internet society of australia

isoc-au

Internet Content Filtering

What it is – and i

STILL isn't…

Paul Brooks

Director – ISOC-AU

pbrooks@layer10.com.au

Official Text….

DBCDE

INTERNET SERVICE PROVIDER CONTENT FILTERING PILOT
TECHNICAL TESTING FRAMEWORK

November 2008

Agenda

Filtering Content from…

- The InterWeb
- The Internet
- Not-the-InterWeb

…in 15 minutes…

Official Text….

"To the extent possible, the aim is to test a range of different types of filtering including:

- ACMA blacklist filtering only (for a blacklist of up to 10,000 URLs); or

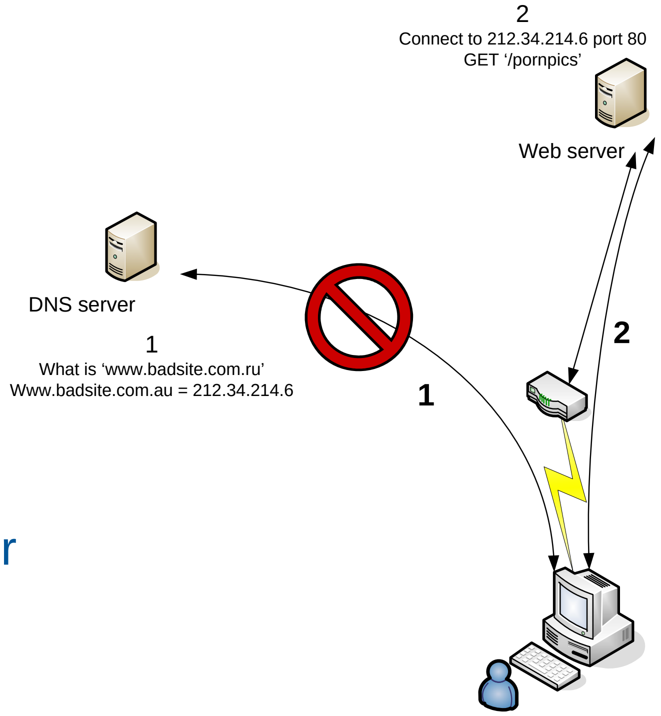
Currently ~1300 URLs

- ACMA blacklist filtering plus the filtering of other content using different approaches to filtering which would, for example, include:
  - Index filtering of different sized blacklists;
  - Dynamic analysis filtering;
  - IP versus URL filtering;
  - DNS poisoning. "

WWW filtering only

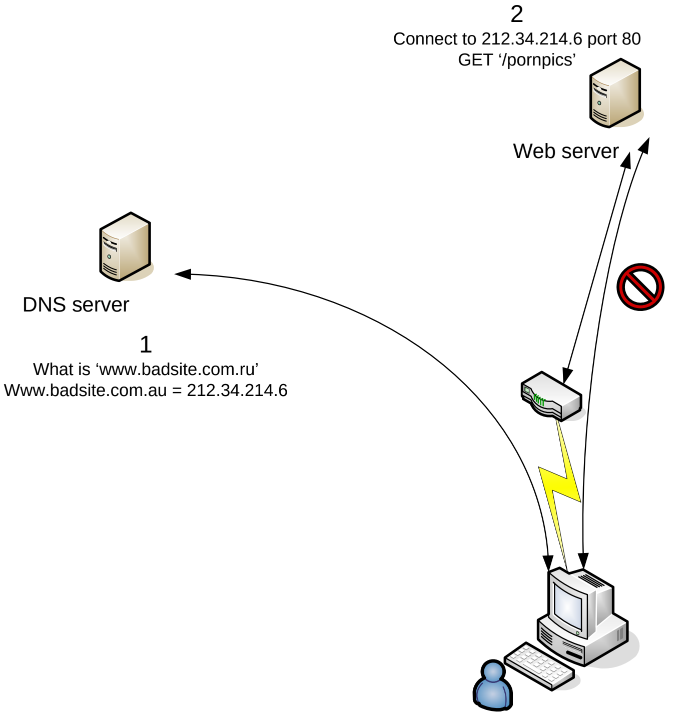
DNS Poisoning (WWW request)

- User asks for ‘www.badsite.com.ru/pornpics’
- DNS request is independent of request for content
- Block DNS request
  - ISP first has to know www.badsite.com.ru is to be blocked – needs prior notification
  - Thousands of names can point to same address
  - User can bypass DNS request by just using the IP address in the browser
  - User can specify an international DNS server
  - Blocks every website on that machine name – www.bigpond.com? Massive collateral damage


2
Connect to 212.34.214.6 port 80
GET ‘/pornpics’
Web server
DNS server
1
What is ‘www.badsite.com.ru’
Www.badsite.com.au = 212.34.214.6
1
2

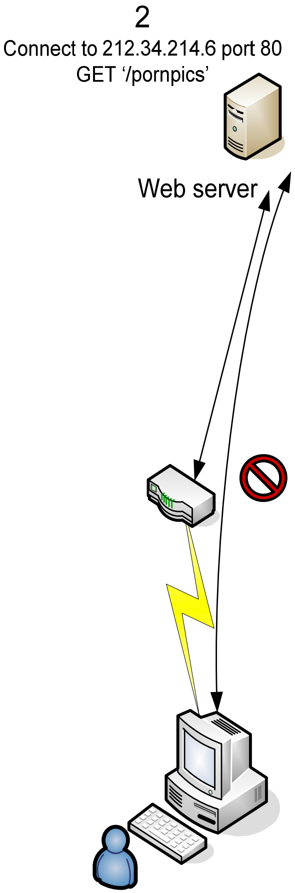
IP vs URL filtering

- User asks for ‘www.badsite.com.ru/pornpics’
- Block IP address
  - ISP first has to know 212.34.214.6 is to be blocked – needs prior notification
  - Thousands of sites can be hosted on the same IP address – massive collateral damage
  - HTTP can use any port number, not just ‘port 80’ – under control of the site – so have to block all connectivity for all applications
  - Golden opportunity for Denial of Service – deliberately host inappropriate content on www.bigpond.com/user/fakename
- **December 2008 – UK ISPs block Wikipedia**


2
Connect to 212.34.214.6 port 80
GET ‘/pornpics’
Web server
DNS server
1
What is ‘www.badsite.com.ru’
Www.badsite.com.au = 212.34.214.6

Dynamic Analysis Filtering

- User asks for ‘www.badsite.com.ru/pornpics’
- Deep Packet Inspection
  - Attempts to look deep into packet contents to identify application, try to classify packets in ‘real time’ and identify signatures of ‘bad stuff’
  - e.g. reconstruct images on the fly – look for excessive flesh tones
- However…
  - Doesn’t scale – bandwidth required and number of images to be analysed increasing faster than Moore’s Law
  - Still images being surpassed by streaming movies
  - Forces all content through a gatekeeper box – poor reliability, poor performance
  - Indiscriminate Blocks medical sites, school swimming carnivals, baby photos…..
  - Defeated by Secure HTTP – encrypted webpages, identical to online banking



Fundamental Issues

- ISP-level filters can’t tell if you are accessing photos of your own kids, or someone else’s
- ISP-level filters can’t tell the age of the user requesting the photo – can only be used for verified illegal content, not for ‘inappropriate’ content
- Easily circumvented using public anonymous proxy sites – the URL the ISP sees is completely different from the eventual URL being accessed
- Easily circumvented by encrypted webpages – HTTPS, SSL encryption

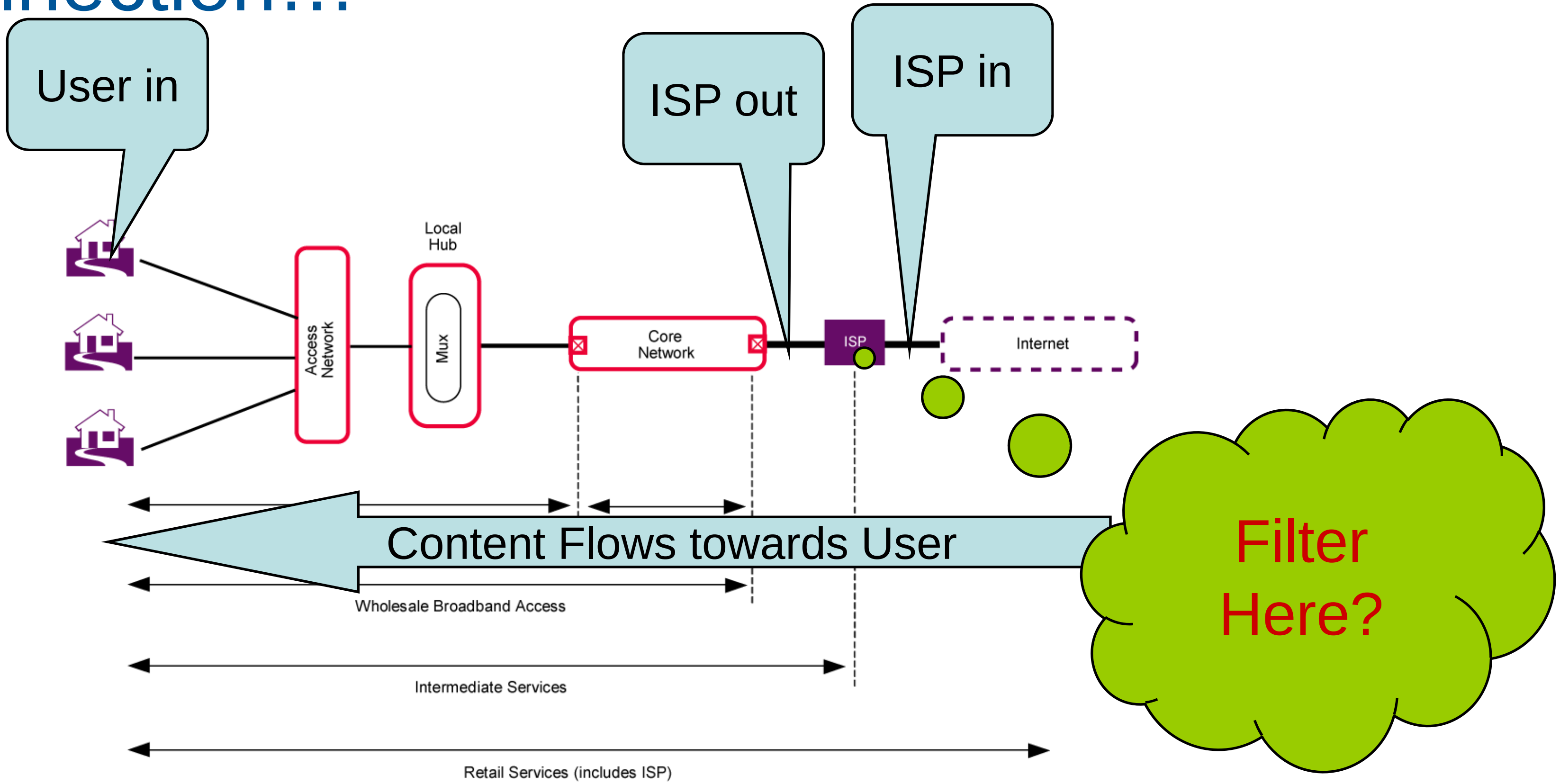
Official Text….

A key component of the Pilot is scalability. The Pilot is, however, limited to the networks and filtering solutions of the participants and their willingness to provide relevant technical information.

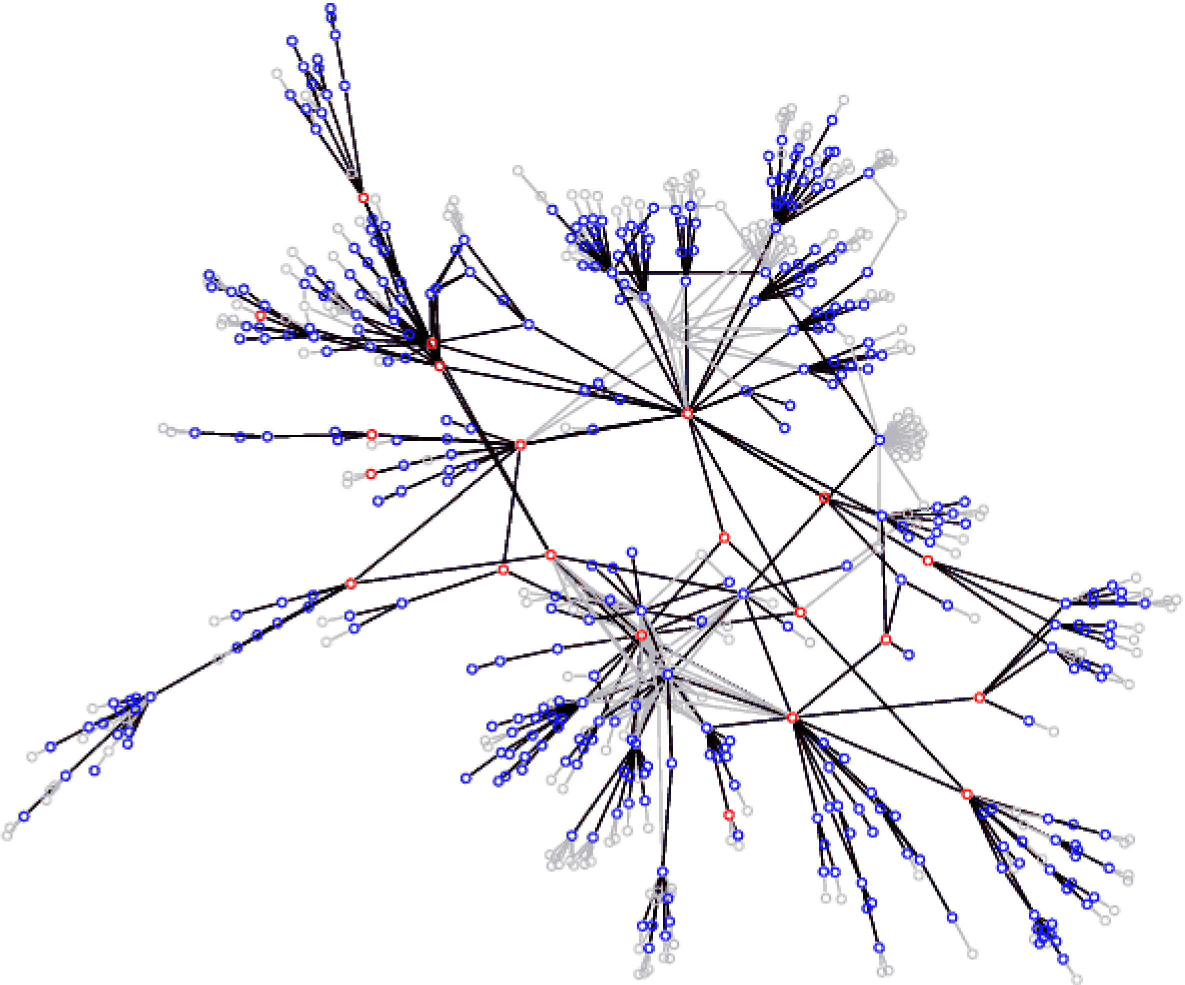
Ideally the Pilot will involve a representative cross section of the industry, for example tier 1, 2 and 3 ISPs, metropolitan, regional and remote ISPs including mobile, wireless and satellite internet service providers and broadband and dial up customers.

The Pilot aims to assess delivery across a variety of internet delivery mediums (wireless and copper to HFC) ranging in speed from 56Kbps through to 12Mbps.

Simplistic ISP Network Diagram for end-user connection…



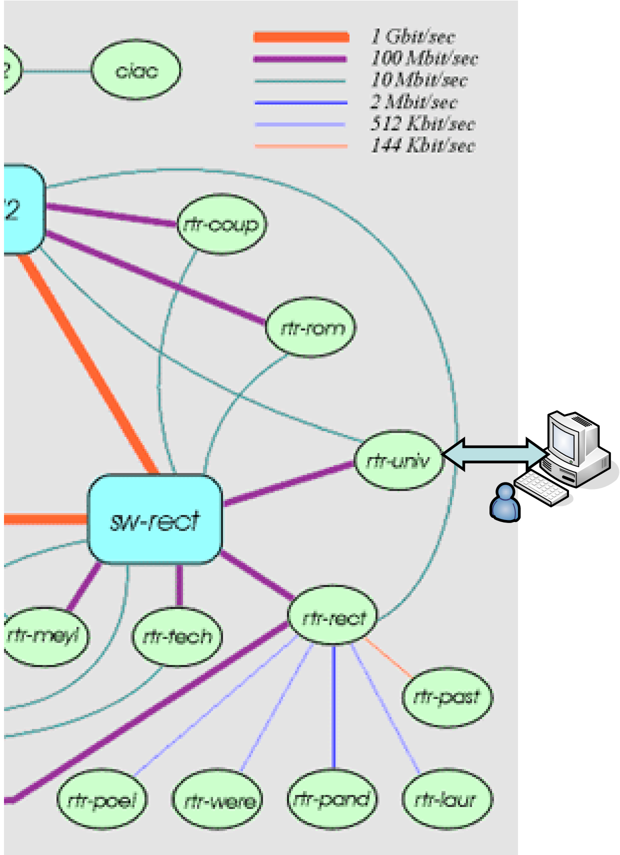
Network Filter – where?



“Medium ISP” Topology Diagram

Network Filter – where?

- Upstream Provider Link?
  - Most ISPs have 3 – 30 upstream providers, often in different states
  - Peering Points – no ‘provider’
- In the ISPs Core?
  - Single point of failure
  - Poor performance of ‘trombone’ traffic paths
  - Huge traffic increase – multiply cost of longhaul transmission
  - Misses content generated by other users of the same ISP
- At the PoP
  - Most ISPs will need 5 - 30 gatekeeper boxes!
- great idea if you sell gatekeeper boxes, not practical in real networks



Official Text – DBCDE, 11/2/09

- "Arrangements for the first phase of the live pilot have been finalised with six ISPs while consultations continue with a number of other ISPs that have applied to take part.
- "The initial round of ISPs are Primus Telecommunications, Tech 2U, Webshield, OMNIconnect, Netforce and Highway 1."

Telstra are conducting their own internal trial, iiNet may be interested in future.

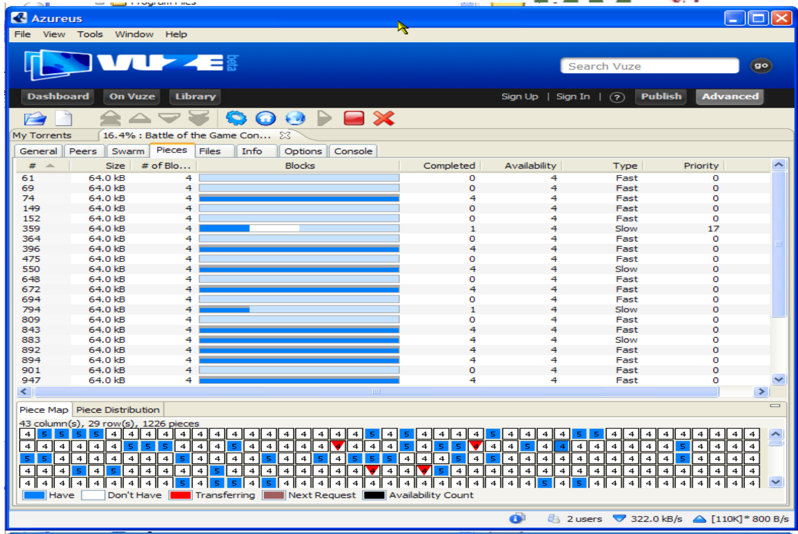
No Tier 1, no mobile/3G/4G, no satellite, no HFC, few national own networks.

Not-the-InterWeb

- The Internet, and inappropriate content, is not just exchanged using HTTP (WWW)
  - Email
  - USENET aka ‘Network News’
  - Peer-to-peer – e.g. bittorrent
  - RSS - Podcasts
  - Instant Messenger – MSN, Yahoo, etc
  - Skype
  - …..and many others

Peer-to-peer transfers

- Files broken into hundreds of small pieces
- Central ‘torrent servers’ only have lists of ‘peers’ with pieces, no content themselves
- Collect pieces from hundreds of PCs while serving your pieces to hundreds that need them
- Looks to the ISP network like hundreds of random connections to other random IP addresses
- Can be encrypted - no way of knowing what is inside the files
- No way to analyse files until all pieces are downloaded
- Cannot be blocked once started – the swarm of active sharers is self-sustaining

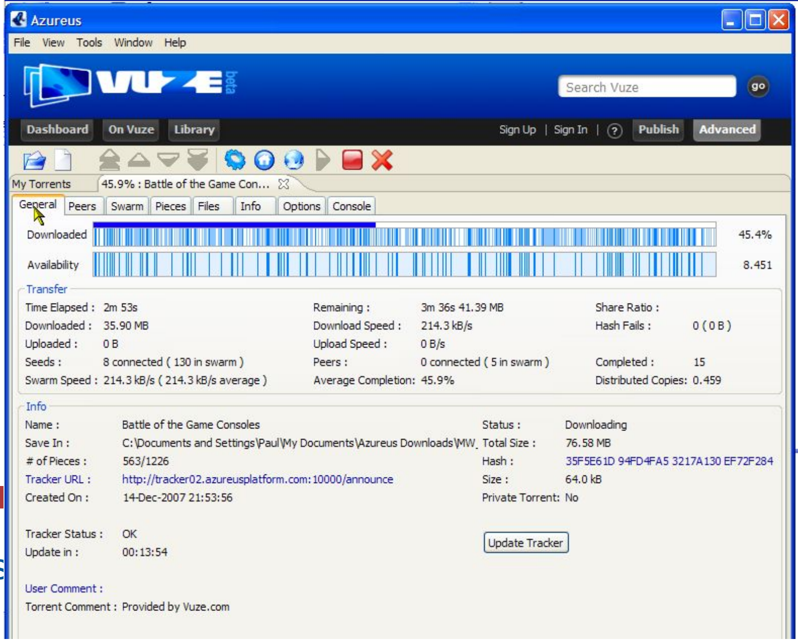


#	Size	# of Blo...	Blocks	Completed	Availability	Type	Priority
61	64.0 kB	4		0	4	Fast	0
69	64.0 kB	4		0	4	Fast	0
74	64.0 kB	4		4	4	Fast	0
149	64.0 kB	4		0	4	Fast	0
152	64.0 kB	4		0	4	Fast	0
359	64.0 kB	4		1	4	Slow	17
364	64.0 kB	4		0	4	Fast	0
396	64.0 kB	4		4	4	Fast	0
475	64.0 kB	4		0	4	Fast	0
550	64.0 kB	4		4	4	Slow	0
648	64.0 kB	4		0	4	Fast	0
672	64.0 kB	4		4	4	Fast	0
694	64.0 kB	4		0	4	Fast	0
794	64.0 kB	4		1	4	Slow	0
809	64.0 kB	4		0	4	Fast	0
843	64.0 kB	4		4	4	Fast	0
883	64.0 kB	4		4	4	Slow	0
892	64.0 kB	4		4	4	Fast	0
894	64.0 kB	4		4	4	Fast	0
901	64.0 kB	4		0	4	Fast	0
947	64.0 kB	4		4	4	Fast	0

43 column(s), 29 row(s), 1226 pieces

🟦 Have ⬜ Don't Have 🟥 Transferring 🟫 Next Request ⬛ Availability Count

2 users 322.0 kB/s [110K]* 800 B/s



Downloaded 45.4%

Availability 8.451

Transfer

Time Elapsed : 2m 53s
Downloaded : 35.90 MB
Uploaded : 0 B
Seeds : 8 connected (130 in swarm)
Swarm Speed : 214.3 kB/s (214.3 kB/s average)

Remaining : 3m 36s 41.39 MB
Download Speed : 214.3 kB/s
Upload Speed : 0 B/s
Peers : 0 connected (5 in swarm)
Average Completion: 45.9%

Share Ratio :
Hash Fails : 0 (0 B)
Completed : 15
Distributed Copies: 0.459

Info

Name : Battle of the Game Consoles
Save In : C:\Documents and Settings\Paul\My Documents\Azureus Downloads\MW_
# of Pieces : 563/1226
Tracker URL : http://tracker02.azureusplatform.com:10000/announce
Created On : 14-Dec-2007 21:53:56

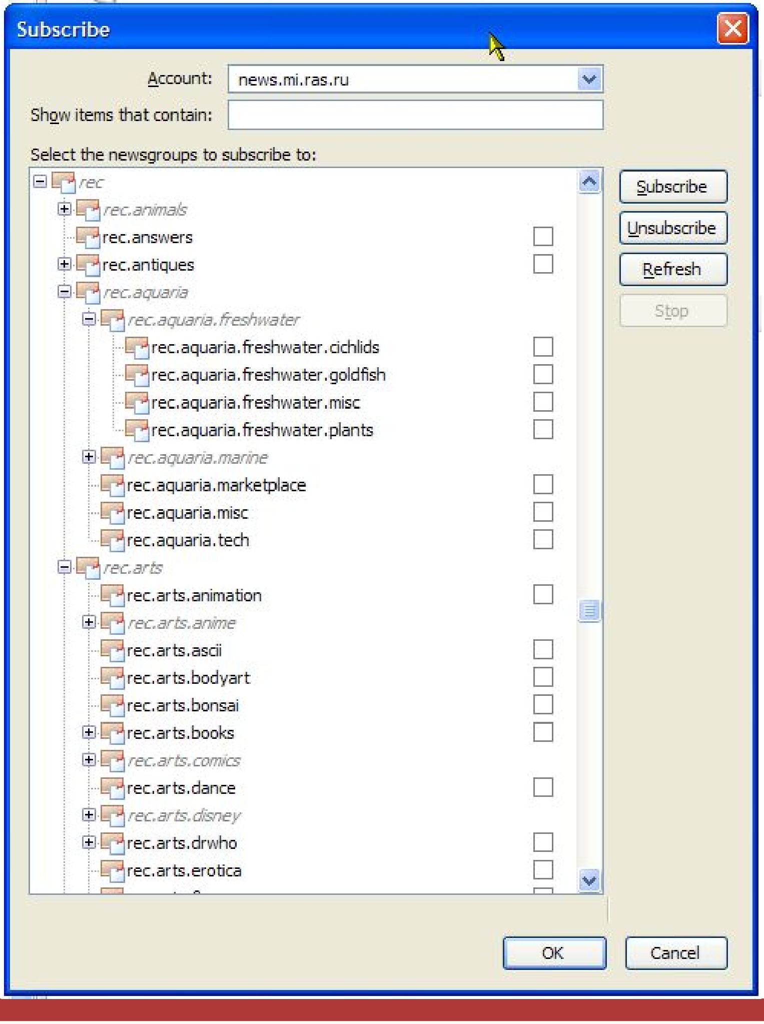
Status : Downloading
Total Size : 76.58 MB
Hash : 35F5E61D 94FD4FA5 3217A130 EF72F284
Size : 64.0 kB
Private Torrent: No

Tracker Status : OK
Update in : 00:13:54

User Comment :
Torrent Comment : Provided by Vuze.com

USENET News

- Message boards, Predates WWW
- >50,000 newsgroups active



Subscribe

Account: news.mi.ras.ru

Show items that contain:

Select the newsgroups to subscribe to:

- rec
  - rec.animals
  - rec.answers ☐
  - rec.antiques ☐
  - rec.aquaria
    - rec.aquaria.freshwater
      - rec.aquaria.freshwater.cichlids ☐
      - rec.aquaria.freshwater.goldfish ☐
      - rec.aquaria.freshwater.misc ☐
      - rec.aquaria.freshwater.plants ☐
    - rec.aquaria.marine
    - rec.aquaria.marketplace ☐
    - rec.aquaria.misc ☐
    - rec.aquaria.tech ☐
  - rec.arts
    - rec.arts.animation ☐
    - rec.arts.anime
    - rec.arts.ascii ☐
    - rec.arts.bodyart ☐
    - rec.arts.bonsai ☐
    - rec.arts.books ☐
    - rec.arts.comics
    - rec.arts.dance ☐
    - rec.arts.disney
    - rec.arts.drwho ☐
    - rec.arts.erotica ☐

USENET news

- Messages are like Email – text encoded attachments
- Images split into dozens or hundreds of messages
- Messages can be distributed across multiple newsgroups
- Until all parts of a binary document (image, program, zip-file, movie) are received, the binary document cannot be reconstructed and analysed
- Even if it is inappropriate content, no way to block it until it has already been distributed

What is the Problem to Solve?

Concerns

- Trial does not have enough large networks to test scalability - may give ‘false positive’
- Retaining ‘state’ in the middle of the network reduces reliability, scalability, and trust – fundamentally breaks Internet architecture
- The problem to be solved hasn’t yet been articulated clearly
  - Are we blocking illegal content, blocking ‘undesirable’ content, and who does the classification, how can incorrect classification be reviewed?
- Technology is no substitute for POS

Parent Over the Shoulder

Thank you

Paul Brooks

Director, ISOC-AU

pbrooks@layer10.com.au

www.isoc-au.org.au