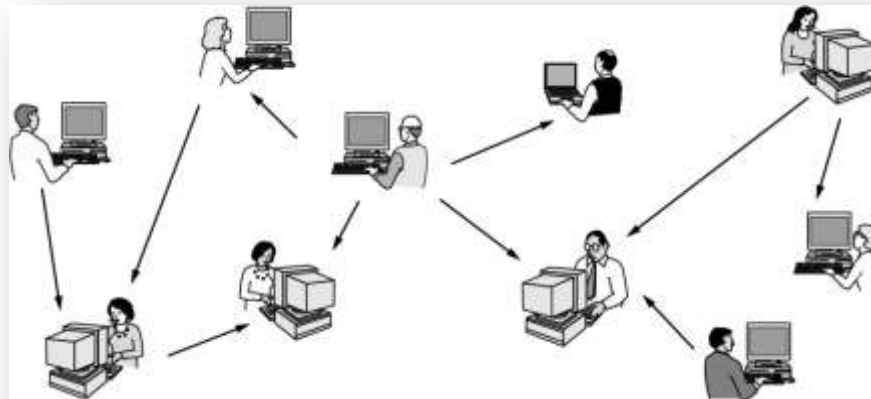


Peer-to-peer: Paradigma

□ Tecnologia peer-to-peer (P2P)

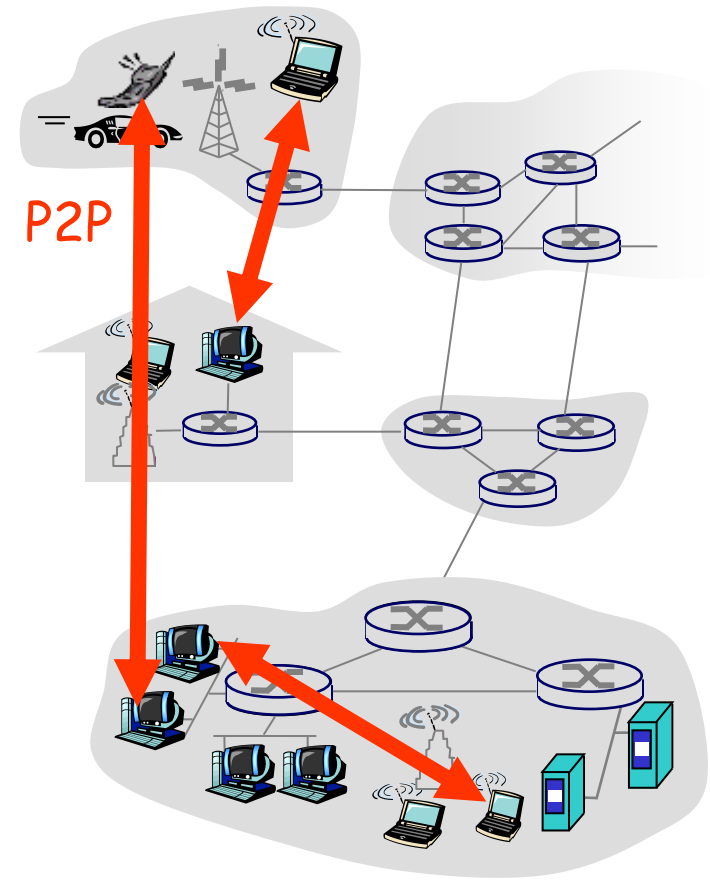
- Surgiu para mudar o paradigma existente (Modelo Cliente/Servidor)
 - Baseiam-se na capacidade de processamento e na largura de banda dos participantes da rede ao invés de se concentrar em um pequeno número de servidores
- Rede P2P pura:
 - São sistemas totalmente distribuídos
 - Não depende de uma organização central ou hierárquica
 - nos quais todos os nós têm as mesmas funcionalidades e tarefas.
 - Qualquer dispositivo pode acessar diretamente os recursos de outro, sem nenhum controle centralizado.



Peer-to-peer: Definição

□ Definições

- IETF: "o compartilhamento de recursos e serviços computacionais diretamente entre sistemas".
- São sistemas distribuídos
 - formados por nodos interconectados capazes de se auto-organizar em redes de sobreposição,
 - mantendo conectividade aceitável e desempenho mediante população transiente de nodos,
 - sem necessitar da intermediação ou apoio de uma entidade central
- Objetivos
 - compartilhar recursos como conteúdo, ciclos de CPU, armazenamento e largura de banda,



Peer-to-peer: Definição

□ Não há consenso

- Existem muitas arquiteturas diferentes para sistemas P2P
 - Diversas arquiteturas impuras: algumas utilizam super-nós outras utilizam servidores centralizados para executar funções auxiliares.
- Muitos sistemas são chamados de peer-to-peer pela forma como são percebidos pelos usuários e não pela sua arquitetura.
 - Exemplo: Voz sobre IP
 - Os usuários realizam chamadas entre si utilizando servidores
 - Após a sessão estabelecida a comunicação (áudio) é entre os pares

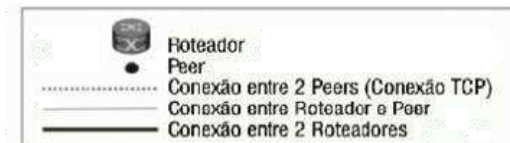
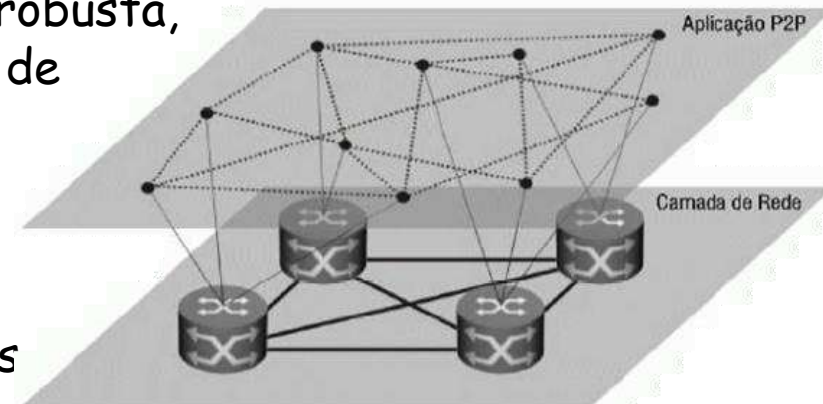
Peer-to-peer: Requisitos

□ Requisitos

- Nós podem estar localizados nas bordas da rede;
- Capacidade dos nós se comunicarem diretamente uns com os outros.
- Nós com conectividade variável ou temporária e endereços também variáveis;
- Capacidade de lidar com diferentes taxas de transmissão entre nós;
- Nós com autonomia parcial ou total em relação a um servidor centralizado;
 - Natureza distribuída da rede peer-to-peer garante robustez no caso de falhas devido a replicação dos dados através de múltiplos pares
 - Em sistemas P2P puros, pela capacidade dos pares de localizar dados sem necessitar de um servidor de indexação centralizado, evita existência de pontos de falha únicos no sistema.
- Assegurar que os nós possuem capacidades iguais de fornecer e consumir recursos de seus peers;
- Rede deve ser escalável;

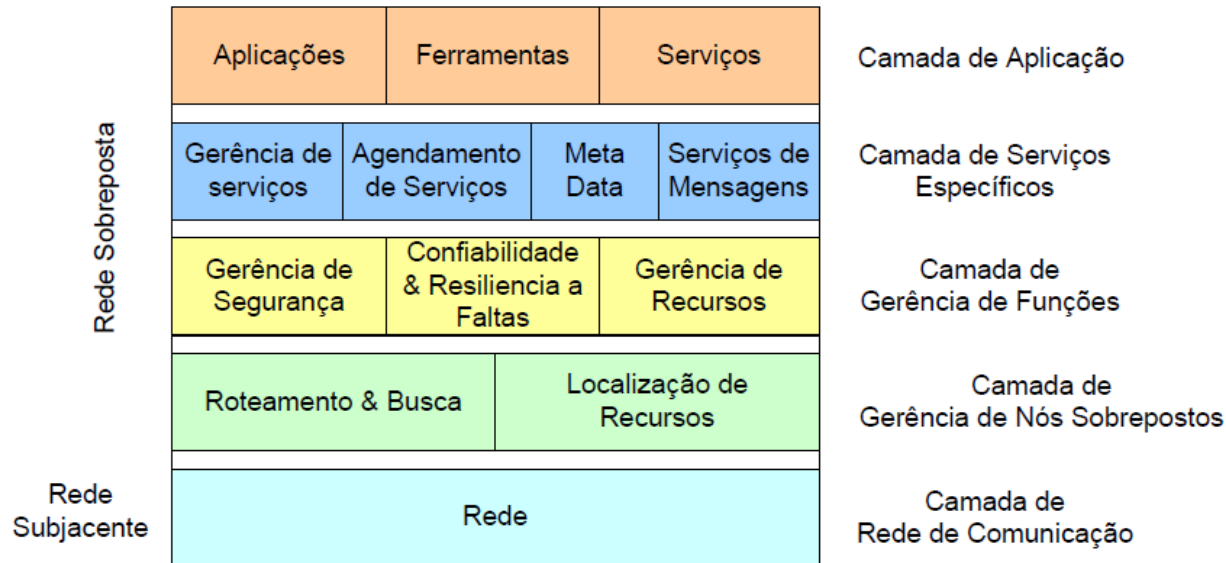
Peer-to-peer: Rede Sobreposta

- Pares formam uma rede auto-organizada sobreposta (overlay network) à rede IP
 - Rede sobreposta é uma rede de computadores construída sobre outra rede
 - Nós da rede sobreposta podem ser vistos como estando conectados através de conexões virtuais ou lógicas, cada uma delas corresponde a um caminho através de um número de conexões físicas na rede subjacente
- Oferecendo várias funcionalidades aos seus componentes:
 - arquitetura de roteamento WAN robusta, busca eficiente de dados, seleção de pares próximos, armazenamento redundante, durabilidade dos dados, nomenclatura hierárquica, autenticação, anonimato, escalabilidade e tolerância à faltas



Peer-to-peer: Rede Sobreposta

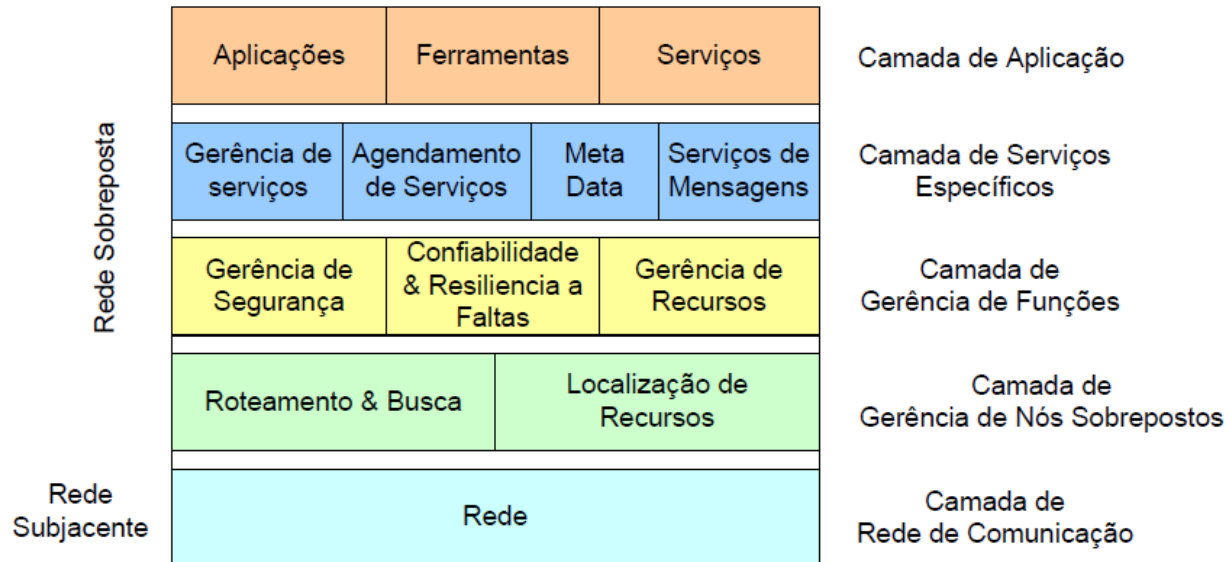
□ Arquitetura de redes P2P sobreposta



- Camada de rede de comunicação descreve as características da rede de computadores conectada através da Internet ou outras redes ad-hoc.

Peer-to-peer: Rede Sobreposta

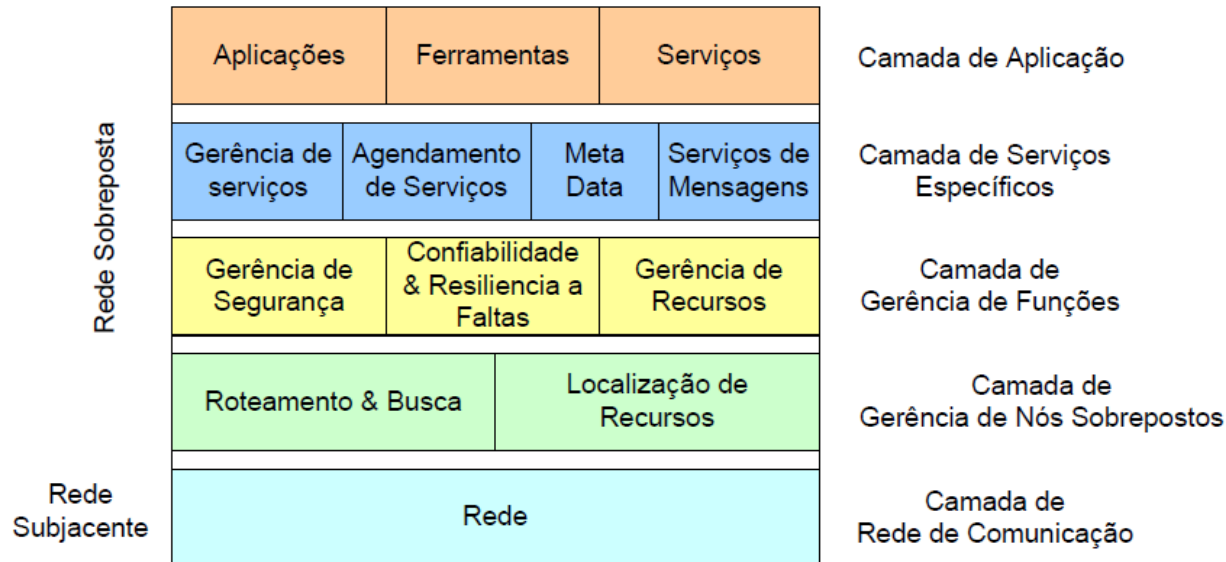
□ Arquitetura de redes P2P sobreposta



- A camada de Gerência de Nós Sobrepostos gerencia os pares, o que inclui detecção de pares e algoritmos de roteamento para otimização do sistema.

Peer-to-peer: Rede Sobreposta

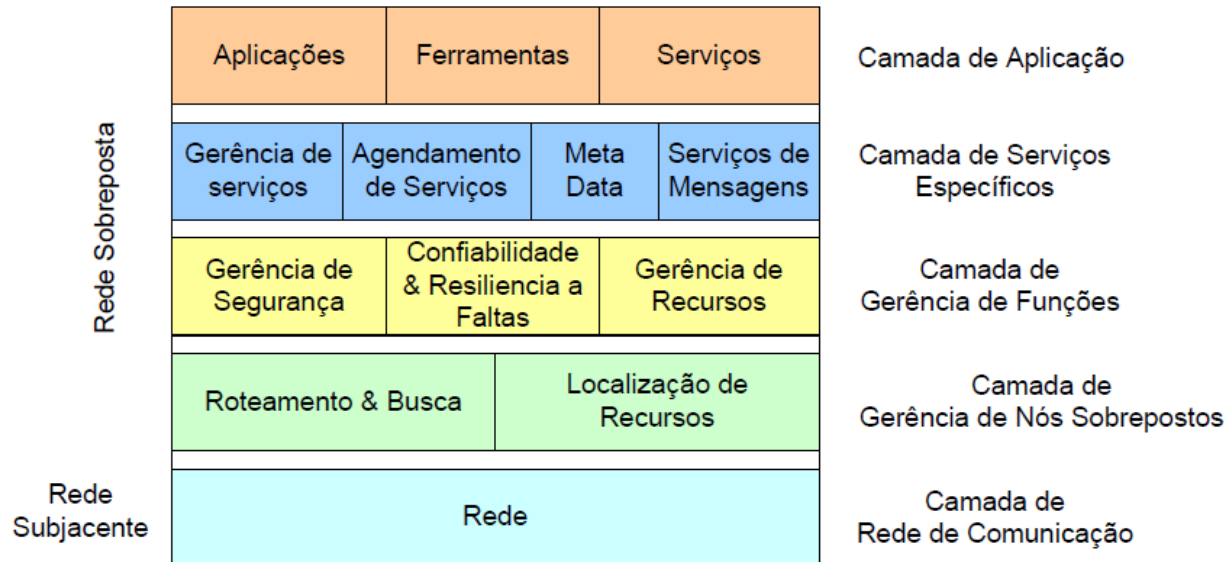
□ Arquitetura de redes P2P sobreposta



- A camada de Gerência de Funções lida com segurança, confiabilidade, resiliência a faltas e aspectos de disponibilidade de recursos agregados para a manutenção da robustez dos sistemas P2P.

Peer-to-peer: Rede Sobreposta

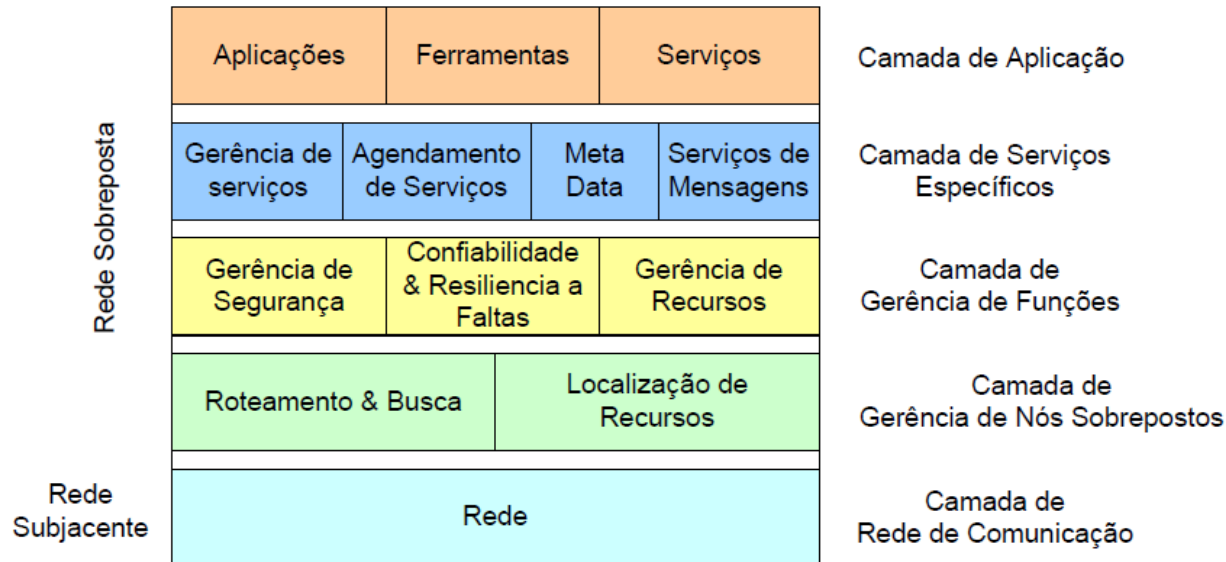
□ Arquitetura de redes P2P sobreposta



- A camada de Serviços Específicos dá suporte à infra-estrutura P2P subjacente e aos componentes específicos das aplicações através de agendamento de serviços, gerência de conteúdo e arquivos.

Peer-to-peer: Rede Sobreposta

□ Arquitetura de redes P2P sobreposta



- A camada de Aplicação está envolvida com ferramentas, aplicações e serviços que são implementados com funcionalidades específicas sobre a infra-estrutura P2P sobreposta

Aplicações P2P

□ Compartilhamento de arquivos

- LimeWire
- Emule
- BearShare
- Kazaa
- Pichat
- Gnutella
- Deluge
- qBittorrent
- BitComet
- Azureus
- Utorrent
- Ares Galaxy
- Kademlia
- SOULSEEK
- YaCy
- FastTrack
- BitTorrent

Aplicações P2P

□ Mensagens instantâneas:

- Jabber (<https://jabber-br.org/>), ICQ, MSN, Skype, K-Lite, Pichat, Pidgin, Usenet, P2P Messenger, SOULSEEK, UfaSoft P2P Instant Messenger

□ Computação distribuída

- SETI@home (Search for Extra-Terrestrial Intelligence)
 - analisar o máximo de sinais de rádio captados por radiotelescópios
- Folding@home
 - Enovelamento de proteína, mal formações,..
- World Community Grid, OurGrid
 - Grade computacionais

Aplicações P2P

- Serviços Web
 - Akamai para CDNs
- Armazenamento distribuído
 - Freenet: para compartilhar arquivos e publicar sites livres (<http://freenetproject.org/>)
 - obfuscação do nó que origina uma consulta e do nó que responde
 - cada nó só conhece seus vizinhos na rota
- Anonimato, resistência a censura
 - Mixmaster remailers
 - Anonimato no envio de mensagens.
 - Publius: replicação e criptografia (<http://cs.nyu.edu/waldman/publius>)
 - Tor: defesa contra monitoramento da rede (<http://www.torproject.org/>)

Aplicações P2P

□ Aplicação de trabalho colaborativo

- Aplicações deste gênero são conhecidas como groupware application
- É um software que suporta diversos pontos, tais como:
 - Comunicação e coordenação de usuários
 - E-mail
 - Calendário
 - Espaço de trabalho
 - Listas de discussão
 - Gerência de documentos
 - Vídeo conferência
- Groove
 - Sistema colaborativo permitindo os usuários se comunicarem e interagirem diretamente de modo síncrono
 - Microsoft SharePoint Workspace

□ Outros...

Peer-to-peer: Classificação

- Modelos de Arquitetura P2P (1ª Classificação):
quanto a distinção das funções realizadas pelos nós
 - Descentralizado
 - Rede onde não há um nó central: todo nó na rede P2P tem a mesma função
 - Nós são autônomos e responsáveis por troca de recursos e por controle (gerenciamento) de recursos.
 - Semicentralizado
 - Existem nós com funções específicas
 - nó central para informações de controle
 - um conjunto de supernós que assume tais funções
 - » onde a queda de um supernó afeta apenas os nós inferiores ligados a ele

Peer-to-peer: Classificação

- Modelos de Arquitetura P2P (2ª Classificação): quanto ao modelo de busca
 - Busca centralizada
 - Rede com um ponto central de busca e nós consultam o ponto central
 - Troca de informações é diretamente entre os nós;
 - Busca por inundação
 - Normalmente a busca é limitada à vizinhança mais próxima do nó que fez a busca
 - busca é escalável, mas não é completa
 - Busca por tabela hash distribuída (DHT)
 - Rede onde os nós têm autonomia e usam uma tabela hash para separar o espaço de busca entre eles.
 - banco de dados tem duplas (chave, valor);
 - » chave: tipo conteúdo; valor: endereço IP
 - pares consultam BD com chave
 - » BD retorna valores que combinam com a chave
 - Exemplo: Kademlia (eMule)

Peer-to-peer: Classificação

□ Modelos de Arquitetura P2P (3ª Classificação)

○ P2P Centralizado

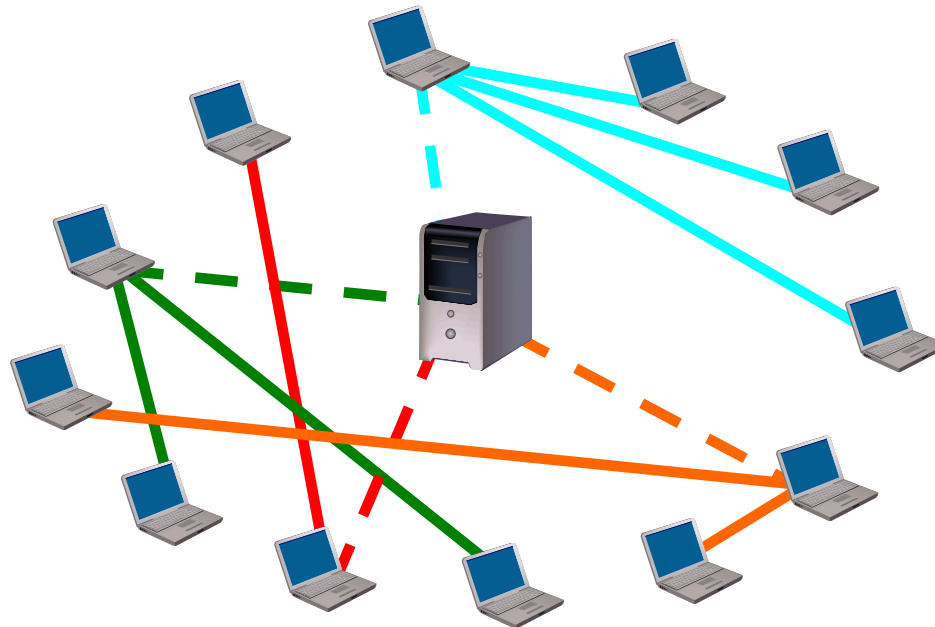
- Modelo centralizado (busca centralizada) onde um nó central mantém um índice central com informações atualizadas.
- Exemplos: sistemas de compartilhamento de arquivos Napster

○ P2P Descentralizada e Estruturada

○ P2P Descentralizada e não Estruturada

P2P Centralizado

- Representante mais conhecido: Napster
 - Solução híbrida: parte centralizada, parte distribuída
 - Índice centralizado, recuperação distribuída
- Serviço de diretório pode limitar desempenho



P2P Centralizado

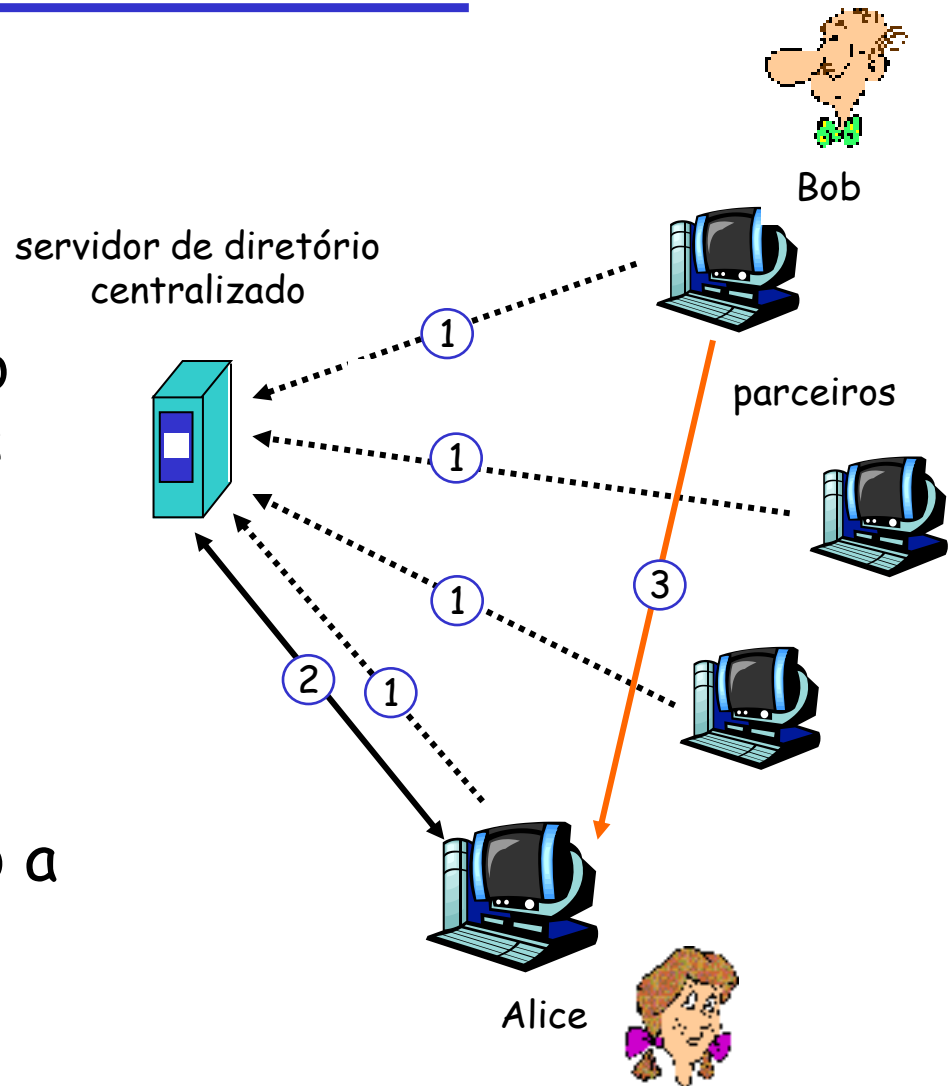
□ Exemplo: Napster

- Máquina de busca dedicada a encontrar arquivos MP3.
- Servidor central é utilizado para armazenar uma lista com as músicas disponibilizadas pelos usuários e onde elas estão localizadas.
- Programa cliente, instalado no computador dos usuários, faz uma consulta ao servidor Napster para obter informações sobre o arquivo desejado.
- Servidor Napster responde se existe o arquivo desejado e onde ele está localizado
 - Caso exista, uma conexão direta é estabelecida com o computador onde o arquivo está armazenado para que seja efetuado o *download*.

P2P Centralizado

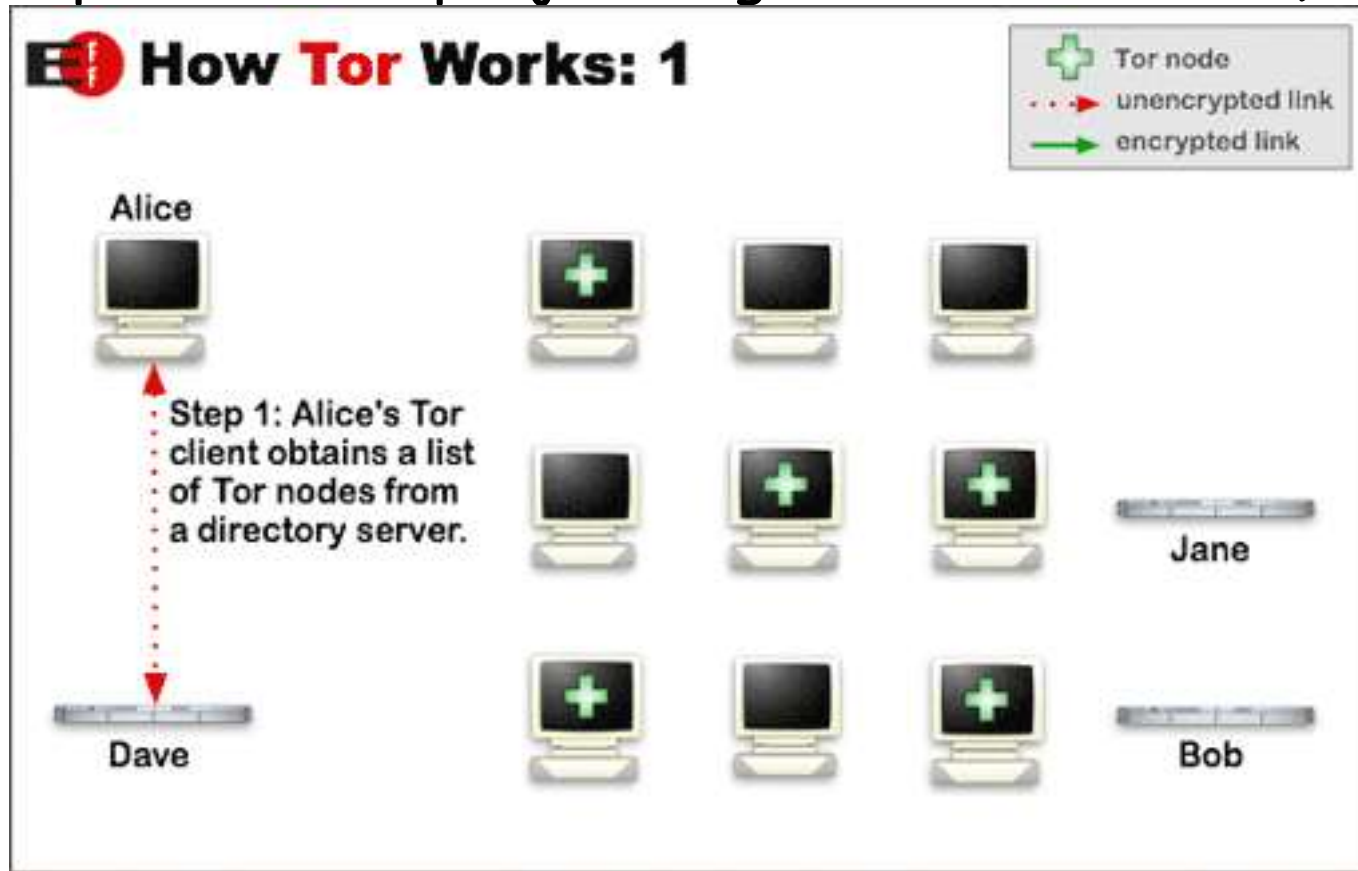
Projeto original do
"Napster"

- 1) quando um parceiro conecta ele informa ao servidor central o seu:
 - endereço IP
 - conteúdo
- 2) Alice consulta sobre a música "Hey Jude"
- 3) Alice solicita o arquivo a Bob



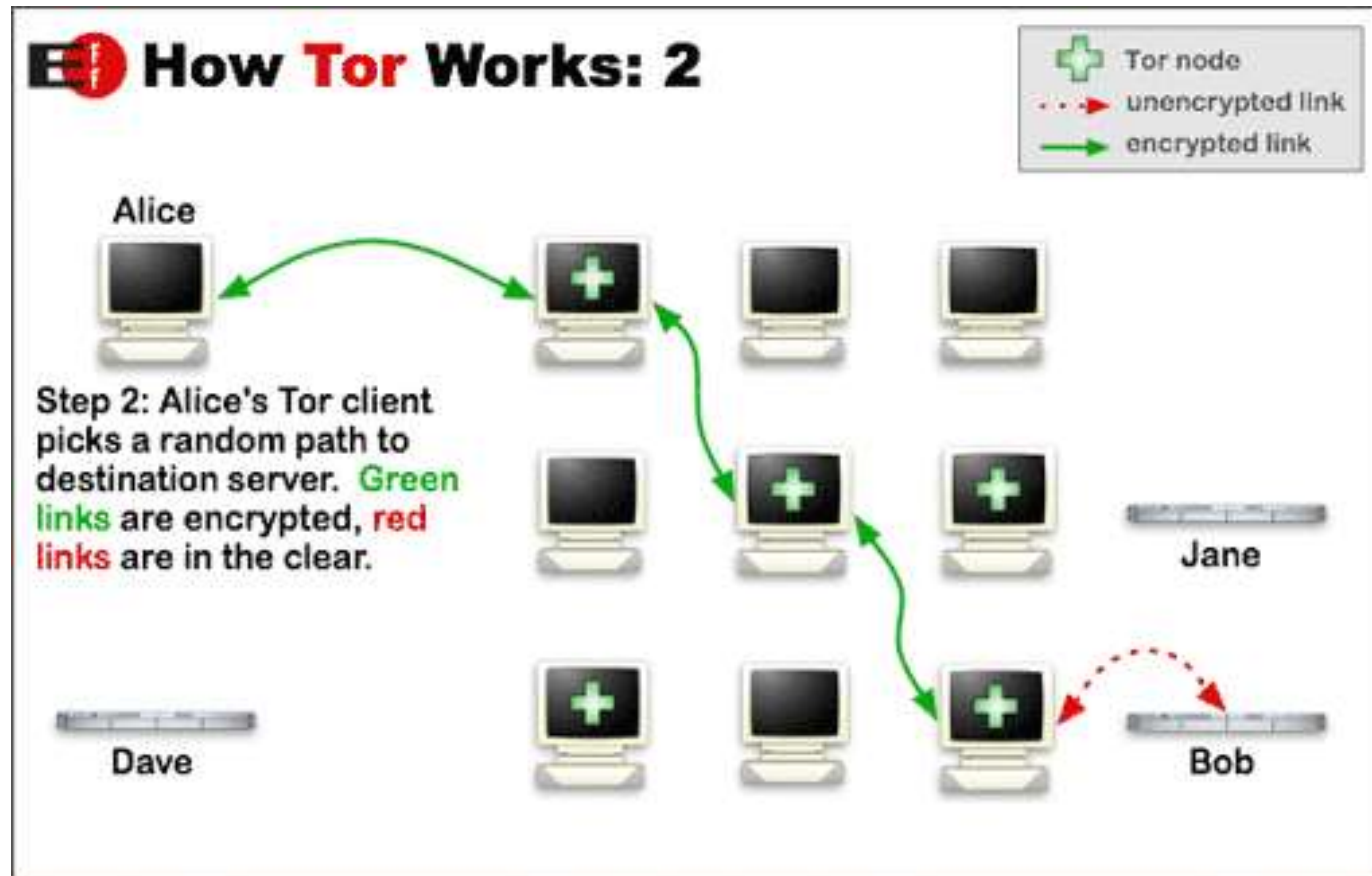
P2P Centralizado

- Como Tor trabalha para garantir anonimato (<http://www.torproject.org/overview.html.en>)



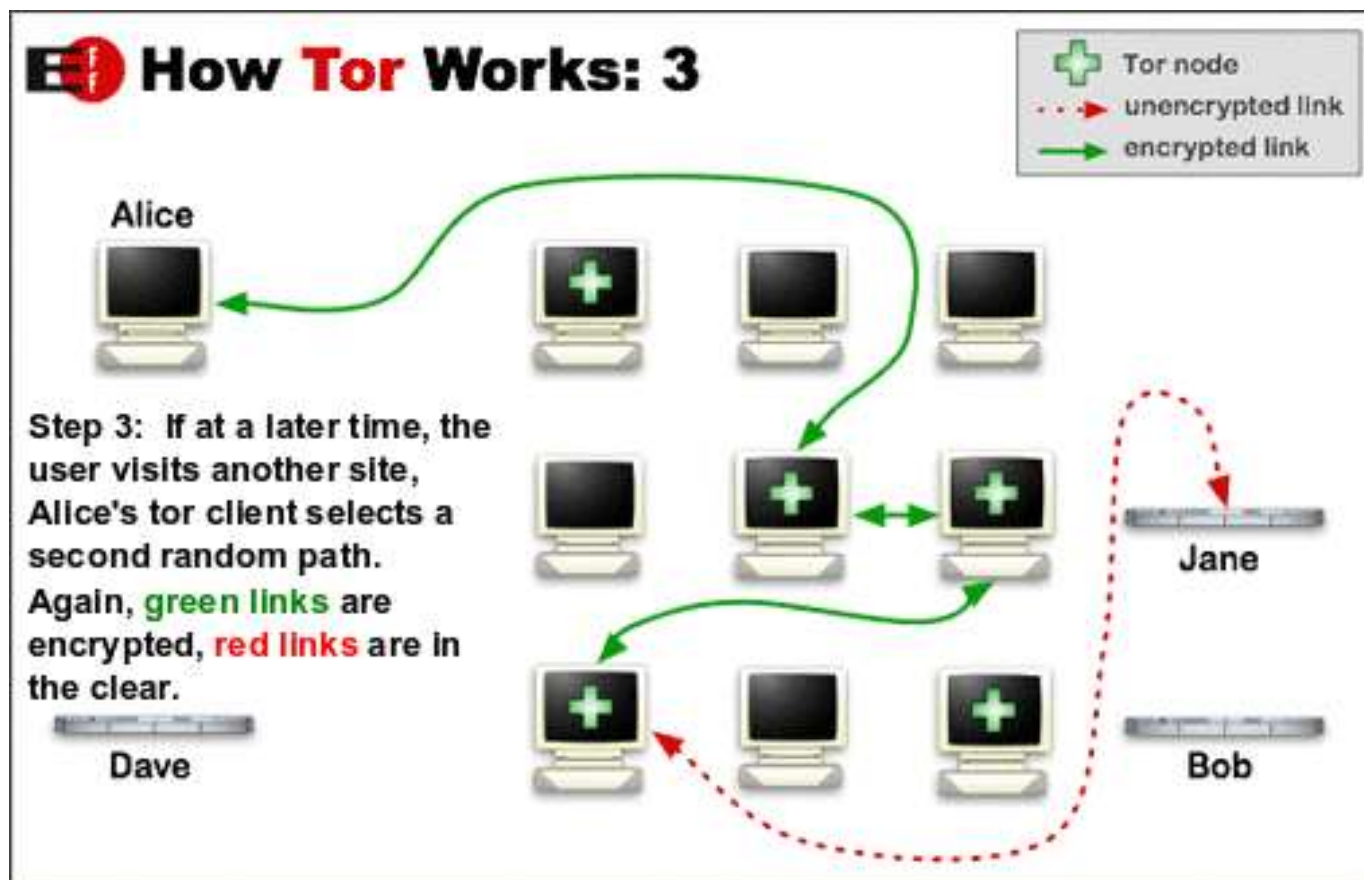
P2P Centralizado

- Como Tor trabalha para garantir anonimato



P2P Centralizado

- Como Tor trabalha para garantir anonimato



P2P Centralizado: Problemas

- Ponto único de falha
- Gargalo de desempenho
- Violação de Direitos Autorais

a transferência de arquivo é descentralizada, mas a localização do conteúdo é altamente centralizada.

Peer-to-peer: Classificação

□ Modelos de Arquitetura P2P (3ª Classificação)

○ P2P Centralizado

○ P2P Descentralizada e Estruturada

- Rede que não possui um servidor centralizado de diretório de informações
 - mas que tem uma estruturação entre os nós.
- Topologia da rede é controlada e os documentos/arquivos são posicionados em locais que posteriormente tornam fácil a sua localização.
- Em geral utiliza busca baseada em DHTs (Tabelas de espalhamento distribuídas)
 - Protocolos e implementações: Chord, Kademlia, Pastry, Tapestry e CAN, Viceroy
 - Aplicações: BitTorrent, eMule, LimeWire,

○ P2P Descentralizada e não Estruturada

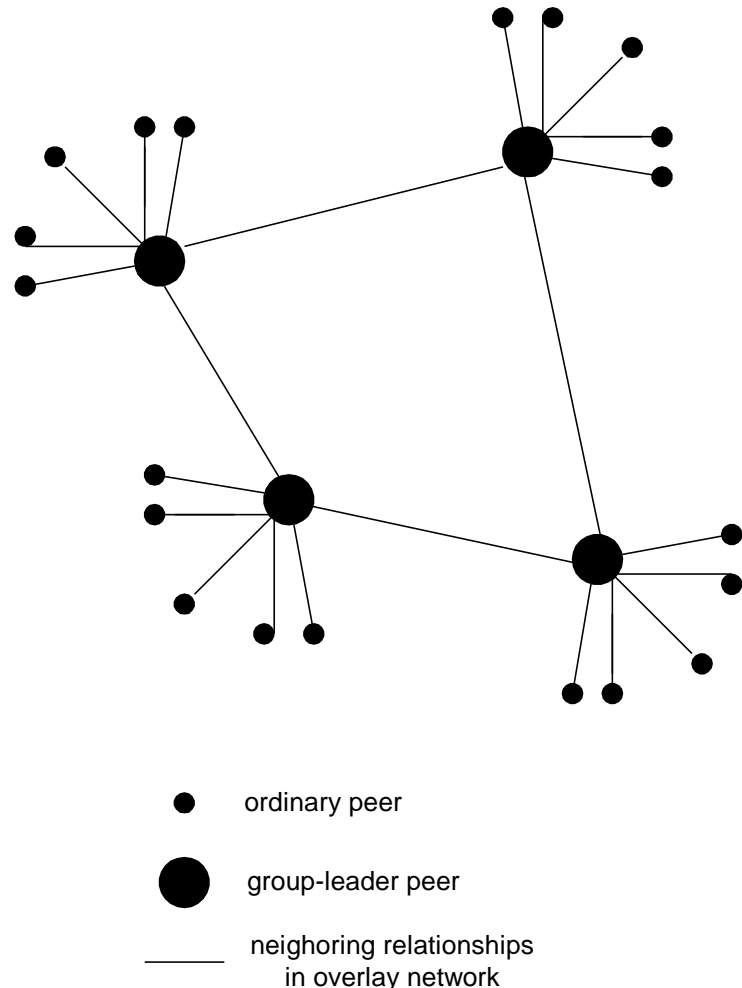
P2P Descentralizada e Estruturada

□ Tabela Hash Distribuída (DHT)

- Tabela Hash (tabela de espalhamento ou de dispersão)
 - Estrutura de dados especial que associa chaves de pesquisa a valores
- As tabelas hash são utilizadas para separar o espaço de buscas
 - Quando uma consulta é realizada, a chave é consultada e a busca é repassada para o nós responsável, que saberá o caminho para o local da informação

P2P Descentralizada e Estruturada

- Cada nó é um líder de grupo ou está associado a um líder de grupo (nó filho)
- Líder de um grupo segue o conteúdo de todos os seus filhos
- Parceiro consulta líder do grupo
 - Líder por sua vez pode consultar outros líderes de grupo.



P2P Descentralizada e Estruturada

□ Gnutella2

- Divide nós em dois grupos: folhas (*leaves*) e concentradores (*hubs*).
 - As folhas mantêm uma ou duas conexões com os concentradores,
 - Os concentradores aceitam milhares de folhas e muitas conexões de outros concentradores.
- Quando uma pesquisa é iniciada
 - nó obtém uma lista de concentradores se necessário, e contacta todos os concentradores da lista até o fim da lista ou se atingir um limite definido na consulta
 - permite um usuário encontrar um arquivo desejado facilmente sem sobrecarregar a rede
- Implementações: Limeware

P2P Descentralizada e Estruturada

□ KaZaA

- Utiliza o conceito de supernós para melhorar o desempenho da rede.
 - Uma máquina que participa da rede (nó) e que possui um alto poder computacional e rápidas conexões com a Internet.
 - Mantêm uma lista contendo os arquivos disponibilizados por outros usuários e o local onde eles estão armazenados.
- Quando uma busca é executada
 - Aplicação KaZaA procura primeiro no supernó mais próximo do usuário que iniciou a consulta, retorna um conjunto de respostas para o usuário e encaminha a consulta para outros supernós.
 - Uma vez localizado o usuário que possui o arquivo, uma conexão é estabelecida diretamente entre os peers para que seja efetuado o download.

P2P Descentralizada e Estruturada

Rede sobreposta

- parceiros são nós
- arcos entre os parceiros e seus líderes de grupo
- arcos entre alguns pares de líderes de grupo
- vizinhos virtuais

Nó de inicialização

- o parceiro que está se conectando ou é alocado a um líder de grupo ou é designado como líder

Vantagens da abordagem

- Ausência de servidor de diretórios centralizado
 - serviço de localização distribuído entre os parceiros
 - mais difícil de ser tirado do ar

Desvantagens da abordagem

- é necessário um nó de inicialização
- os líderes de grupo podem ficar sobrecarregados

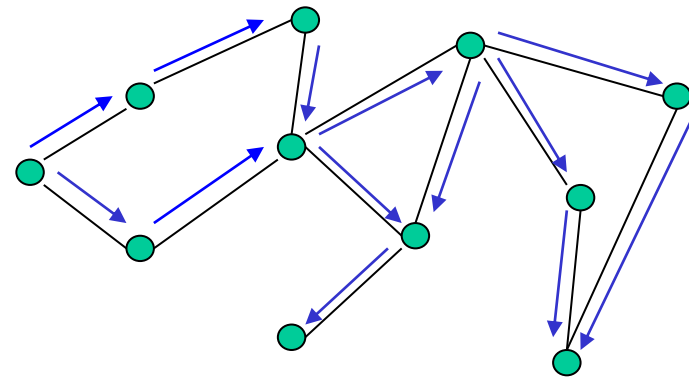
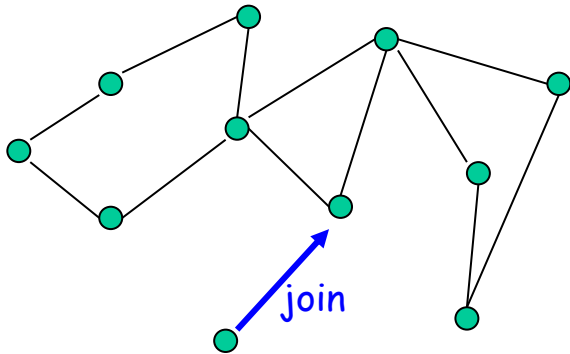
Peer-to-peer: Classificação

- Modelos de Arquitetura P2P (3ª Classificação)
 - P2P Centralizado
 - P2P Descentralizada e Estruturada
 - P2P Descentralizada e não Estruturada
 - Rede que não possui servidor centralizado, nem controle preciso sobre a topologia e localização/busca de documentos.
 - Compreende os dois tipos (descentralizado e semi-centralizado) e a busca por inundação.

P2P Descentralizada e não Estruturada: Inundação de consultas

□ Gnutella

- sem hierarquias
- usa nó de inicialização para aprender sobre os demais
- mensagem de união (join)
- Envia consulta aos vizinhos
- Vizinhos encaminham a consulta
- Se o parceiro consultado possui o objeto, ele devolve a mensagem para o parceiro que fez a consulta



P2P Descentralizada e não Estruturada: Inundação de consultas

Prós

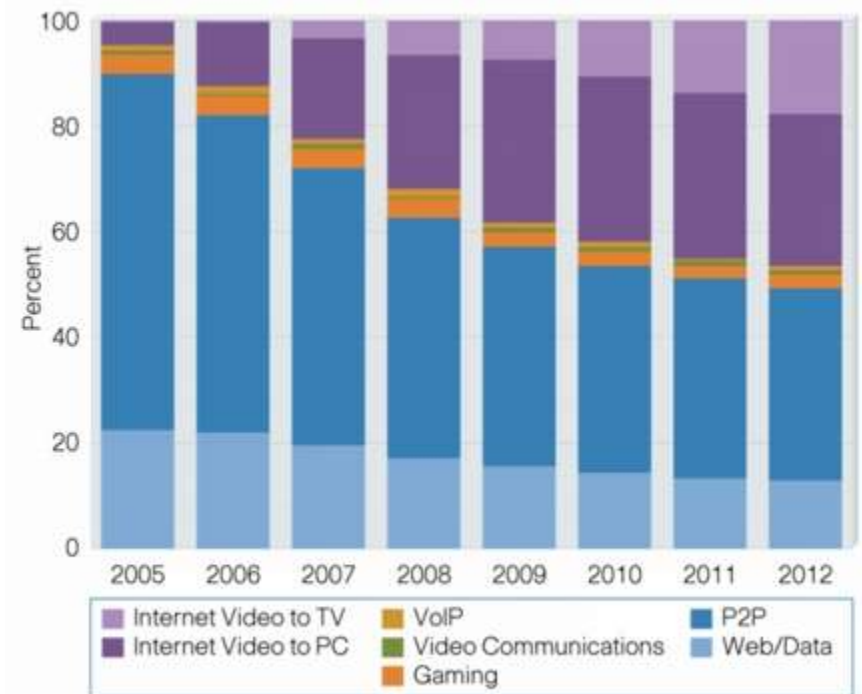
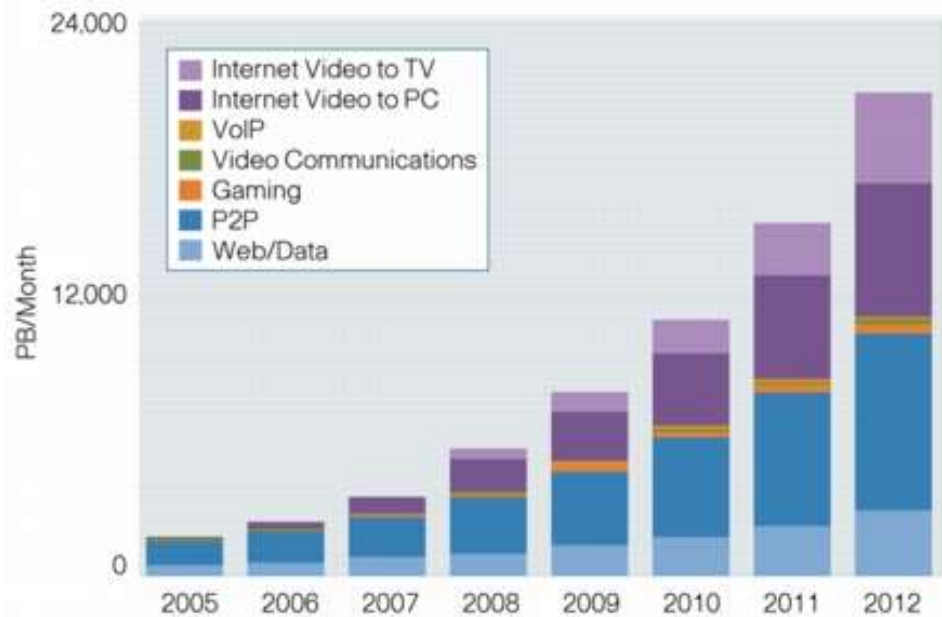
- ☐ os parceiros têm responsabilidades similares: sem líderes de grupo
- ☐ altamente descentralizado
- ☐ nenhum parceiro mantém informação de diretório

Contras

- ☐ tráfego excessivo de consultas
- ☐ raio da consulta: pode não ter o conteúdo quando estiver presente
- ☐ nó de inicialização
- ☐ manutenção da rede sobreposta

Peer-to-peer

- Um dos maiores consumidores de recursos da Internet

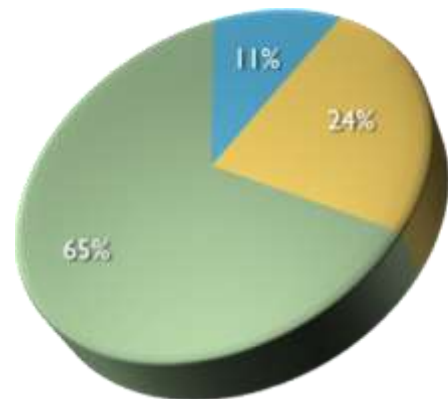
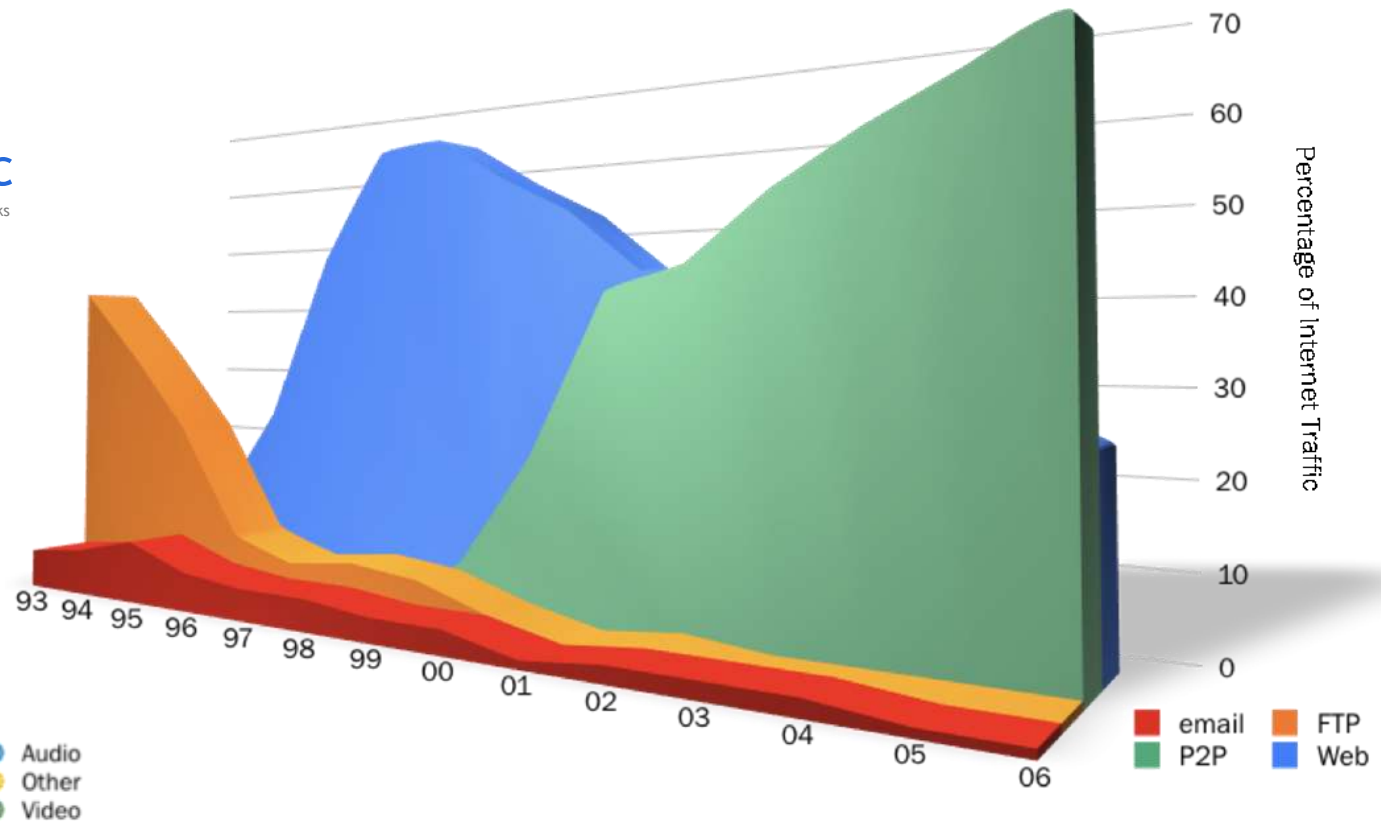


Source: Cisco, 2008

Peer-to-peer

- Um dos maiores consumidores de recursos da Internet

CacheLogic
Advanced Solutions for P2P Networks

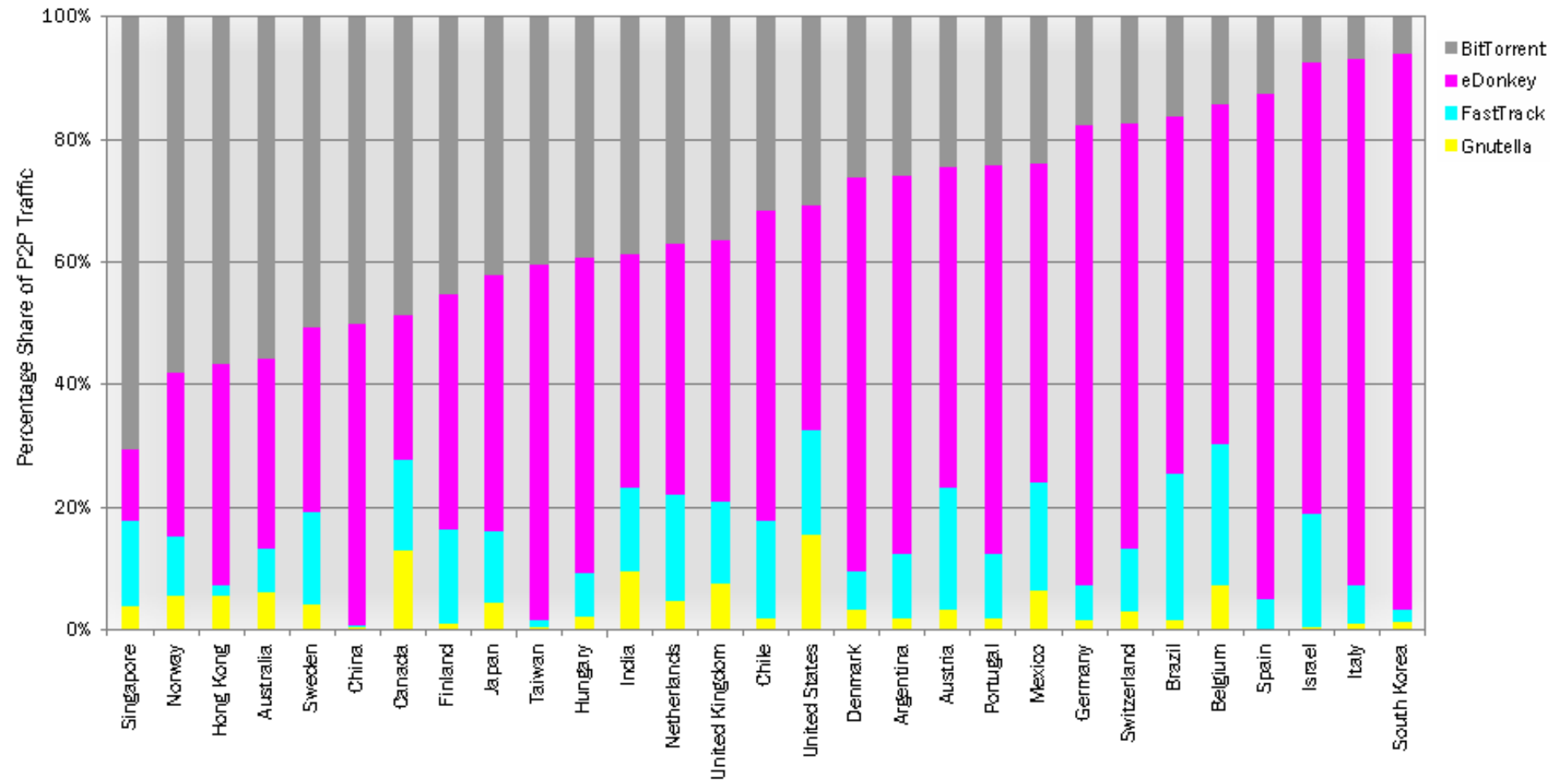


Peer-to-peer

□ Um fenômeno global

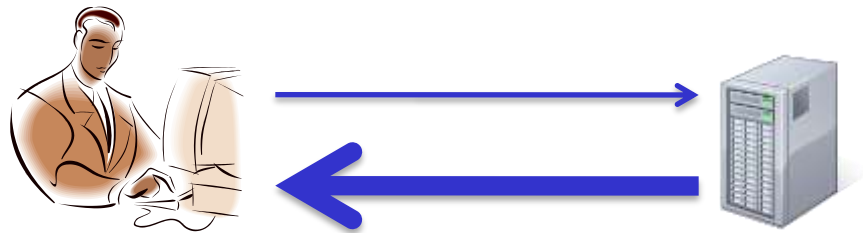
CacheLogic

Advanced Solutions for P2P Networks



P2P: Problema para a Internet

- P2P é tráfego Simétrico
 - Normalmente o tráfego da Internet é assimétrico



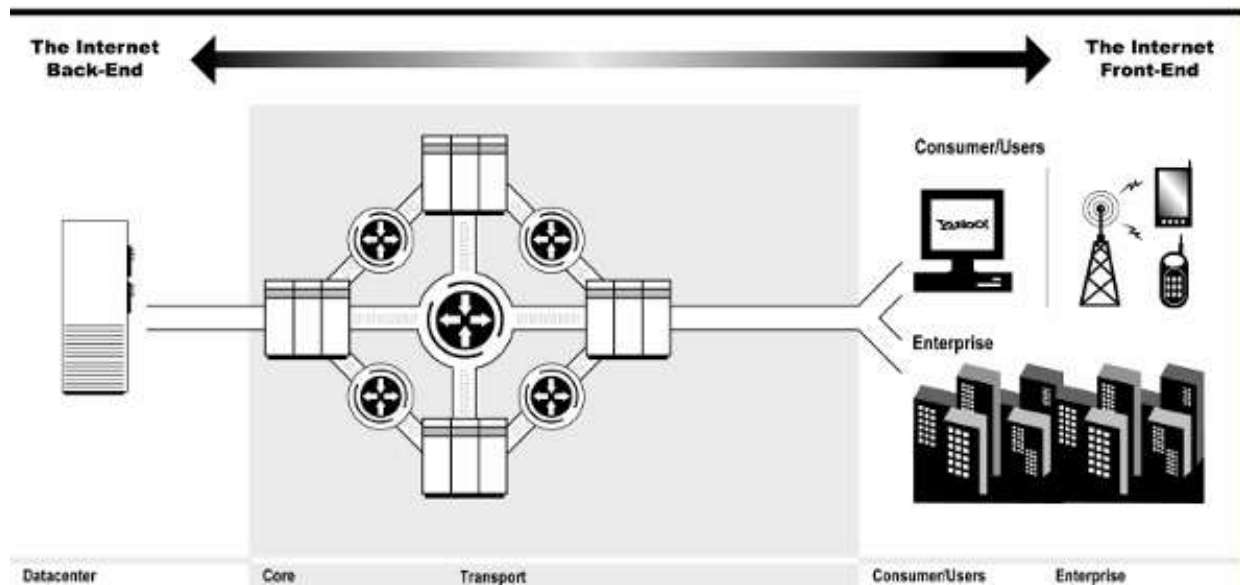
- P2P é uma aplicação simétrica



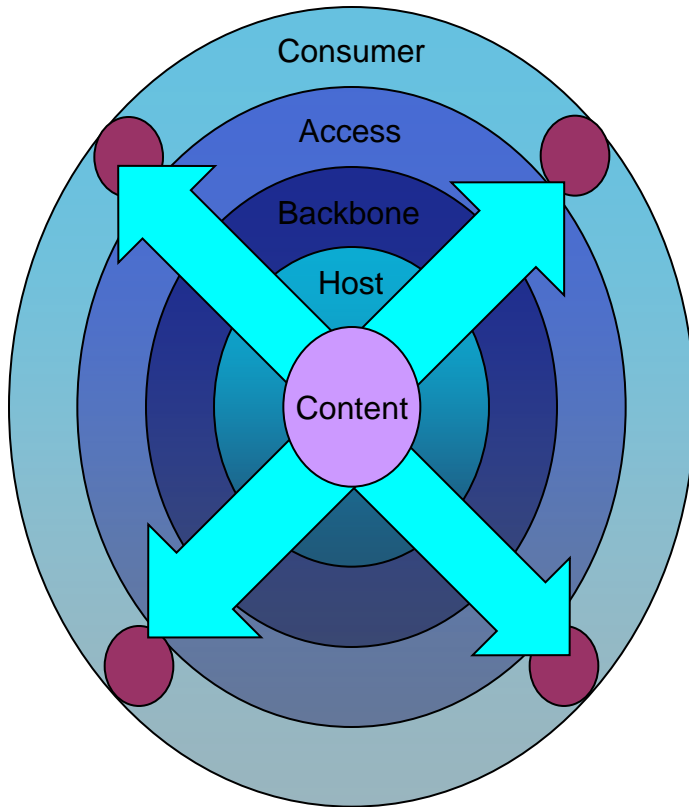
Internet de hoje não é projetada para tráfego simétrico

- A Internet foi construído com a premissa de distribuição núcleo-para-borda
 - Centros de dados de alta capacidade, High Capacity Data Centres, núcleos de alta capacidade, bordas pequenas assimétricas
 - Custo e desempenho otimizado pelos provedores
- P2P coloca nas bordas o conteúdo da distribuição nas redes de acesso

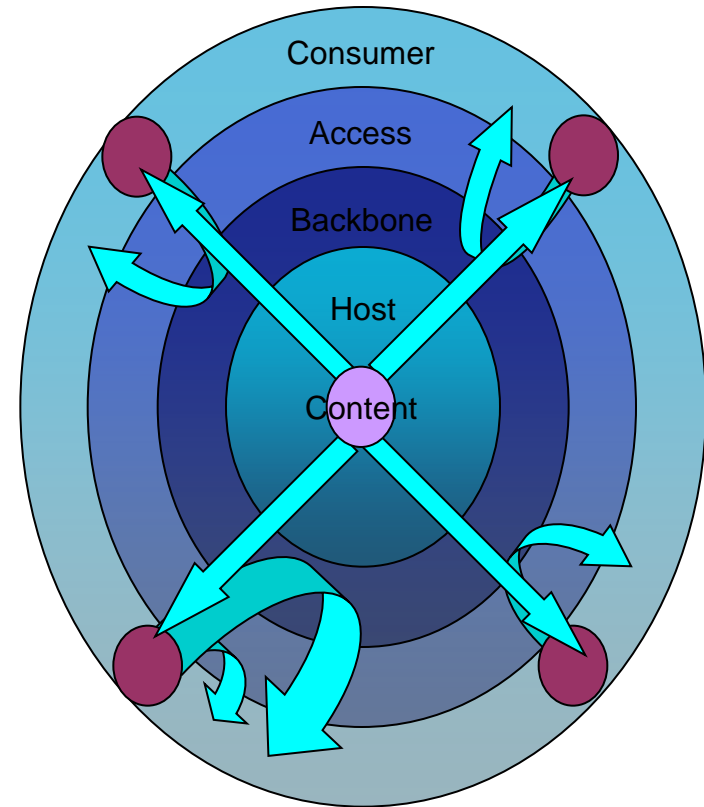
The Morgan Stanley Internet Ecosystem



Internet de hoje não é projetada para tráfego simétrico



Fluxos de Conteúdo tradicionais

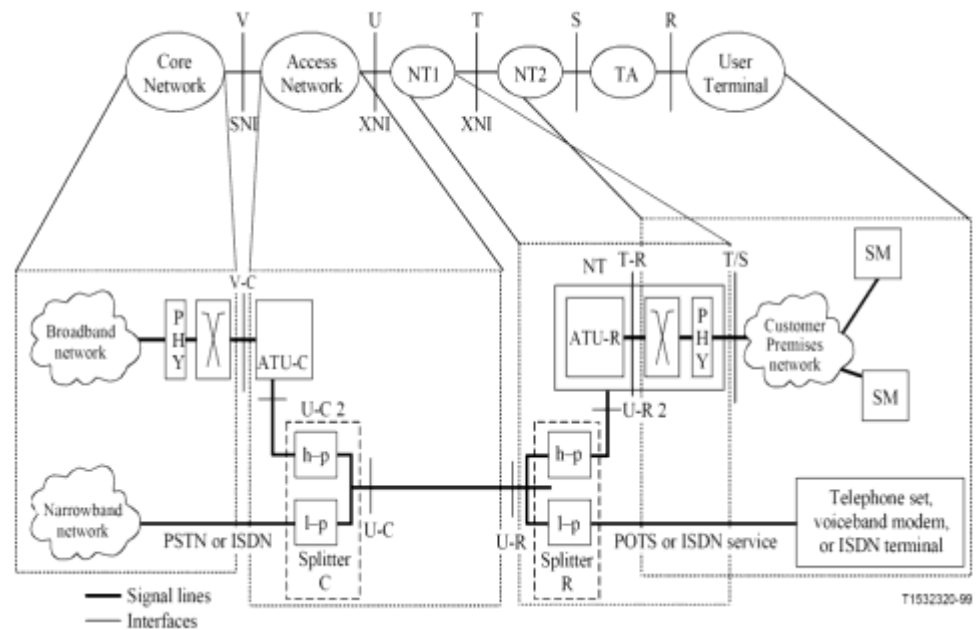


Fluxos de conteúdos no P2P

Redes de Acesso Banda Larga

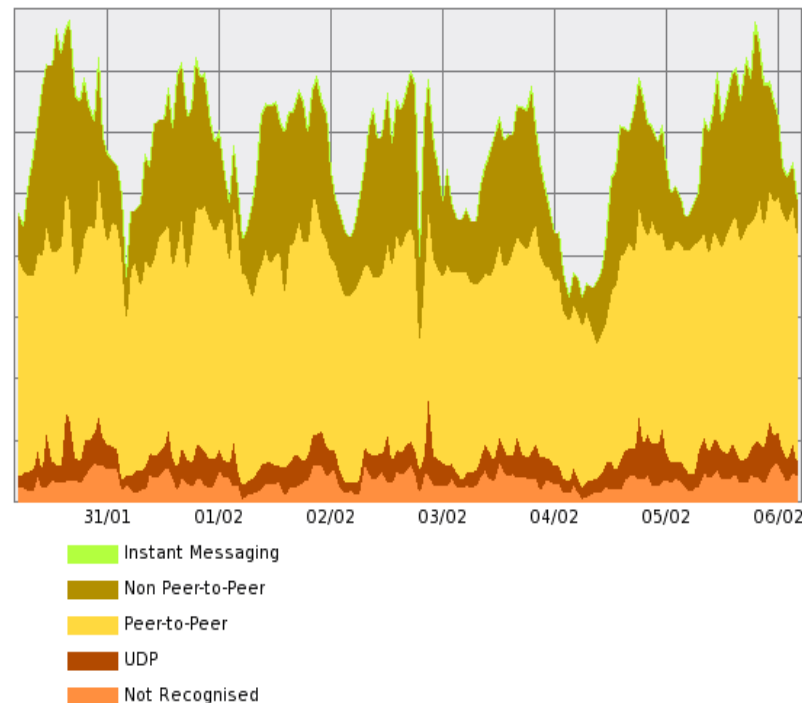
Projeto- DSL

- ADSL e Cable modem foram projetados para tráfego assimétrico
 - Web, VoD
- ADSL
 - Download < 7Mbps
 - Upload < 800kbps
 - 11:1
- ADSL2+
 - Download < 24Mbps
 - Upload < 1Mbps
 - 24:1
- Assimetria está aumentando



Impacto do P2P nos Provedores de Serviço - Tránsito/Peering

- 50-65% do tráfego downstream é P2P
- P2P usa muita banda
 - Em muitos países 90% do tráfego P2P atravessam as fronteiras internacionais
- Aplicações P2P consomem agressivamente a banda disponível



Impacto do P2P nos Provedores de Serviço - Transito/Peering

- Maior parte dos protocolos tradicionais são orientados a download (HTTP)
 - Isto torna o P2P um problema grave para os provedores de última milha
 - Responsável por 75-90% do tráfego upstream da rede de última milha

