

# Algoritmos de encaminhamento para comunicação multi-ponto e sua utilização em ambientes de grande escala

José Legatheaux Martins

Departamento de Informática da Faculdade de  
Ciências e Tecnologia da UNL

## Algoritmos **distribuídos** para **comunicação multi-ponto** e sua utilização em ambientes de **grande escala**

- **Trata-se de um problema fundamental em redes de computadores e sistemas distribuídos**
- **Com muitas aplicações concretas:**
  - Difusão multimédia e teleconferências
  - Difusão de objectos (documentos, eventos, ...)
  - Espaços de trabalho colaborativos (CSCW)
  - Sistemas tolerantes a falhas com base em replicação — comunicação em grupo
  - Localização de recursos
  - Distribuição de carga entre servidores equivalentes
  - Aquisição distribuída de dados (Telemetria, ...)

2

## Enunciado preliminar do problema

- **Enunciado informal**
  - Seja  $R$  uma rede com  $N$  nós, interligados por  $C$  canais, a cada um dos quais está associado um custo (um real não nulo e positivo)
  - Seja  $G$  um grupo de  $T$  nós e  $E$  um nó especial, designado por emissor
  - Pretende-se transmitir uma mensagem de  $E$  para os  $T$  membros de  $G$  de forma óptima em  $R$ , difundindo a mensagem por  $R'$ , uma sub-rede de  $R$
- Se  $T = N$  — Difusão generalizada ou *broadcasting*
- Se  $T < N$  — Difusão restringida a  $G$  ou *multicasting*
- Critérios de optimização mais frequentemente usados:
  - (C1) A soma dos custos dos canais de  $R'$  é mínima
  - (C2) O custo do caminho seleccionado para chegar de  $E$  a cada membro de  $G$  é mínimo
- Com estes critérios de optimização,  $R'$  é necessariamente uma árvore

3

## Algoritmos centralizados

- **Broadcasting e critério 1** — Árvore de cobertura mínima - *Minimal Spanning Tree* - *MST* (algoritmo bem conhecido)
- **Broadcasting e critério 2** — Árvore de caminhos mínimos - *Shortest Path Tree* - *SPT* (Dijkstra)
- **Multicasting e critério 1** — Árvore de Steiner — (*Minimal*) *Steiner Tree* (existem apenas boas aproximações da solução óptima)
- **Multicasting e critério 2** — Árvore de caminhos mínimos (*SPT* seguido da supressão sucessiva de nós folha e arcos não pertencentes a  $G$ )

4

## Formulações particulares do problema

| Número de emissores | Número de receptores | Caracterização sintética | Designação comum  |
|---------------------|----------------------|--------------------------|---|
| 1                   | N                    | 1 para N (todos)         | <i>Broadcasting</i> com um só emissor   |
| 1                   | $T < N$              | 1 para T (alguns)        | <i>Multicasting</i> com um só emissor   |
| 1                   | $K < T$              | 1 para K entre T         | <i>Multicasting</i> com um só emissor e com filtragem ( <i>publish/subscribe, geocasting, directed diffusion, ...</i> ) |
| 1                   | 1                    | 1 para 1 entre T         | <i>Anycasting</i>   |

5

## Formulações particulares ... (continuação)

| Número de emissores | Número de receptores | Caracterização sintética | Designação comum   |
|---------------------|----------------------|--------------------------|--|
| M                   | N                    | M para N (todos)         | <i>Broadcasting</i> com vários emissores                 |
| M                   | $T < N$              | M para T (alguns)        | <i>Multicasting</i> com vários emissores                 |
| M                   | $K < T$              | M para K entre T         | <i>Multicasting</i> com vários emissores e com filtragem |
| T                   | 1                    | T para 1                 | <i>Agregação e concentração de dados</i>                 |

6

## Agenda

- Algoritmos para difusão — *broadcasting*
- Algoritmos para difusão restrita — *multicasting*
- Casos especiais
- Aplicação ao modelo IP Multicasting e o estado da sua implementação
- Difusão em redes lógicas
- Observações finais

O objectivo é discutir a viabilidade dos algoritmos conhecidos e algumas lições que se podem tirar de algumas experiências concretas em grande escala

7

## Difusão generalizada ou *broadcasting*

- 1) Inundação com supressão de duplicados
  - Algoritmo tipo "força bruta", não óptimo mas muito robusto
  - As mensagens não duplicadas geralmente chegam por uma SPT
- 2) Selecção dos arcos de uma SPT através de inundação e calculo distribuído do custo de encaminhamento
  - Usada em LANs e MANs — IEEE 802.1D - Spanning Tree Protocol
- 3) Detecção de duplicados pelo Reverse Path Forwarding Check
  - Encaminha por uma SPT ou por uma Reverse SPT se os canais são assimétricos
  - Yogen K. Dalal and Robert M. Metcalfe. "Reverse Path Forwarding of Broadcast Packets," *Communications of ACM*, 21(12):1040-1048, 1978
- 4) Construção distribuída de uma R(SPT) enviando mensagens de adesão para o emissor
  - Evolução natural do algoritmo proposto no paper anterior

8

## Agenda

---

- Algoritmos para difusão — *broadcasting*
- Algoritmos para difusão restrita — *multicasting*
- Casos especiais
- Aplicação ao modelo IP Multicasting e o estado da sua implementação
- Difusão em redes lógicas
- Observações finais

9

## Algoritmos para *multicasting*

---

- Estes algoritmos envolvem um problema novo: a *gestão da filiação do grupo de receptores*
- O estado sobre a filiação pode ser mantido *centralizado, replicado por todos os nós, ou distribuído apenas pelos nós do grupo*
- Na forma clássica de apresentação, quase todos os algoritmos propostos dependem de os *nós da rede conhecerem uma forma óptima de encaminhamento ponto a ponto*
- Estes algoritmos envolvem, tal como os anteriores, a opção entre *uma árvore por emissor ou uma árvore partilhada por todos os emissores*

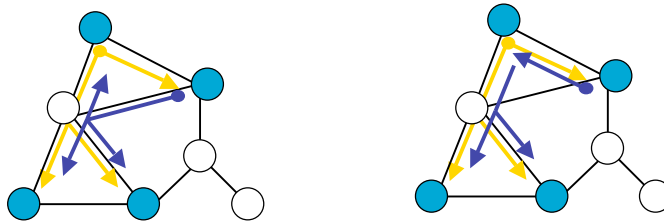
10

## Construção distribuída de árvores (R)SPT para Multicasting

- 1) Gestão centralizada da filiação no emissor e difusão sem estado na rede  
Proposto: por Dalal & Metcalf, aplicado ao IP em Lorenzo Aguilar, "Datagram Routing for Internet Multicasting." *Computer Communication Review* 14(2): 58-63 (1984)
- 2) Multicasting por inundação pelo emissor, detecção de duplicados por RPF e poda dos ramos inúteis  
Stephen E. Deering and David R. Cheriton. "Multicast Routing in Datagram Internetworks and Extended LANs." *ACM Transactions on Computer Systems (TOCS)*, 8(2):85-110, 1990
- 3) Filiação do grupo e configuração da rede integralmente replicadas em todos os nós da rede; cálculo da SPT pelos nós usando uma extensão do algoritmo de Dijkstra  
J. Moy. "Multicast Extensions to OSPF." IETF RFC 1584 (Proposed Standard), March 1994
- 4) Construção da árvore por iniciativa dos subscritores dirigindo mensagens de enxerto (*graft messages*) ao emissor  
Adaptado de: Yogen K. Dalal and Robert M. Metcalfe. "Reverse Path Forwarding of Broadcast Packets," *Communications of ACM*, 21(12):1040-1048, 1978

11

## Uma árvore por emissor versus partilha de uma árvore



- Mantém mais estado nos nós
- Mas cada um dos emissores encaminha por uma (R)SPT

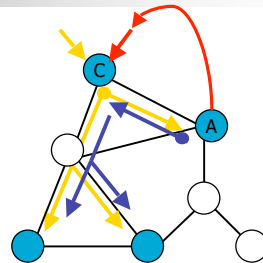
- Requer menos estado nos nós
- A melhor árvore é a de Steiner, mas esta varia com a filiação, o que introduziria instabilidade no encaminhamento

● pertence a G    ○ não pertence

12

## Do uso de uma árvore partilhada

- Neste caso, geralmente usa-se uma árvore construída por iniciativa dos subscritores com raiz num nó *C*, dito *centro* ou *core*. Como escolhê-lo ?
- Se a filiação e o padrão de tráfego variarem, só é possível uma solução baseada numa heurística
- Existe uma solução trivial se todos os emissores estiverem próximos de *C*
- A árvore pode ser usada de forma bidireccional ou todos os emissores injectarem o seu tráfego na árvore a partir do nó *centro* ou *core*



- Tráfego com origem em *C* →
- Tráfego com origem em *A* e encaminhado de forma bidireccional pela árvore →
- Tráfego dirigido a *C* por *A* para ser difundido para o grupo *G* pela árvore →

A. J. Ballardie, P. Francis and J. Crowcroft, "Core Based Trees", Proc. ACM SIGCOMM'93, San Francisco, CA, 1993

13

## Comparação (se vários emissores)

| Algoritmo                  | Sem estado na rede          | Calculo isolado por cada nó             | Inundação, teste RPF e poda                 | Subscrição junto do emissor       | Árvore partilhada                 |
|----------------------------|-----------------------------|---|---|-----------------------------------|-----------------------------------|
| Filiação                   | Centralizada nos emissores  | Replicada por inundação em todos os nós | Distribuída com recenseamento por inundação | Distribuída pelos nós de <i>G</i> | Distribuída pelos nós de <i>G</i> |
| Estado mantido por         | Emissores e mensagens       | Replicado por todos os nós              | Todos os nós da rede                        | Só nos nós das árvores            | Só nos nós da árvore              |
| Árvore(s)                  | SPTs com raiz nos emissores | SPTs com raiz nos emissores             | SPT(R)s com raiz nos emissores              | SPT(R)s com raiz nos emissores    | SPT(R) com raiz no nó centro      |
| Complexidade computacional | Baixa                       | Elevada                                 | Baixa                                       | Baixa                             | Baixa                             |
| Complexidade do controlo   | Baixa                       | Elevada                                 | Elevada                                     | Baixa                             | Baixa                             |
| Escala de aplicação        | Baixa                       | Baixa                                   | Baixa                                       | Baixa                             | Aceitável                         |

14

## Utilização conjunta dos algoritmos (PIM-SM)

---

- Dado um nó designado por "rendez-vous" (RV) é possível construir uma árvore que cubra os membros de  $G$  por iniciativa dos subscritores
- Os nós emissores distintos de RV podem começar por lhe dirigir o tráfego através de um túnel.
- Se um emissor  $E$  particular se revelar importante, por um dado critério, os membros de  $G$  podem decidir construir uma nova árvore de difusão com raiz em  $E$
- Neste último caso, o nó RV e a árvore inicial servem apenas para que o emissor seja conhecido dos membros de  $G$ , evitando assim o uso de inundação para esse efeito

D. Estrin, D. Farinacci, A. Helmy, D. Thaler, S. Deering, M. Handley, V. Jacobson, C. Liu, P. Sharma, and L. Wei. "Protocol Independent Multicast-Sparse Mode (PIM-SM): Protocol Specification." IETF RFC 2362, June 1998

15

## Agenda

---

- Algoritmos para difusão — *broadcasting*
- Algoritmos para difusão restrita — *multicasting*
- **Casos especiais**
- Aplicação ao modelo IP Multicasting e o estado da sua implementação
- Difusão em redes lógicas
- Observações finais

16



## Casos especiais

---

### 1 para 1 entre T ou Anycasting

- Problema típico das CDNs e da localização de recursos; realizável a nível aplicativo (e.g. Akamai)

M. Gritter, and D.R. Cheriton, "An Architecture for Content Routing Support in the Internet," USENIX SIT, 2001

### 1 para K entre T ou difusão filtrada

- Entregar a um subconjunto de  $G$  interessados em mensagem que satisfaçam o predicado  $P$

Sérgio Marco Duarte, "DEEDS — A Distributed and Extensible Event Dissemination Service," Doctoral Dissertation, FCT/UNL, 2005

### Difusão filtrada seguida de agregação (K para 1 )

- Paradigma "Difusão Dirigida" para redes de sensores
- C. Intanagonwiwat, R. Godivan and D. Estrin, "Directed Diffusion: a Scalable and Robust Communication Paradigm for Sensor Networks," ACM Mobicom, Boston, 2000

17

## Agenda

---

- Algoritmos para difusão — *broadcasting*
- Algoritmos para difusão restrita — *multicasting*
- Casos especiais
- **Aplicação ao modelo IP Multicasting e o estado da sua implementação**
- Difusão em redes lógicas
- Observações finais

18

## Encaminhamento IP Multicasting

| Protocolo        | Algoritmo   | Intra AS | Inter AS | Estado real (*)  |
|------------------|---|----------|----------|--|
| DVMRP            | Inundação, RPF, poda                                    | Sim      |          | Histórico  |
| MOSPF            | Calculo isolado por cada nó                             | Sim      |          | Inactivo   |
| CBT              | Subscrição num nó core                                  |          |          | Nunca entrou em produção                               |
| BGMP             |   |          | Sim      | Nunca entrou em produção                               |
| PIM-DM           | Inundação, RPF, poda                                    | Sim      |          | Utilização bastante reduzida                           |
| PIM-SM           | Difusão a partir do nó RV e subscrição junto do emissor | Sim      | Sim      | Utilização generalizada Intra AS isto é, dentro do ISP |
| PIM-SSM e IGMPv3 | Subscrição junto do emissor que é único                 | Sim      | Sim      | Utilização generalizada Intra AS isto é, dentro do ISP |
| Bidir-PIM        | PIM-SM com árvore bidireccional                         | Sim      | Sim      | Em normalização (*)                                    |
| MBGP+PIM-SM+MSDP | Vários  |          | Sim      | Utilização reduzida ou em regressão                    |

(\*) P. Savola, "Overview of the Internet Multicast Routing Architecture," IETF Draft, Work in progress, draft-ietf-mboned-routingarch-02.txt, October 2005

19

## Problemas e impasses do IP Multicasting

- **Identificação:** os endereços IP Multicast são identificadores únicos — tal exige mecanismos de geração e localização actualmente ausentes da arquitectura base da Internet
- **Protecção e segurança:** o modelo da Internet continua a não incorporar segurança na base — o modelo IP Multicast exacerbou este problema
  - Os problemas da identificação e segurança são parcialmente resolvidos por PIM-SSM
- **Viabilidade económica:** o custo de um grupo depende do seu âmbito e do custo do seu *control plane*
- **Problemas arquitecturais:** não existe uma solução única (a la TCP) para o transporte e as implementações integralmente a nível da periferia são pouco eficazes

Hugh W. Holbrook and David R. Cheriton. "IP Multicast Channels: EXPRESS Support for Large-scale Single-source Applications." In SIGCOMM '99

J. Saltzer, D. Reed and D. Clark, "End-to-end arguments in system design," ACM Transactions on Computer Systems, 2(4):195-206, 1984

20

## Agenda

---

- Algoritmos para difusão — *broadcasting*
- Algoritmos para difusão restrita — *multicasting*
- Casos especiais
- Aplicação ao modelo IP Multicasting e o estado da sua implementação
- Difusão em redes lógicas
- Observações finais

21

## Redes lógicas (CDNs, Overlay, P2P, ...)

---

- **Visão concreta.** Uma rede lógica é um conjunto de nós de nível aplicacional, interligados por canais implementados por ligações de transporte (UDP, TCP, HTTP, ...)
  - Os nós da rede lógica estão na periferia da Internet
  - Estabelecem ligações entre si de forma arbitrária ou segundo uma estratégia pré-definida
- **Visão abstracta.** Uma rede lógica é um grafo, com  $N$  nós, com  $N \cdot (N-1)$  arcos, cada um dos quais com um custo **finito** associado
  - A rede diz-se homogénea se todos os canais têm o mesmo custo
  - e heterogénea no caso contrário

Qualquer rede lógica tem subjacente uma **rede de suporte** heterogénea e de capacidade limitada — este último caso, o normal, é mais difícil de tratar teoricamente

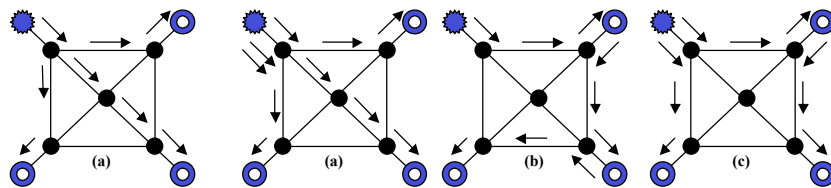
22

## Que canais dos $N.(N-1)$ possíveis seleccionar ?

- **Redes lógicas ad hoc ou não estruturadas** (sistemas P2P de primeira geração, e.g. Gnutella) — os canais são, grosso modo, os que “aparecerem”
- **Redes lógicas estruturadas logicamente** (DHTs — *Distributed Hash Tables*, e.g. Chord, Pastry, Tapestry, ...) — os canais são escolhidos em função dos identificadores dos nós; a função custo é definida no espaço dos identificadores
- **Redes lógicas optimizadas em função da capacidade real disponível via a rede de suporte** — os canais de maior capacidade têm prioridade
- **Redes lógicas aleatórias** — os canais variam constantemente pois as mensagens são difundidas através de caminhos escolhidos aleatoriamente

23

## Novas funções de custo: **pressão e extensão**



☀ Emissor (lógico)    ⦿ Receptor (lógico)    ● Nó (suporte)    → Cópia da mensagem M

- Se se pretender comparar a solução de encaminhamento multi-ponto na rede lógica, com a solução baseada numa árvore óptima na rede de suporte, os seguintes novos critérios têm cabimento:
  - **Pressão (*stress*)**: o número máximo de mensagens que atravessam cada canal
  - **Extensão (*stretch*)**: o quociente entre custo real que a rede lógica exige e o custo que seria possível na rede de suporte
- **A avaliação destes parâmetros nem sempre é fácil**

24

## Que novos algoritmos de encaminhamento ?

---

- Sistemas “começar pela malha” (*mesh first*)
- Sistemas “começar pela árvore” (*tree first*)
- Encaminhamento usando a topologia de uma DHT (solução implícita)
- Encaminhamento epidémico ou probabilístico

25

## Exemplos de sistemas *mesh first*

---

- **Narada (CMU)** cria uma malha aleatória mas depois otimiza-a para o cenário de aplicação e para o algoritmo “inundação, RPF, poda”
- **Scribe (MS-C.)** constrói uma “árvore de difusão, com raiz no emissor, por iniciativa dos subscritores” sobre a DHT Pastry
- **SplitStream (MS-Cambridge)** generaliza a aproximação seguida pelo sistema Scribe. Para distribuir melhor a carga pelos nós introduz uma floresta de árvores Scribe na DHT Pastry
- **DEEDS (FCT/UNL)** promove a heterogeneidade disponibilizando os algoritmos mais adequados para cada caso com base numa aproximação de redes activas

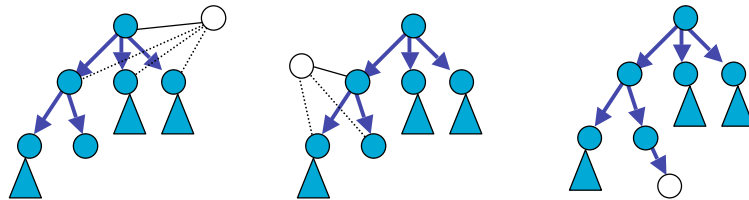
A novidade está nos algoritmos para formação da malha e não nos algoritmos de encaminhamento multi-ponto. A qualidade da solução do ponto de vista da rede de suporte está dependente da qualidade da malha e mais secundariamente do algoritmo.

26

## Exemplo de um sistema *tree first*

John Jannotti, David K. Gifford, et al. "Overcast: Reliable Multicasting with an Overlay Network." (OSDI 2000) —MIT (a pedido da Cisco)

- Cada nó que se junta a um grupo começa por estabelecer um primeiro canal até ao emissor (a raiz da árvore) e avalia o seu custo
- Em seguida, avalia o custo dos canais até aos filhos da raiz
- Em função dos resultados, o nó reposiciona-se (sucessivamente) de forma a que o canal que o liga à árvore seja sempre o de menor custo, sem aumentar demasiado o custo do canal para a raiz, e tentando minorar a *pressão (strees)* sobre a rede, mesmo sacrificando a *extensão (stretch)*.



27

## Outros exemplos de sistemas *tree first*

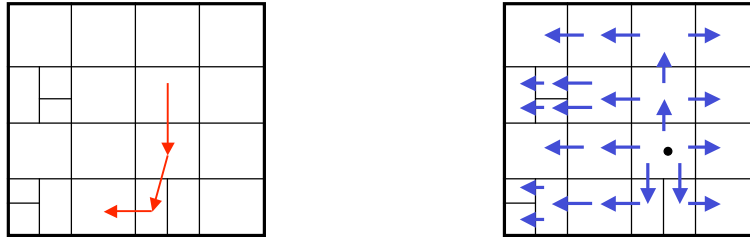
- **NICE (Maryland)** — os nós organizam-se em *clusters* de  $K$  a  $2.K$  nós; o nó centro do *cluster* assume o papel de um nó interior da árvore de difusão (bidireccional); o critério de *clustering* é a latência
- **ALMI (Washington-StLouis)** — o nó *rendez-vous* tem conhecimento completo da árvore corrente e distribui pelos nós o encargo de testarem novos canais aleatoriamente; em função dos resultados, a árvore é reorganizada através de um algoritmo centralizado de cálculo de uma aproximação da árvore de Steiner; a árvore é usada de forma bidireccional
- **Bullet (Duke)** — começa por construir uma árvore mas depois os nós estabelecem ligações horizontais e a árvore degenera numa malha para distribuir a carga entre os diferentes ramos

Todos estes sistemas dispõem de um *control plane* relativamente complexo para adaptar e otimizar dinamicamente a(s) árvore(s) de difusão e necessitam de detectar ciclos durante a reconfiguração das mesmas

28

## Exemplo de difusão por inundação otimizada topologicamente numa DHT (CAN - Berkley)

Sylvia Ratnasamy, et al. "ApplicationLevel Multicast Using Content-Addressable Networks." *Networked Group Communication*, volume 2233 of *Lecture Notes in Computer Science*, pages 14–29, Springer Verlag, 2001



Uma avaliação por simulação revela uma solução melhor que a difusão por vários canais ponto a ponto do emissor para cada receptor, mas apresenta *stress* e *stretch* absolutamente decepcionantes e necessita de detectar duplicados

29

## Difusão por inundação epidémica

Ken Birman et al., "Bimodal multicast," *ACM Transactions Computer Systems*, 1999 (Cornell)

P. Eugster, R. Guerraoui, et al. "Lightweight Probabilistic Broadcast," *ACM Transactions Computer Systems*, 2003 (EPFL+MS-C)

- Cada mensagem a difundir é passada a  $K$  membros do grupo, escolhidos aleatoriamente, cada um desses passa-a outros  $K$ , .... Exige um método de detecção de duplicados
  - Requer a troca periódica de mensagens de controlo com membros também escolhidos aleatoriamente que permitem conhecer novos membros e detectar mensagens ainda não conhecidas
  - Foram desenvolvidos modelos e simulações que permitem ter garantias probabilísticas da fiabilidade e do desempenho da entrega
  - Foram introduzidas muitas optimizações de complexidade de memória e de privilegiar os canais de maior capacidade (MS-C, EPFL, FCUL, Uminho, ...)
- Tem a robustez dos percursos aleatórios de grafos o que lhe confere propriedades muito interessantes

30

## As redes lógicas são uma solução viável ?

---

- Os algoritmos e sistemas que se configuram em função da rede de suporte, têm custos de controlo muito acrescidos por não terem acesso ao estado da rede de suporte e melhoram muito se poderem colocar nós em pontos específicos da rede de suporte
- As DHTs são excelentes do ponto de vista da escala para a localização e encaminhamento para identificadores. Está por provar que na sua forma actual consigam fazer difusão com *stress* e *stretch* razoáveis
- Sistemas como SplitStream e Bullet usam florestas de árvores ou árvores complementadas com malhas para tirarem partido da capacidade de *ingress* ("upstream") dos participantes num sistema P2P. É possível fazê-lo com menos complexidade ?
- A difusão epidémica tem excelentes propriedades de robustez e adaptação. De forma natural explora todos os canais *ingress* ("upstream") disponíveis. É aplicável de facto em grande escala ?

31

## Agenda

---

- Algoritmos para difusão — *broadcasting*
- Algoritmos para difusão restrita — *multicasting*
- Casos especiais
- Aplicação ao modelo IP Multicasting e o estado da sua implementação
- Difusão em redes lógicas
- **Observações finais**

32



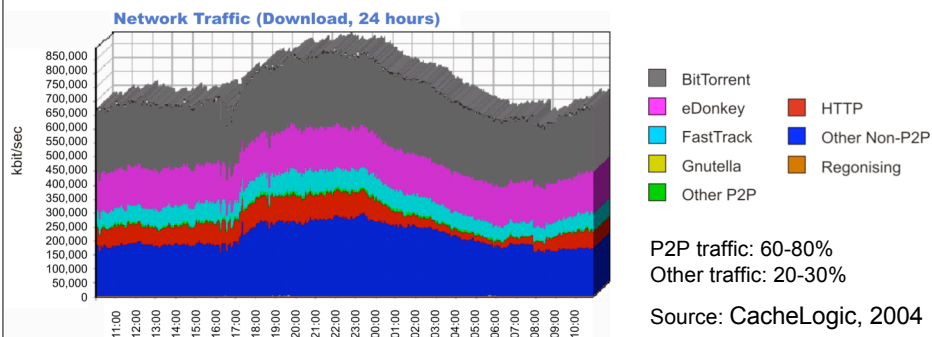
## A situação actual

- Actualmente muitos ISPs usam IP Multicasting baseado nos protocolos PIM-SM e SSM para implementarem IP TV dentro da sua rede
- Fora das redes de investigação, alguns grandes clientes poderão contratar um serviço específico a um ISP e uma multinacional poderá contratá-lo a vários; a activação será manual e o serviço caro
- No entanto, os clientes normais não podem utilizar IP Multicasting para a suas aplicações ("Johnny still can't multicast")
- Mas podem usar uma solução de nível aplicacional baseada numa rede lógica.

33

## O Exemplo BitTorrent

- Aplicação P2P de *broadcast* maciço de ficheiros "grandes"
- Implementação pragmática de um algoritmo de difusão epidémica
- Muito popular e eficaz com números e sucesso impressionantes (mais de 10<sup>7</sup> milhões de utilizadores em alturas de pico)



34

## Como o BitTorrent usa a rede de suporte

### R. Bindal et al., "Improving Traffic Locality in BitTorrent via Biased Neighbor Selection," ICDCS 2006, Lisboa

Estudo por simulação usando o código real do BitTorrent. Rede com 14 ISPs interligados em estrela; cada ISP com 50 utilizadores de canais assimétricos de 1 Mbps *download* / 100 *upload* Kbps

| ISP interconnect bottleneck | Download time 95 <sup>th</sup> percentile | Traffic redundancy | Neighbor biased selection               | Download time 95 <sup>th</sup> percentile | Traffic redundancy |
|-----------------------------|---|--------------------|---|---|--------------------|
| No bottleneck               | 1.4                                       | 47                 | Regular BitTorrent                      | 1.4                                       | 47                 |
| 2.5 Mbps                    | 1.6                                       | 32                 | Biased (1 outside the ISP, N clustered) | 1.2                                       | 3                  |
| 1.5 Mbps                    | 2.1                                       | 25                 |   |   |                    |
| 0.5 Mbps                    | 3.5                                       | 22                 |   |   |                    |

*Download time* é o tempo normalizado de fazer o *download* do ficheiro com a capacidade de *upload*, isto é, a 100 Kbps

35

## Cenário idealizado: junção dos dois mundos

- Os utilizadores estão dispostos a pagar para fazerem o *download* rápido de ficheiros a partir de fornecedores
- Um grupo IP Multicast é activado se o número de utilizadores sobe para além de um certo limite e suporta a difusão dos blocos usando técnicas já desenvolvidas para canais assimétricos de *broadcasting* (estudados em sistemas móveis)
- Um algoritmo epidémico de cooperação P2P é usado para compensar os blocos em falta, como proposto inicialmente por K. Birman no protocolo bimodal multicast
- Resultado: *download time* < 1, *traffic redundancy* próxima de 1, canais de interligação e rede de acesso do ISP racionalmente usados
- Mas muitos problemas concretos por resolver

36

## Porque não ?

---

- A Internet é cada vez mais heterogénea via as novas redes móveis de acesso; a periferia está cada vez mais "às cegas" (e.g. TCP móvel, mobilidade, ...). **A Internet do modelo uniforme e ponto a ponto está um pouco ultrapassada ....**
- Do ponto de vista da comunicação multi-ponto parece continuar a ser necessário **poder enriquecer a rede em pontos específicos** (e.g. poder replicar pacotes onde faz sentido, compensar erros onde faz sentido, receber indicações do *core* sobre os custos, com segurança e com um modelo económico viável)
- Do ponto de vista da comunicação multi-ponto existem **bloqueios arquitecturais** ao nível da **identificação**, ao nível de **segurança**, ao nível do **modelo económico**, ao nível dos serviços e estrutura do *core* e **na forma como a periferia se relaciona com o core**

37