuses vagues d'enthousiasme que l'Internet a connu. Le peer-to-peer sera-t-il capable d'atteindre une maturité et une crédibilité suffisantes pour être adopté par les entreprises ?

Une valeur non marchande

Le terme peer-to-peer est souvent remplacé par l'acronyme P2P par analogie au B2B (Business to Business) et autres B2C. B2B n'est qu'une expression à la mode (buzzword) dans le jargon Internet pour parler de l'e-business ; en d'autres termes cela correspond à l'échange de produits, services ou information entre entreprises sur le réseau mondial.

P2P introduit une confusion et laisse supposer qu'il existe un marché similaire encore inexploité. En réalité un peer représente une entreprise, un utilisateur ou un périphérique connecté au réseau. Plutôt qu'un marché, P2P est un modèle de communication qu'il conviendrait d'opposer à C2S pour « Client to Server ».

Il est donc certain que le peer-to-peer ne va pas révolutionner l'économie. En revanche, il pourrait bien contribuer à offrir de nouvelles solutions technologiques aux acteurs de cette économie.

Le potentiel du peer-to-peer n'est apparu que récemment et les entreprises restent prudentes. Il faut bien comprendre que parler du peer-to-peer comme d'un marché n'a pas de sens [STDG0801] puisqu'il s'agit plus d'une approche technologique que d'une opportunité de business.

Le cabinet Gartner Group prévoit que 30 pourcent des entreprises utiliseront des systèmes peer-to-peer d'ici l'an prochain. En revanche le cabinet Frost & Sullivan se contente d'annoncer que près de 6 millions d'entreprises feront appel au peer-to-peer d'ici 2007. Les avis sont donc partagés ; mais qu'ils soient optimistes ou plus réservés, aucun cabinet d'étude ne prévoit une fin funeste au peer-to-peer.

Pourtant, aujourd'hui, peu d'entreprises positionnées sur le peer-to-peer ont dégagé des revenus consistants. Mais ceci risque de changer si le peer-to-peer est capable de répondre aux exigences des acteurs de l'économie Internet.

Facteurs d'adoption

Quelles sont les exigences pour l'adoption et la diffusion des systèmes peer-to-peer dans l'activité économique des services Internet ?

Des entreprises comme Daimler Chrysler, Intel ou Ford ont commencé à utiliser des systèmes peer-to-peer pour permettre à leurs équipes géographiquement disséminées de collaborer. Le succès d'une innovation exige qu'elle apporte des réponses tangibles à des problèmes technologiques, commerciaux et économiques.

Soutenir les innovations

Il est encore difficile de déterminer avec exactitude les lacunes que l'utilisation du peer-to-peer permettra de

combler, et certains se demandent si le peer-to-peer sera capable de générer des revenus dans un système complètement décentralisé. Pour faciliter l'adoption du peer-to-peer, les entreprises doivent penser clairement et stratégiquement aux façons de transformer son utilisation en profits.

Les difficultés économiques de l'année 2000 auront appris aux entreprises que le constant renouvellement, comme la réinvention des processus de gestion, est une question de survie. C'est pourquoi le maintien des investissements en recherche et développement, ainsi que les expérimentations de marchés, sont un facteur clé dans l'adoption du peer-to-peer.

Les frasques judiciaires de Napster ont laissé dans l'esprit des investisseurs une image négative du peer-to-peer. La réduction du peer-to-peer au concept de Napster limite le pouvoir de persuasion du modèle. Un travail éducatif est donc nécessaire pour convaincre les capitaux risqués de prendre des risques à nouveau.

Standardisation

Les standards sont encore un facteur qui devrait faciliter l'adoption. Les grands comptes de l'industrie participent activement à leur élaboration aux côtés d'entreprises innovantes et d'organismes internationaux. Si cette composition hétérogène des groupes de travail implique une certaine lenteur dans la définition des standards, elle permet de s'assurer de leur compatibilité ainsi que de leur robustesse dans des environnements hétérogènes.

Des réponses à ces questions commencent à apparaître au sein des groupes de travail. Par exemple, le Peer-to-peer Working Group [P2PWG], initié par Intel et comprenant un grand nombre d'acteurs majeurs du secteur s'attache à trouver des solutions communes pour favoriser l'adoption du peer-to-peer.

Ses objectifs primaires sont la standardisation, la coopération mutuelle et le partage de connaissance. Le P2PWG publie de nombreux documents concernant la sécurité, les standards et des études de cas. Les membres du groupe espèrent ainsi résoudre les questions de sécurité, définir des architectures et des API communes, et ouvrir de nouvelles perspectives à l'industrie.

Capitaliser la connaissance

Les concepts du peer-to-peer ne sont certes pas nouveaux mais l'idée de son adoption et de sa diffusion, elle, est plus récente. La prise de conscience du potentiel économique des technologies peer-to-peer enthousiasme les uns, qui y voient d'ores et déjà le modèle gagnant de demain, alors que les autres y voient juste une mode passagère sans véritable avenir.

Il convient en réalité de se poser des questions essentielles sur les exigences auxquelles le modèle doit satisfaire pour qu'il puisse libérer son potentiel. Comment se modèle va-t-il intégrer à celui existant ? Le peer-to-peer saura-t-il tirer parti des systèmes existants ? Comment le peer-to-peer va-t-il affecter le volume des transactions ?

Investissements

A ce jour, environ 560 millions de dollars ont été investis dans des sociétés peer-to-peer. Mais la tendance est récente puisque 84 pourcent de ces investissements ont été réalisés durant les dix huit derniers mois [STDG0801].

Microsoft a augmenté la crédibilité du peer-to-peer en investissant 51 millions de dollars dans la société Groove Networks. En acquérant 20 pourcent du capital, le géant de Redmond compte ainsi intégrer Groove à sa stratégie Internet puisque le logiciel offre une compatibilité avec la plateforme .NET. Microsoft n engage pas son argent dans une startup immature puisque plus de 100 partenaires ont d ores et déjà investi dans la société fondée par Ray Ozzie, le créateur de Lotus Notes. La société a déjà réussi à lever un total de 117 millions de dollars.

De même, Sun semble aussi croire au concept du peer-to-peer puisque la société construit sa propre plateforme peer-to-peer portant le nom de JXTA. Les inventeurs du langage Java ont notamment démontré leur confiance dans le peer-to-peer en achetant la technologie de recherche peer-to-peer de la société Infrasearch Inc, en février 2001.

Le grid computing constitue la technologie peer-to-peer dans laquelle les investisseurs discernent le plus clairement un potentiel commercial. Aussi la société Entropia a reçu 30 millions de dollars de Moore Capital Management et REE Ventures en février 2001. United Devices a fait appel à de nombreux fonds de capitaux risqués pour aujourd'hui totaliser 18 millions de dollars.

Malgré tout, les investissements levés par les acteurs du peer-to-peer restent minces. Le fait est que les investisseurs ne voient pas toujours la pertinence des projets peer-to-peer qu'ils reçoivent. L'époque des investissements à tour de bras dans les projets innovants est définitivement terminée, et le coup de semonce de l'année 2000 a recentré les capitaux risqués vers leur objectif premier : la rentabilité.

Les entreprises qui veulent réussir doivent désormais démontrer leur capacité à trouver un modèle de revenu plutôt que d'agiter une innovation technologique prétendument révolutionnaire. Les fournisseurs de solutions peer-to-peer sont, pour la plupart, des petites entreprises qui doivent se doter d'une équipe managériale décente pour rassurer les investisseurs.

Distribution et marketing

Si les entrepreneurs restent perplexes en ce qui concerne la pertinence du peer-to-peer, le marketing, en revanche, voit de nouveaux horizons s'ouvrir. Avec Napster, le peer-to-peer a démontré sa capacité à établir des communautés actives de participants. Les annonceurs peuvent s'appuyer sur des réseaux peer-to-peer pour améliorer leur communication.

Réseaux de distribution

Les systèmes peer-to-peer peuvent ouvrir la voie à de nouveaux modèles de distribution. S'appuyant sur l'expérience de l'e-business, ils court-circuitent certains réseaux de distribution classiques.

Avantages

En plus des facilités qu'apporte le commerce électronique, plusieurs facteurs laissent penser que le peer-to-peer pourrait servir les industries de la musique et du cinéma.

Dans les systèmes peer-to-peer, l'information circule rapidement et de manière efficace. Les contenus les plus prisés sont rapidement répandus alors que les autres disparaissent automatiquement. Un réseau de distribution

classique doit supporter des coûts importants pour alimenter les points de distribution. Avec le peer-to-peer, c'est le consommateur qui véhicule les contenus rendant ces coûts pratiquement nuls.

Les réseaux peer-to-peer s'organisent spontanément en communautés aux intérêts convergents, et ce quelle que soit la répartition géographique de ses membres. Cette caractéristique permet d'éviter les coûts engendrés par des études de marchés locaux. Il n'est plus la peine de se demander si la nouvelle star du hip-hop marseillais va conquérir le marché italien. Les transalpins amateurs iront d'eux-mêmes rechercher le dernier album de leur star favorite.

Les réseaux de distribution supportent mal que la chaîne d'approvisionnement soit brisée. Or les réseaux peer-to-peer comblent assez bien cette faiblesse. Si un contenu ne peut pas être obtenu par un chemin un autre existe certainement.

L initiative BMG

Les majors de l'industrie musicale américaine ont pris en grippe le phénomène Napster et enrichissent leurs avocats dans de nombreuses actions légales contre les

systèmes d'échange de fichier. Le groupe allemand Bertelsmann ne montre pas la même hostilité puisqu'il a conclu des accords avec Napster en octobre 2000.

Les deux sociétés ont mis en place un système de souscription payant donnant accès aux utilisateurs au catalogue d'artiste de Bertelsmann. En cherchant à tirer profit de la manne de mélomanes que Napster a su drainer, plutôt que de mener une guerre judiciaire, Bertelsmann tente de tirer parti du modèle de distribution en assurant la juste compensation des artistes.

Si cette initiative semble aller dans le bon sens, elle devra cependant faire ses preuves en terme de rentabilité et convaincre les utilisateurs.

Le streaming comme alternative

Si les systèmes d'échange de fichiers devaient échouer à fournir des infrastructures correctes assurant à la fois une distribution efficace et une rétribution des artistes, le streaming apporterait peut-être une alternative intéressante. En ne permettant pas à l'utilisateur final de stocker le flux multimédia qu'il reçoit, le streaming apporte un certain contrôle sur la distribution du contenu.

Désormais, certains CDN offrent l'opportunité de profiter du peer-to-peer pour la diffusion de flux en streaming [voir 2.1.5]. La combinaison d'un système de partage de fichier du type Napster avec le streaming semble aussi être une alternative intéressante pour promouvoir des artistes sans répandre des copies illicites sur le réseau.

Malgré tout, aucun modèle de distribution ne sera acceptable avant d'avoir trouvé des solutions pour le problème d'acquiescement des droits d'auteur. De plus, si un utilisateur paye pour obtenir un contenu, il est en droit d'exiger qu'on lui assure que ce qu'il reçoit est authentique.

Affiliation

L'Internet fournit un canal où l'utilisateur final contrôle la profondeur et la durée de l'expérience interactive [MAN0601]. Le marketing peut exploiter ces caractéristiques. L'un des moyens imaginés est d'utiliser les espaces collaboratifs du peer-to-peer pour permettre l'extension d'une marque en ligne en « connectant » les clients d'une marque à ceux d'une autre.

La notion de communauté qui entoure le peer-to-peer apporte certains avantages au marketing :

1. Tirer pleinement parti de la richesse des formats multimédia pour communiquer plus efficacement sans être cantonné aux formats traditionnels (de simples bannières publicitaires clignotantes..).
1. Permettre au contenu de s'organiser automatiquement. Les contenus les plus intéressants sont diffusés rapidement au sein de la communauté si ils sont populaires.
1. Permettre une communication plus ciblée au sein de communautés dont on connaît mieux le comportement.

Certaines études ont montré que les gens sont très exposés à la publicité et s'en protègent en filtrant massivement les messages auxquels ils sont exposés. En revanche ils écoutent leurs amis [ROSEN01].

Le marketing viral peut tirer parti de l'esprit communautaire pour mettre en place des systèmes d'affiliation ou de recommandation par bouche à oreille. Les utilisateurs d'une même communauté ont généralement des centres d'intérêt connexes.

Il serait ainsi aisé d'imaginer des systèmes dans lesquels les utilisateurs toucheraient une compensation en conseillant les autres membres sur des produits ou des artistes qui pourraient les intéresser.

Dans cette optique, le consommateur n'est plus une cible lointaine cachée derrière un ordinateur dont on ignore s'il tient cas de l'information qu'on lui livre. Il devient un participant actif et critique. Cette vision ouvre une nouvelle discussion sur l'impact social du peer-to-peer.

Spamming

Origine du terme

Le spamming est une pratique qui trouve son origine en 1978 lorsqu'un employé de DEC tenta d'envoyer email à l'ensemble des adresses de la côte ouest des Etats-Unis en utilisant la liste d'adresses d'Arpanet pour inviter les gens à une présentation de la nouvelle machine DEC-20.

Le terme devint populaire en avril 1994, lorsque deux avocats de Phoenix, nommés Canter et Siegel, postèrent un message pour annoncer leurs services dans le cadre de la loterie américaine pour la green card. Les gens ont rapidement adopté le terme « spam ». Depuis, le terme s'est généralisé pour désigner des envois non sollicités sur Internet.

Les Napster Parasites

Les réseaux peer-to-peer d'échange de fichier offrent un terrain propice à ce genre d'activités. Ils n'offrent plus le

contenu de leur disque dur aux autres membres du réseau mais aussi à des entreprises qui utilisent le réseau comme un outil marketing qui tentent de vendre plus efficacement et plus rapidement de la musique. Ces sociétés sont dénommées les Napster Parasites [BROWN0201].

La société BigChampagne [BCPGN] n'hésite pas à s'introduire dans la communauté de Napster, à des fins commerciales. La société utilise une technique qui consiste à parcourir les disques des utilisateurs du logiciel dans le but d'établir les goûts musicaux de l'utilisateur. Leur système est alors capable d'envoyer des messages publicitaires instantanés ciblés (et non sollicités) pour inciter l'utilisateur à acheter de la musique similaire [MANCO0201].

Et la technique semble fonctionner. La société a expérimenté son système avec l'artiste Glen Phillips qui a accepté de se prêter au jeu. BigChampagne a pu localiser 54 000 utilisateurs de Napster possédant des chansons de l'artiste. Parmi ceux-ci, 22 000 sont allés visiter le site de l'artiste suite à la réception du message. De plus, 90 pourcent de ces utilisateurs ont fourni leur adresse e-mail sur le site.

Si BigChampagne réussit à convaincre les majors de l'industrie musicale que son système peut être profitable à la distribution musicale, les utilisateurs de Napster pourraient bientôt être inondés de messages. Pour éviter ce phénomène, BigChampagne assure qu'elle garde une trace des utilisateurs ayant déjà reçu les messages publicitaires pour éviter d'envoyer le même message plusieurs fois. Si la société prétend de cette façon protéger la vie privée des utilisateurs, d'autres sociétés pourraient être moins scrupuleuses.

Mais, comme l'avoue le fondateur de BigChampagne, Eric Garland, mais ce n'est qu'un début, le véritable défi sera d'être capable d'analyser le comportement des utilisateurs finement.

Les utilisateurs de Napster sont donc scrutés méticuleusement lorsqu'ils utilisent le système. De telles sociétés exercent leur activité en toute légalité, et profitent du silence bienveillant des majors de l'industrie de la musique pour s'introduire sur les disques dur des utilisateurs. Ce superbe système dans lequel tout le monde semble gagnant nous ferait presque oublier qu'il constitue une violation ouverte de la vie privée.

Chapitre 3 - Impact social

Ce dernier volet clos notre étude en tentant de présenter les répercussions sociologiques que le phénomène peer-to-peer implique. Nous considérons tout d'abord l'internaute en son rôle au sein d'Internet ainsi que son comportement. Cette étude s'achève sur des considérations d'ordre moral et éthique liées au phénomène.

Le peer-to-peer et l'internaute

Replacé au cœur du réseau, l'internaute fut le premier à connaître les effets du peer-to-peer de masse avec les systèmes de partage de fichiers. Ceux-ci modifient sensiblement, la donne et mettent un pouvoir certain entre les mains de leurs utilisateurs, générant parfois des dérives nuisibles.

Un statut différent

L'internaute enfin actif

L'internaute est généralement un acteur passif sur Internet. Nombreux sont ceux qui désirent communiquer leurs

idées, leurs passions ou le fruit de leur travail sur Internet. La publication de ses propres Suvres sur le réseau mondial n'est pourtant pas une chose aisée.

Tout d'abord il est impératif de se conformer au standard de publication de pages web : HTML, ce qui a de quoi rebuter les plus enthousiastes. De plus, il faut faire appel à un hébergeur distant qui se réserve un droit de censure. Le système de Freenet, bien que toujours en développement, permettra bientôt à chacun de publier très facilement ses Suvres sans risquer la censure.

En ce qui concerne l'impact, Freenet ne garantit pas la visibilité des Suvres. Il assure simplement qu'un contenu reste présent dans le système tant qu'il est demandé. Au bout d'un certain temps, si un contenu n'est plus demandé, le système l'efface localement.

Imaginons que vous êtes un amateur de cuisine nippone et que vous avez publié une recette de sushis dont vous êtes particulièrement fier. Maintenant, votre recette connaît un grand succès auprès des utilisateurs de Freenet qui se délectent de votre merveilleuse création culinaire. Freenet recopiera votre recette sur les machines de tous les amateurs qui ont recherché votre recette. Lorsque tout le monde aura goûté vos sushis jusqu'à l'écoeurement, votre recette commencera à disparaître des machines des utilisateurs. Elle deviendra alors difficile à trouver, car peu populaire. En revanche, cette disparition n'est pas forcément globale puisque les utilisateurs japonais ne se lassent pas si facilement des sushis. Votre recette aura peut-être disparu de Freenet en Europe mais les utilisateurs japonais pourront toujours se délecter.

Le système harmonise et répartit donc naturellement les ressources sur le réseau, sans intervention externe. Sans faire appel à une quelconque autorité de censure, le système tient compte des centres d'intérêts des utilisateurs en rapprochant les contenus des endroits où ils sont les plus demandés.

Les Suvres publiées par les utilisateurs sont donc primées en fonction de l'intérêt qu'elles ont pour les membres de la communauté. C'est une solution de publication efficace accessible à tous.

Le partage de connaissances

La société Amoweba [AMWB] propose le logiciel Human Links, un outil de recherche unique tirant parti du peer-to-peer. Constatant que les moteurs de recherche classiques ne sont pas capables d'indexer la totalité du web, et fournissent souvent des réponses peu pertinentes, la société essaye d'humaniser la recherche d'information sur Internet.

Les utilisateurs ont la possibilité de partager leurs adresses web favorites avec les utilisateurs ayant les mêmes centres d'intérêt. Cette approche humaine permet d'effectuer des recherches pertinentes non pas dans un index monolithique et impersonnel comme celui d'un moteur de recherche classique.

A mesure que le réseau Human Links s'agrandira, les recherches se feront plus précises. Il s'agit d'un bon moyen de trouver des pages non indexées par les moteurs de recherche et de filtrer les sites sans intérêt. L'objectif n'est pas de pouvoir tout trouver mais plutôt de trouver le meilleur.

Cette application est un très bon exemple de la philosophie d'échange de connaissance d'Internet. L'utilisateur devient réellement acteur et tire partie de la puissance intrinsèque du réseau.

Comportements néfastes

Les systèmes d'échanges de fichiers peer-to-peer peuvent facilement souffrir de l'anarchie générale. L'anonymat aidant, les membres sont facilement tentés d'avoir des comportements malveillants. Sans un minimum de contrôle, on peut voir apparaître des activités comme la diffusion de virus ou le freeloading.

Des vers dans mon Gnutella !

En février 2001, le réseau Gnutella a été infecté par un virus de type worm. Un worm est un virus qui est capable de se répandre de pair en pair sur un réseau. Le worm en question a été appelé Gnutella Mandragore.

Ce worm était capable de monitorer les connexions Internet et répondait à toutes les requêtes quelles qu'elles soient. Il simule la présence d'un fichier correspondant à la requête et si ce fichier est téléchargé, c'est en réalité le vers qui se transmet et infecte la nouvelle machine. Ainsi, les nSuds qui sont infectés deviennent rapidement surchargés (puisqu'ils répondent à toutes les requêtes).

Si il n'avait aucun effet néfaste sur les machines des utilisateurs, si ce n'est les polluer et ralentir leur trafic réseau, ce worm a démontré la vulnérabilité d'un réseau où l'information circule librement.

En plus d'être un lieu de partage, le réseau Gnutella est aussi un terrain de jeu parfait pour répandre des programmes malveillants. D'autres pourraient se servir de ce principe pour utiliser le réseau à des fins de propagande par exemple. Ce genre d'activité est non seulement néfaste à l'activité physique du réseau mais a aussi l'inconvénient de briser la chaîne de confiance entre les utilisateurs.

Combattre le freeloading

Les réseaux peer-to-peer ne peuvent fonctionner correctement que s'il y a participation active de leurs membres. Ainsi, une des pratiques les plus néfastes au fonctionnement de ces réseaux est le freeloading. Les freeloaders bénéficient des ressources partagées sur le réseau, sans pour autant partager les leurs.

De tels peers sont comme des fuites dans le réseau. Une grande quantité de bande passante est dépensée pour acheminer les ressources jusqu'à eux, sans qu'il n'y ait de contribution au reste du réseau en retour.

Au début de Gnutella on a vu apparaître des sites web proposant à leurs visiteurs de télécharger des fichiers sur le réseau. Bien évidemment, ces sites ne proposaient que le service de recherche et de téléchargement sans permettre aux visiteurs de soumettre leurs fichiers pour les introduire dans le réseau. Cette activité a rapidement déclenché l'indignation des membres actifs, et le protocole a finalement été rectifié pour interdire l'accès au réseau à ce genre de sites.

Quand Mojo Nation réinvente le commerce

Une autre approche, moins radicale, pour combattre le freeloading est celle de Mojo Nation. Le système tente d'accorder des crédits (Mojo) aux utilisateurs qui favorisent le fonctionnement du réseau.

Un peer disposant d'une forte bande passante et partageant un grand nombre de fichiers se verra accorder plus de Mojo qu'un peer à faible connexion et partageant peu de fichiers. Le Mojo ainsi collecté, sert de monnaie d'échange

pour télécharger des fichiers.

Chaque transaction est assimilable à un achat ou une vente. Lorsqu'un peer A fournit une ressource à un peer B, A « vend » sa ressource à B. Le coût (en Mojo) est établi en fonction de la qualité de service qu'il fournit.

Naturellement, les peers contribuant au réseau, c'est-à-dire disposant et délivrant des ressources intéressantes pour les autres, seront récompensés en Mojo. Le Mojo ainsi acquis leur permet « d'acheter » des ressources présentes sur le réseau. Le coût des ressources est fonction de leur rareté et de leur facilité d'accès. Télécharger un fichier coûtera plus de Mojo si il est trouvé sur peu de peers et/ou si ces peers ont une connexion lente.

Cette approche favorise les utilisateurs contribuant favorablement au réseau et dissuade les freeloaders. Il s'agit d'une approche commerciale, mais la monnaie n'a de valeur qu'à l'intérieur du système. Ce mécanisme est assez intéressant socialement car il incite les membres à ne pas profiter du réseau sans y participer, puisque un freeloader sera rapidement à cours de Mojo.

Morale et éthique

Le peer-to-peer, s'il a vocation à stimuler l'échange et le partage, connaît de nombreux détracteurs car il ouvre la porte à des dérives encore difficilement contrôlables.

Le problème des droits d'auteur

Napster doit autant sa popularité à son logiciel qu'à ses péripéties judiciaires. La société a été poursuivie en justice par les principaux majors de l'industrie musicale et la RIAA [RIAA] pour avoir incité ses utilisateurs à échanger des fichiers protégés par le copyright.

Mais la société n'est pas la seule à avoir connu des problèmes judiciaires, ses successeurs, le trio Morpheus / Kazaa / Grokster, sont également poursuivis aux Etats Unis pour les mêmes raisons. Et même aux Pays Bas, la société Fast Track - qui a créé le réseau que les trois sociétés en question utilisent - a été sommée par la justice néerlandaise d'empêcher aux utilisateurs de son réseau de télécharger du matériel protégé par le copyright.

Auparavant, le manque d'alternatives offrait des profits confortables aux majors de l'industrie musicale. L'arrivée du format MP3, avait, déjà en son temps, inquiété l'industrie musicale. Et les utilisateurs ont rapidement pris goût à la musique gratuite. Avec les applications de partage de fichier, l'accès à ce genre de musique devient un jeu d'enfant.

Le problème est épineux, puisque peu d'applications de partage de fichiers ont été conçues, dès le départ, dans le souci d'offrir une juste rémunération aux artistes.

Même si Napster s'est réconcilié avec BMG en essayant de mettre en place un système payant, il y a fort à parier que la majorité des utilisateurs refusera de payer pour quelque chose qu'ils peuvent avoir gratuitement.

Pornographie

Ce n'est un secret pour personne, les applications de partage de fichiers peer-to-peer ont pu être construites et leur popularité avec l'échange de fichiers musicaux, mais aujourd'hui la majorité du trafic est certainement générée par du contenu à caractère pornographique.

Outre l'impact sur le trafic du réseau, cela pose des problèmes aux parents désireux de protéger l'exposition de leurs enfants à ce genre de contenu. Les logiciels proposent en général des systèmes de filtrage se basant sur des listes de mots-clés. Mais en pratique, ceux-ci sont généralement peu efficaces.

Le réseau Gnutella, par exemple, est une mine d'or pour les amateurs de vidéos pornographiques [SHW0701]. Un rapport [WAXLAR0701] du gouvernement américain a montré que trois des mots les plus recherchés étaient porn, xxx et sex.

Une recherche aussi anodine que « Britney Spears », renvoie un nombre de réponses assez impressionnant. Cependant, peu des fichiers renvoyés ne sont effectivement en rapport avec la chanteuse américaine.

Outre le préjudice moral que cela peut porter aux artistes, cela pénalise les utilisateurs qui recherchent autre chose que des vidéos pornographiques. Cette forme de pollution numérique remet en cause l'utilité et la pertinence des recherches avec ce genre d'outils.

Autres activités illicites

Si les réseaux peer-to-peer sont un bon moyen de faciliter le partage d'informations intellectuellement stimulantes, ils sont aussi un facilitateur pour les activités illicites.

Stimulé par le sentiment de pouvoir tout trouver, l'internaute est bien souvent à la recherche de documents controversés ou amoraux. On trouve notamment sur Internet des sites spécialisés dans les contenus outrageants comme des snuff movies.

De même, les autorités ont fréquemment recours à Internet pour démanteler des réseaux de pornographie infantile.

On peut aussi imaginer que les réseaux de terrorisme puissent librement diffuser des informations critiques comme des plans de construction de bombes par exemple. Dans ce cas, les réseaux peer-to-peer peuvent être considérés par les gouvernements comme des éléments d'insécurité publique.

Dès lors que ces informations sont publiées sur des sites, les autorités disposent de moyens pour retrouver les auteurs en consultant les historiques des fournisseurs d'accès par exemple. Pour les réseaux décentralisés la tâche est bien plus ardue puisque certains systèmes ne permettent pas de déterminer l'origine des données présentes sur le réseau.

Conclusion

Issu d'une volonté d'accessibilité équitable à l'information et à la connaissance, l'Internet a rapidement démontré son potentiel économique. Pour répondre à une logique marchande, les services web n'ont cessé de progresser en s'appuyant sur le modèle client/serveur.

Architecturé de manière asymétrique, le client/serveur repoussait les clients légers en bordure du réseau. Ne permettant pas à tous de tirer pleinement parti des ressources globales d'Internet, ce modèle a peut-être trouvé une

alternative avec le peer-to-peer.

En intégrant les fonctions de client et de serveur sur le même poste, le peer-to-peer permet à tous de tirer partie des ressources globales du réseau.

Ce concept, redécouvert récemment, a pourtant connu des applications bien avant Napster, sans qu'il ne soit clairement identifié. Aujourd'hui encore, les avis divergent pour distinguer ce qui est de ce qui n'est pas peer-to-peer. Néanmoins, les experts s'accordent à dire qu'il n'existe pas une architecture unique mais plusieurs, répondant chacune à des problématiques différentes.

Avec l'expérience, les concepts se raffinent et les protocoles peer-to-peer sont désormais prêts à satisfaire des exigences professionnelles variées. Les applications sont multiples. Qu'il s'agisse de calcul distribué, de partage de fichiers, de travail collaboratif, d'applications militaires, ou de réseaux de distribution, la technologie est prête à démontrer son potentiel.

Pourtant, les entreprises restent prudentes car elles exigent des standards robustes et sûrs avant de se lancer. Nul doute que le peer-to-peer saura convaincre et s'intégrer à l'Internet d'aujourd'hui, puisque le marketing s'est déjà emparé du concept avec des résultats prometteurs.

Pour l'internaute aussi le peer-to-peer amène un grand changement. S'il peut désormais participer activement à la dynamique d'Internet, il convient de nuancer l'enthousiasme ambiant. Les réseaux de partage de fichiers offrent, certes des possibilités énormes, mais génèrent aussi des comportements qui dérangent l'éthique. En observant le phénomène on est en droit de se demander si le peer-to-peer saura conserver la philosophie de partage qui est à l'origine d'Internet.

Avec l'Internet, les utilisateurs comme les entreprises disposent d'un outil unique qui pénètre chaque jour leur vie quotidienne. Le peer-to-peer est une étape de plus dans la construction de la société de l'information. Paradis des libertés, ou enfer anarchique dominé par l'individualisme, l'avenir d'Internet ne peut être envisagé que selon l'orientation que lui donneront les entreprises où les particuliers. Alors ? Demain nous promet-il un Internet tirelire où encyclopédique ?

Glossaire

.NET (dot-net) : Plateforme stratégique de Microsoft dont l'objectif est la convergence des ordinateurs personnels avec le web.

API (Application Programming Interface) : Bibliothèque de fonctions prêtes à être utilisées pour programmer des applications.

ASP (Application Service Provider) : Entreprise qui propose des services applicatifs accessibles depuis Internet plutôt que de les installer localement sur des machines.

B2B (Business to Business) : Acronyme utilisé pour désigner le marché des services fournis par des entreprises pour d'autres entreprises.

B2C (Business to Consumer) : Acronyme utilisé pour désigner le marché des services fournis par des entreprises pour les particuliers.

Bande Passante : Désigne la quantité de données qu'une connexion réseau est capable de transférer par unité de temps.

BizTalk : Solution Microsoft alternative à ebXML.

Broadcast : Mode de transmission réseau dans lequel les paquets sont destinés à l'ensemble des nœuds.

Cache : Technique consistant à stocker des données temporairement pour en faciliter l'accès plus tard.

Calcul distribué : Forme de calcul qui fait appel à des ordinateurs distants pour résoudre un problème.

CDN (Content Delivery Network) : Réseau de distribution de contenu.

Client/Serveur : modèle de communication réseau asymétrique dans lequel le client émet des requêtes vers un serveur qui y répond. Un serveur est généralement une machine disposant de ressources importantes ; il est capable de servir plusieurs clients de façon concurrente.

COM (Component Object Model) : Framework de Microsoft pour développer et supporter des objets composants de programmes.

CRM (Customer Relationship Management) : Terme désignant les méthodologies, les logiciels, les capacités d'Internet utilisables par une entreprise pour gérer des relations client d'une façon organisée.

DNS (Domain Name System) : Système destiné à faciliter le nommage des entités sur Internet, en remplaçant les adresses IP numériques par des noms plus faciles à retenir.

ebXML (Electronic Business XML) : Langage, basé sur XML, dont l'objectif est de faciliter l'interopérabilité des activités de commerce électronique.

ERP (Enterprise Resource Planning) : Terme désignant un large ensemble d'activités supportées par des applications multi module facilitant la gestion des entreprises.

Flooding : Technique consistant à répandre une transmission sur l'ensemble d'un réseau. Contrairement au broadcast, ce sont les nœuds traversés qui décident de propager la transmission.

Freeloading : Pratique qui consiste à utiliser les ressources du réseau sans partager les siennes.

FTP (File Transfer Protocol) : Protocole utilisé pour transférer des fichiers sur le réseau.

GPS (Global Positioning System) : Constellation de satellites permettant de localiser géographiquement un émetteur.

Grid computing : Forme de calcul distribué.

HTTP (Hyper Text Transfer Protocol) : Protocole destiné à faciliter la navigation sur le World Wide Web.

Hashage (Hashing) : Technique qui consiste à transformer une suite d'octets (comme un fichier) en une clé qui représente la chaîne de départ.

Intranet : Réseau privé interne à une organisation qui ne dépend pas d'Internet.

IP : Protocole de communication d'Internet.

Killer Application : Terme désignant une classe d'applications qui remportent un grand succès.

LAN (Local Area Network) : Réseau d'ordinateurs et de matériels étendu sur une zone limitée comme un bâtiment.

Legacy Systems : Systèmes hérités d'une époque où les langages, plateformes et techniques différaient de ceux de maintenant.

Méta-données (metadata) : Ensemble de données utilisées pour décrire le contenu d'autres données.

Mp3 (MPEG-1 Audio Layer3) : Format de compression de son permettant de conserver un niveau de qualité acceptable.

Network Discovery : Processus par lequel un peer tente de localiser les autres membres du réseau.

Noeud : Terme utilisé pour qualifier un point de connexion dans un réseau. Des liens séparent les noeuds.

P2P : voir peer-to-peer.

Pair : voir Peer.

Pair à pair : voir peer-to-peer.

Peer : Participant d'un réseau peer-to-peer.

Peer-to-peer : lisez ce document !

RSA : Système de cryptage et d'authentification utilisant des clés privées et publiques.

SCM (Supply Chain Management) : Gestion des mouvements des matériels, de l'information et de l'argent dans un processus allant du fournisseur au consommateur (en passant par le fabricant, le grossiste et le détaillant).

SMTP (Simple Mail Transfer Protocol) : Protocole d'acheminement des e-mails.

SOAP (Simple Object Access Protocol) : Protocole d'accès à des objets distants se basant sur HTTP et XML.

Streaming : Ensemble de techniques permettant la diffusion de flux multimédia sur Internet sans que le destinataire du flux n ait besoin d attendre que la totalité du flux soit transmise avant de le jouer.

Super-noeud (supernode) : Noeud d un réseau peer-to-peer assurant des fonctions de serveur local. On trouve les super-noeuds dans de nombreux systèmes peer-to-peer mais sous des noms différents (ex : root node)

Swarm distribution : Technique qui consiste à répartir la fonction de transmission de données sur un réseau. Plusieurs peers participent collectivement au transfert en transmettant une petite partie des données.

TCP : voir TCP.

TCP/IP (Transmission Control Protocol / Internet Protocol) : Protocole de transmission et de contrôle des paquets IP. Il s agit du protocole de base pour la communication sur Internet.

Telnet : Protocole d accès à des machines distantes.

Téraflop : Unité de mesure de puissance de calcul représentant 1000 milliards d opérations à virgule flottante par seconde.

Téraoctet : Unité de mesure de capacité de mémoire représentant 1000 milliards d octets.

UDDI (Universal Description Discovery and Integration) : Annuaire Internet basé sur XML destiné à permettre aux entreprises de présenter leur activité et rendre leurs systèmes inter opérables pour réaliser du commerce électronique.

UDP (User Datagram Protocol) : Protocole alternatif à TCP qui n effectue pas de contrôle sur l ordre et la quantité des paquets qu il achemine. Ce contrôle est laissé à la charge de l application.

Usenet : Une collection de messages, organisés en groupes de discussion (newsgroups) soumis par des utilisateurs partout dans le monde. Usenet est organisé en groupes de discussion.

Web service : Service proposé par un serveur web à destination d utilisateurs ou de programmes connectés à Internet.

Workflow : Terme utilisé pour décrire les tâches, les procédures, les gens, les organisations, les informations et outils nécessaires à un processus de gestion (business process).

WSDL (Web Services Description Language) : Langage, basé sur XML, destiné à décrire des services web et à fournir les moyens d y accéder.

XML (eXtensible Markup Language) : Métalangage de description qui est abondamment utilisé sur Internet dans des protocoles applicatifs.

XML-RPC (XML Remote Procedure Call) : Protocole applicatif d invocation de fonctions implantées par sur des systèmes distants.