Sistemas de redes *peer to peer*

Alberto Martinez Scremin
Barbara Sabrina Herrera
Bianca Caruso da Paixão
Daniel Tamaki

Informática I

Índice:

1. Introdução

Neste trabalho iremos apresentar os sistemas de rede, com enfoque para a arquitetura *peer to peer*. Primeiramente, iremos definir e explicar o funcionamento dessa arquitetura, e compará-la com outros tipos de rede, diferenciando-os e apresentando vantagens e desvantagens de cada uma.

Posteriormente, nos aprofundamos nesse sistema, assim, iremos falar um pouco sobre a história do p2p. Além disso, também serão mostrados os mecanismos de busca em uma rede *peer to peer* e os sistemas de segurança nessa rede.

Finalmente, serão apresentados questionamentos sobre a legalidade e a ilegalidade de certas práticas nessa arquitetura.

2. O que é *peer to peer* ?

Segundo a definição do dicionário Português-Inglês, *peer to peer* significa, de pares em pares. Isso quer dizer que os computadores da rede estão todos interligados em uma cadeia descentralizada, onde cada um possui funções equivalentes não havendo uma hierarquia entre eles. Todos os usuários são clientes e servidores, funcionando, assim, de forma totalmente independente e livre da existência de um servidor central. (ver anexo 9.1)

Uma pesquisa realizada recentemente pela Xerox mostrou que 70% dos usuários das redes p2p sequer compartilham arquivos, enquanto que 1% dos usuários compartilham 50% dos arquivos disponibilizados. Assim, esta pesquisa mostra o que já havia sido percebido na prática: que redes *peer to peer* puras na prática não funcionam como esperado na teoria. Então, algumas aplicações, como por exemplo o Napster, encontraram como solução para tal problema a criação do modelo peer to peer misto, onde os serviços críticos são realizados por servidores, mesclando, assim, a arquitetura p2p e cliente/servidor. (ver anexo 9.3)

A vantagem de uma arquitetura de rede descentralizada é que ela é bem mais difícil de ser interrompida, pois não existe mais um ponto de falha. Porém a busca neste tipo de rede é muito lenta e não é garantido que a consulta terá algum resultado, porque o arquivo desejado pode estar a uma distância muito grande para ser alcançado.

A busca realizada sobre determinado arquivo é enviada para todos os computadores presentes na rede, os quais tomam conhecimento dos demais usuários presentes por meio de um *Host*[1]*,* este não possui bancos de dados como nomes de usuários ou arquivos, ele apenas monitora os pontos de presença na rede. Ele informa o IP dos usuários dos softwares que estão na rede. Então, o pedido do arquivo é repassado para todos os usuários da rede os quais repassam para outros usuários, assim por diante até que seja encontrado o arquivo desejado.

3. Arquitetura Cliente/Servidor

Diferentemente da arquitetura citada acima, esta estrutura funciona de forma mais centralizada e é clara a participação de um servidor central no ato de compartilhamento de arquivos. (ver anexo 9.2) Nessa arquitetura, quando um usuário (cliente) pretende encontrar um determinado arquivo, ele envia o nome deste arquivo para o servidor central, onde há um banco de dados com todos os demais usuários (clientes) do software e com o registro de quais arquivos cada um deles está disponibilizando para *download*. Assim, o servidor central tem a função de informar aos clientes, quais os outros clientes possuem o arquivo solicitado e também estabelece a conexão direta entre os dois computadores, sem que a transmissão do arquivo passe pelo servidor.

[1] Host é o computador central, também chamado de servidor, que controla e armazena os programas e dados utilizados por outros computadores em uma rede. No caso do compartilhamento de dados sob arquitetura *peer to peer*, ele funciona como o “motor de arranque” para o funcionamento da rede, uma vez dentro da rede, o usuário não precisa mais se comunicar com o Host.

Outro aspecto importante é que por meio desta arquitetura podem ser tomadas providências técnicas, como o uso de filtros, com o intuito de inibir a requisição de arquivos protegidos por direitos autorais, além de poder identificar o principal responsável pelo ato infrator, o servidor.

Assim, em um software que funcione sob arquitetura Cliente/servidor, a tarefa de monitorar, impedir ou restringir o acesso de uma máquina a outra é realizada facilmente, ao passo que em uma rede sob a arquitetura *peer to peer* esta prática torna-se extremamente difícil. Isso porque em uma estrutura não hierarquizada, a identificação dos responsáveis é dificultada pela ausência de um servidor central, bem como, por se tratar de uma rede com milhares de usuários conectados simultaneamente, o que é praticamente impossível de ser monitorada integralmente.

4. A história do *peer to peer*

O peer to peer sempre existiu, porém não era reconhecido como tal. Servidores com endereços de IP fixos ou determináveis podiam sempre se comunicar com outros servidores para usufruir os serviços disponibilizados. Um número de aplicações pré-P2P, como, por exemplo, os correios eletrônicos, possuíam a capacidade de oferecer um serviço em uma rede distribuída. Entretanto, um aplicativo ganhou destaque na época, a USENET.

A USENET surgiu em 1979, e foi desenvolvida por estudantes universitários dos EUA. Ela permitia que dois computadores trocassem informação, nos primórdios, antes da conexão oferecida pela Internet. A primeira interação consistia na capacidade de um computador de se ligar a outro, via modem, pesquisar documentos e transferir esses documentos para armazenamento local. Nesta aplicação não existia nenhuma autoridade central, a distribuição dos documentos é gerada por cada usuário e o conteúdo da rede é replicado para todos os usuários dela. Entretanto, algumas propriedades da USENET ajudam a distinguí-la do primeiro sistema *peer to peer*.

Com o surgimento do ICQ, o tradicional correio eletrônico foi substituído, já que esse programa oferecia aos seus usuários a capacidade de comunicar-se de forma mais rápida e permitia também que seus utilizadores fossem notificados quando seus amigos se ligavam. Além disso, ele permitia também a troca de arquivos, que embora seja classificada como uma aplicação p2p, o ICQ usa uma arquitetura mista entre p2p e cliente/servidor. Seu serviço de notificação de novos usuários é centralizado em um servidor, mas todos os outros serviços são puramente *peer to peer*.

O Napster surge em 1999, pelo estudante Shawn Fanning, oferecendo aos seus usuários a possibilidade de compartilharem arquivos em *mp3*. Como uma arquitetura mista, semelhante à do ICQ, era suportado por um servidor central, que armazenava as listas de músicas das diferentes máquinas e fazia a conexão entre quem buscava um arquivo e quem possuía o arquivo. Era genial, rápido, mas também fácil de frear. Sem o servidor, o Napster não existia, e a justiça norte-americana conseguiu fechá-lo por infringir a lei de direitos autorais.

Aproveitando a deixa do *Napster* surge o Gnutella e a FastTrack que levaram o compartilhamento de arquivos a um passo mais longe: eliminaram a necessidade de um servidor central para realizar a pesquisa, elas ligam os computadores ponto a ponto, o que as torna praticamente impossíveis de fechar. Para funcionar, essas redes usam o endereço IP, o RG do computador, que permite localizar cada uma das máquinas na *web*. Cada computador se conecta a até cinco máquinas para fazer buscas. Essas últimas a mais outras cinco e assim por diante, formando uma rede de milhares de computadores em minutos.

O FastTrack deixou essa conexão ainda mais rápida ao transformar alguns dos computadores da rede - os que tinham mais velocidade e maior capacidade de processamento - em subservidores, que armazenam dados de diversos outros computadores, criando verdadeiros atalhos para a informação.

Outros sistemas de partilha de arquivos aparecerem, cada um tentando resolver os problemas e aprimorar os sistemas atuais. Temos, de um lado, redes mais confiáveis e que conseguem garantir anonimato a seus usuários. É o caso de redes como a Freenet, a Entropy e da GNUnet. De outro, estão as redes privadas,

mais seguras, que conectam apenas amigos. É o caso da Mute, da Waste e do programa Grouper.

5. Segurança

A segurança é um componente essencial para qualquer sistema de computação e é especialmente relevante para sistemas P2P. Navegar pelas redes P2P pode ser não muito seguro, pois existem várias ameaças dentro da rede, como vírus que vem com arquivos e outros. O P2P compartilha muitos problemas de segurança e soluções com o resto da rede e sistemas distribuídos. Por exemplo, dados corrompidos, transferência não confiável, problemas de latência e problemas de identificação são alguns deles.

5.1. Principais ameaças da rede P2P:

As principais ameaças que existem para os usuários da rede P2P são:

- a contaminação por vírus, *worms* e *malware* nos arquivos baixados por P2P;
- a invasão através de vulnerabilidades em aplicações de redes P2P;
- a publicação de informações sensíveis através de compartilhamento de arquivos acidentalmente;
- os processos judiciais em decorrência de violação de direitos autorais através da obtenção de software, filmes, músicas, livros obtidos via redes P2P. As ações judiciais, principalmente no exterior, têm sido mais freqüentes devido a atuação mais efetiva de associações e empresas na monitoração das atividades em redes P2P.

5.2. Outros tipos de ameaças da rede P2P:

Alguns outros tipos de ameaças podem ser citados como: roubos, onde empresas perdem milhões em propriedades, tal como código fonte; obstrução da

largura de banda, aplicações como KaZaa, Gnutella e Freenet tornam essas ações possíveis para um computador que compartilha arquivos com outros na Internet; *Bugs*, para que as aplicações de compartilhamento de arquivo funcionem apropriadamente o software correto deve ser instalado no sistema dos usuários. Se este *software* contém um *bug* isto pode expor a rede a inúmeros riscos; quebra de criptografia; falta de confidenciabilidade. O KaZaa e o Gnutella dão acesso direto à arquivos armazenados nos discos rígidos de outros clientes, como conseqüência é possível para um *hacker* descobrir qual o sistema operacional o computador hospedeiro possui e conectar a pastas que não estão permitidas para compartilhamento obtendo assim acesso a informações sigilosas e arquivos de sistema.

Interoperabilidade é a área de maior interesse de segurança a respeito de redes P2P. A introdução de diferentes plataformas, diferentes sistemas e diferentes aplicações trabalhando juntas numa dada infra-estrutura abre uma série de questões a respeito de segurança que é associado com a interoperabilidade. Quanto maiores forem as diferenças dentro de uma infra-estrutura, maiores serão os problemas de segurança associados, Negócio privado numa rede pública Muitas companhias conduzem negócios particulares através de redes públicas. Isto leva a exposição do sistema a diversos riscos de segurança.

Também sempre existirão usuários maliciosos que têm a intenção de ganhar acesso clandestino às redes corporativas. E não importando qual o protocolo de segurança é colocado para impedir a ação desses usuários, com o tempo eles encontrarão um jeito de burlar esses protocolos. Então, o que os mecanismos de segurança precisam fazer é se manterem à frente desses *hackers* criando cada vez maiores e melhores protocolos. Mas isto não é tão simples de ser realizado.

5.3. Mecanismos de segurança da rede P2P:

Todos os sistemas de segurança atuais são baseados em criptografia que utiliza tanto chave simétrica e privada quanto chave assimétrica e pública ou, às vezes, uma combinação das duas.

Técnicas da chave privada são baseadas no fato de que o remetente e o destinatário compartilham um segredo, que é alterado através de várias operações de criptografia como a encriptação de decriptação de mensagens e a criação e verificação de dados de autenticação de mensagens. Essa chave pública deve ser trocada através do um processo fora do campo anterior para a comunicação pretendida.

Técnicas da chave pública são baseadas nos pares de chaves assimétricas. Geralmente cada usuário está em posse de apenas um par de chaves. Uma das chaves do par está publicamente disponível enquanto que a outra está privada. Porque uma chave está disponível não é necessária uma troca de chave fora da banda, entretanto existe a necessidade de uma infra-estrutura para distribuir a chave pública autenticavelmente. Já que não é necessário um pré-compartilhamento dos segredos anteriores para a comunicação, as chaves públicas são ideais para apoiar segurança entre partes previamente conhecidas.

Par de chaves assimétricas são diferentes de uma chave da porta da frente, que permite ao seu portador bloquear ou desbloquear esta porta com igual facilidade, a chave pública utilizada na criptografia é assimétrica. Isto significa que a chave consegue encriptar uma mensagem com extrema facilidade, porém decriptar como um todo, é considerado muito difícil.

5.4. Protocolos da rede P2P:

Mecanismos para estabelecer uma criptografia forte e verificável são muito importantes. Estes mecanismos são padrões de protocolos de autorização que permitem que pares assegurem que eles estão falando com o sistema remoto pretendido.

5.4.1. Protocolo *Secure Sockets Layer* (SSL):

Para uma proteção das informações transmitidas em uma rede P2P, alguns sistemas empregam o protocolo padrão *Secure Sockets Level* (SSL). Este protocolo garante que um arquivo e eventos enviados cheguem intactos e

invisíveis para qualquer um que um que não seja o destinatário. Além disso, porque todos os pares usam SSL, ambos os lados automaticamente provam quem eles são para cada um antes de transferirem qualquer informação pela rede. O protocolo providencia mecanismos de comunicação que são à prova de adulterações e confidenciais, usando as mesmas técnicas utilizadas pela maioria dos *webmaster* que protegem usuário em transações financeiras transmitidas pela Internet.

5.4.2. Tecnologia IPSec:

A maioria das VPN´s (*Virtual Private Networks*) utilizam a tecnologia IPsec, o *framework* que envolve vários protocolos que tem sido o padrão para vários especialistas. O IPsec é útil porque ele é compatível com os mais diferentes tipos de *hardware* e *software* para VPN, e é o melhor para redes com clientes com acesso remoto. O IPsec requer pouco conhecimento para os cliente, porque a autenticação não é baseada no usuário, o que significa que um *token*[2] (como um Secure ID ou um Crypto Card) não é utilizado. Ao invés disso, a segurança vem do IP da estação ou do seu certificado (e.g. X.509), que estabelece a identidade do usuário e garante sua integridade na rede. Um túnel IPsec basicamente age como uma camada de rede que protege todos os pacotes de dados que passa por ela qualquer que seja a aplicação.

5.4.3. Infra-estrutura de chave pública (PKI):

A *Public Key Infrastruture* (PKI), toda caracterizada pelo X.509 sobre um *backbone*[3] de rede SSL. Essa combinação de X.509, autenticação PKI e encriptação de transporte SSL é o estabelecimento de um padrão criptográfico para o *e-commerce* na Internet. Este padrão permite certificados de segurança da Endeavors, ou de qualquer outra reconhecida autoridade do certificado X.509,

---

[2] *Token* é o arquivo utilizado em redes tipo anel que quando lançado na rede , informa que um determinado micro deseja falar com outro , evitando assim colisões de informações. Em computação também é conhecido como uma seqüência de caracteres com um significado coletivo.

[3] No contexto de redes de computadores, o backbone (traduzindo para português, *espinha dorsal*) designa o esquema de ligações centrais de um sistema mais amplo. Em termos de composição, o backbone deve ser concebido com protocolos e interfaces apropriados ao débito que se pretende manter.

para fornecer uma identidade verdadeira a qualquer par que surja na rede. O uso da encriptação de segurança ponto-a-ponto SSL habilita cada par de *peers* que se comunicam a terem uma única chave. A vantagem é que quando um *peer* deixa a comunidade, todas as suas chaves únicas se tornam inválidas, mas as chaves entre os outros membros da comunidade não são afetadas.

5.5. Segurança do usuário na rede P2P:

O usuário para ter segurança pode lançar mão de alguns recursos tais como: *Firewalls*, IP dinâmico e *NAT*, que crescem além da necessidade na arquitetura da Internet de se tornar um sistema escalável e seguro, mas novas aplicações P2P desafiam essa arquitetura demandando que os participantes sirvam recursos como também os usem. Devemos estar preocupados em nos defendermos de nós mesmos. Os pontos mais importantes para isso são: controle de conexão, controle de acesso, controle de operação, antivírus e é claro a proteção dos dados armazenados nos nossos computadores.

5.6. Futuro da segurança da rede P2P:

As redes P2P têm recebido cada vez mais atenção ao longo da sua existência. Cada vez mais as pessoas tendem a compartilhar seus recursos e a necessitarem de cada vez mais segurança, pois a industria da infração eletrônica também acompanha essa evolução.

5.6.1. Biometria:

A biometria envolve o uso de uma característica pessoal para autenticar o usuário. Características que são comumente usadas incluem uma imagem facial da pessoa, uma assinatura, uma impressão digital ou um padrão de retina. A vantagem da característica biométrica é que os usuários não precisam lembra nenhuma senha ou armazenar nenhuma chave, que são as maiores fraquezas num sistema de autenticação convencional.

5.6.2. Criptografia por chave quântica:

Muitos sistemas de encriptação modernos dependem da dificuldade em realizar um ataque de força bruta às chaves privadas, devido às restrições de processamento e de tempo. Embora ainda no estágio teórico, a melhoria de desempenho dado por um hipotético computador quântico renderia uma quantidade bem inferior de algoritmos. Obviamente novos algoritmos de criptografia serão necessários. Criptografia quântica utiliza os estados dos fótons como chaves para codificar informações. De acordo com Heisenberg, é impossível descobrir ambos o momento e a posição de um fóton em um dado instante de tempo. Entretanto, em teoria um invasor não poderia descobrir uma chave privada baseando-se na informação do estado do fóton; o invasor teria que ter o fóton atual para decifrar qualquer dado encriptado com uma chave. Infelizmente este conceito é, para o momento, muito complexo para ser implementado. Os cientistas da IBM construíram o primeiro protótipo funcional de uma chave de distribuição quântica (QKD) no final dos anos 80. Ainda está muito longe, mas já pudemos observar grandes progressos nestas áreas. E embora nós não possamos ver o QKD no mercado por ainda um bom tempo, a tecnologia soa incrivelmente promissora.

6. A questão legal

A questão da legalidade e ilegalidade dos arquivos compartilhados é discutida, desde a popularização do Napster. Ele despertou a atenção da indústria fonográfica e de alguns artistas, que se sentiram prejudicados, visto que muitos de seus trabalhos estavam disponíveis na rede, antes mesmo de serem lançados. Logo, o Napster foi fechado pela justiça norte americana por infringir os direitos autorais.

A prática do compartilhamento de arquivos, por si só, não constitui uma violação dos direitos autorais, desde que os arquivos compartilhados não estejam

protegidos. Entretanto, o que vem acontecendo é a violação desses direitos e, assim, os autores não recebem nada pela troca, de suas obras, *online*.

Além disso, muitos filmes, músicas, livros, fotografias, desenhos, *software* são colocados na rede e compartilhados, antes mesmo do lançamento destes no mercado. Assim, o direito ao ineditismo, que é um direito moral do autor, é quebrado.

7. Mecanismos de busca

Os mecanismos de busca surgiram logo após o aparecimento da Internet, com a intenção de prestar um serviço extremamente importante: a busca de qualquer informação na *web,* apresentando os resultados de uma forma organizada, e também com a proposta de fazer isto de uma maneira rápida e eficiente. A seguir serão apresentados diferentes mecanismos de busca para redes p2p e suas respectivas descrições.

7.1. Transmissão de inundação de consultas (*Flooding broadcast of queries*)

A implementação original do Gnutella é um exemplo deste mecanismo de busca. Quando um par realiza uma consulta, ela é transmitida para todos os seus pares vizinhos. Se um par vizinho não fornecer resultados então esta consulta é transmitida para os pares vizinhos do vizinho. Se o recurso for encontrado, aquele par envia uma mensagem para o par que originou a consulta indicando que encontrou resultados para a consulta, e então estabelece uma conexão P2P.

7.2. Sistemas de seleção de encaminhamento

Este mecanismo envia a consulta para todos os pares e encaminha seletivamente a consulta para específicos pares que são considerados possíveis locais onde o recurso pode ser encontrado.

Se um par tiver banda larga suficiente para processar suas potencialidades ele se torna um super par. Os pares com conexão linha discada também podem consultar super pares.

Este mecanismo utiliza algoritmo de controle de fluxo (FCA-*Flow Control Algoritm*), uma forma inteligente de controlar o fluxo de como o par encaminhará e responderá a uma mensagem e um esquema sensível de prioridade para descartar mensagens que não caibam na conexão. Dessa maneira reduzindo a limitação da conexão de banda larga para a escalabilidade.

7.3. Redes de “hash table” descentralizada

Aqui, cada arquivo armazenado no sistema possui uma identidade (id) única do seu conteúdo, sendo utilizado para identificá-lo e localizá-lo.

Sendo assim, a sua localização se torna mais rápida e dessa maneira é impossível realizar uma busca nebulosa (*fuzzy*) dentro da rede. Se um par procura um arquivo de um outro par, ele obrigatoriamente deverá ter uma id para poder receber o arquivo.

7.4. Índices centralizados e repositórios

Este tipo de busca é utilizado pelo Napster. Índices de todos os pares e seus recursos são armazenados em um grande servidor central. A consulta é enviada ao servidor , o qual procura pelo índice correspondente. Se a consulta fornecer resultado, o servidor envia mensagem informando onde o par que realizou a consulta poderá conseguir o arquivo. Apesar deste mecanismo de busca

apresentar melhor desempenho, a banda e o hardware necessários para esta rede p2p de grande porte exige um custo muito alto para o sistema.
Sua deficiência consiste em que ocorrendo alguma falha no servidor toda a rede vem abaixo.

7.5. Índices distribuídos e repositórios

A idéia de índices distribuídos é que cada nó da rede contém um índice de arquivos locais assim como índices de arquivos armazenados de alguns pares vizinhos. Quando o nó recebe uma consulta, primeiramente ele verifica se pode ser feita naquela localidade. Se não puder, ele usa o índice local para determinar para qual nó a consulta será encaminhada. Em cada máquina local o índice não é estático, mudando de acordo com o movimento dos arquivos no sistema.

7.6. Relevância dirigida à “crawlers” de redes

Ele utiliza um banco de dados acumulados pelo par para assim determinar quais recursos podem ou não ser relevantes. Com o tempo é acumulada uma grande gama de informações, as quais são analisadas para determinar quais elementos comuns o par achou relevante. O *crawler* então atravessa a rede inserindo novas informações aos documentos em html, que são posteriormente comparados com o perfil destilado do par anterior. O tempo necessário para atravessar um número grande de índices é muito longo, não sendo este sistema aconselhável para redes de grande porte.

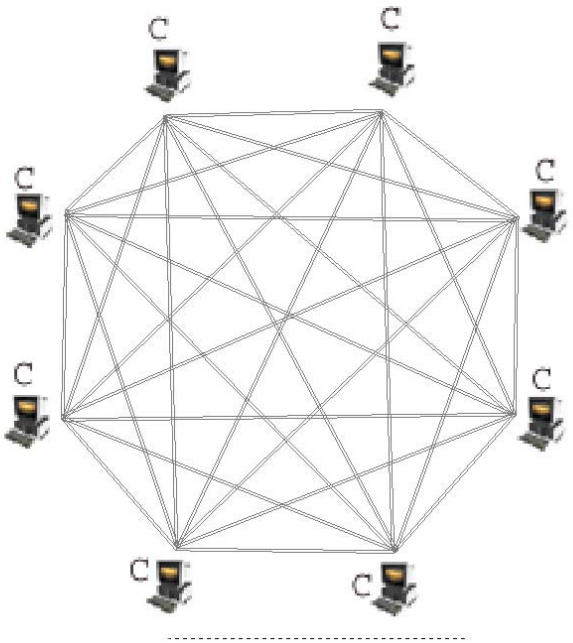
8. Conclusão

Existe uma grande motivação para o desenvolvimento e a utilização de aplicações P2P. Pelo lado técnico, elas representam a possibilidade do uso de sistemas de compartilhamento sem uma unidade central que podem se organizar

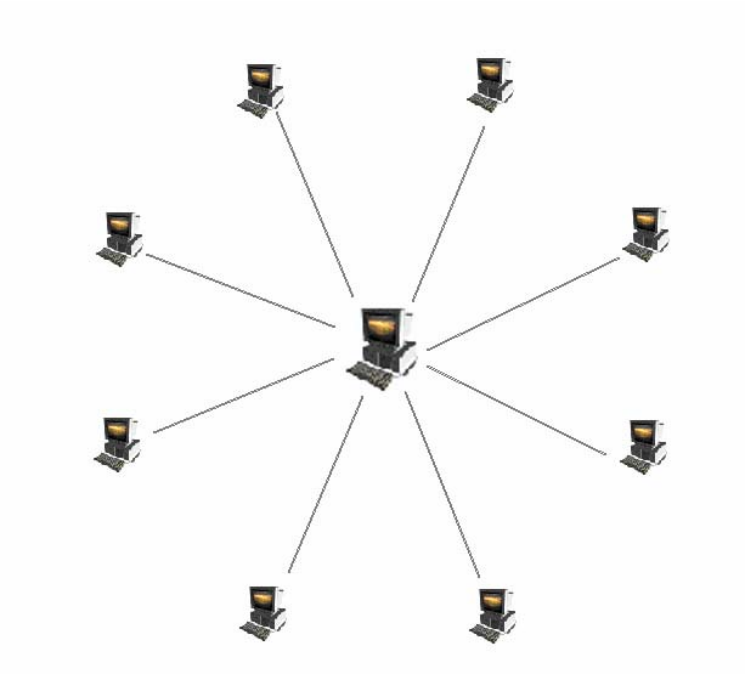
e funcionar de forma descentralizada. Por outro lado, existem razões não técnicas que também impulsionam o desenvolvimento das mesmas, sendo o anonimato a principal destas. As razões pelas quais as pessoas buscam o anonimato são basicamente duas, para o compartilhamento de material protegido por direitos autorais e para a livre expressão de idéias.

9. Anexos

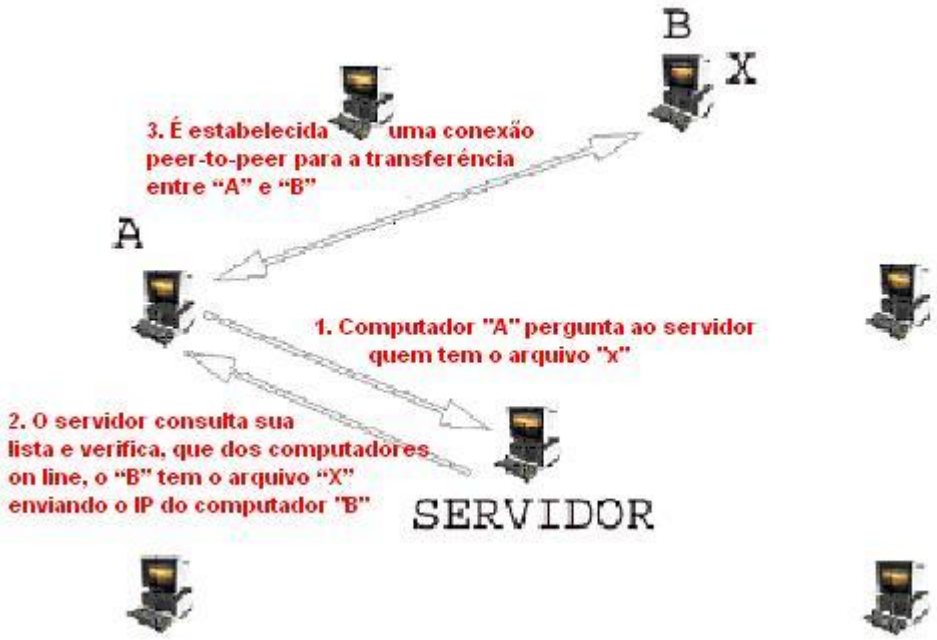
9.1. Esquema *peer to peer*



9.2. Esquema Cliente/servidor



9.3. Esquema misto (*peer to peer* e cliente/servidor)



10. Bibliografia

Redes p2p e BitTorrent: http://idgnow.uol.com.br/internet/2006/05/16/idgnoticia.2006-05-15.2495285982/

Universidade Federal do Rio de Janeiro: http://www.gta.ufrj.br/

http://www.wikipedia.org/

Online computer dictionary for computer and Internet terms and definitions: http://www.webopedia.com/

Seguranças em rede p2p: http://twiki.im.ufba.br/pub/MAT570/LivroseArtigos/p2p_sbrc2006_C5.pdf/