

Internet Study 2008/2009

Hendrik Schulze, Klaus Mochalski

For the third year in a row, after 2006 and 2007, ipoque has conducted a comprehensive study measuring and analyzing Internet traffic in eight regions of the world. The study includes statistical data about popularity and user behavior for all common network protocols. This covers most applications used in today's Internet such as Web browsing, media streaming, P2P file sharing, one-click file hosting, instant messaging, Internet telephony and online games. BitTorrent and eDonkey downloads have been analyzed to classify the transferred files according to their content type. Some of the key findings are: P2P still produces most Internet traffic worldwide although its proportion has declined across all monitored regions – losing users to file hosting and media streaming; regional variations in application usage are very prominent; and Web traffic has made its comeback due to the popularity of file hosting, social networking sites and the growing media richness of Web pages.

Introduction

This study uses the same methodology as the 2007 Internet Study¹ to classify network traffic according to protocol and protocol class. Several of ipoque's ISP and university customers agreed to provide anonymized traffic statistics collected by PRX Traffic Managers installed in their networks. Protocols and applications are detected with a combination of layer-7 deep packet inspection (DPI) and behavioral traffic analysis. All classified network flows are then accounted per subscriber providing the raw data for this study. All comparisons and trend analyses, if not noted otherwise, refer to the 2007 study. It is important to note that the results are not statistically representative. The monitoring sites were picked based on accessibility and may not be typical for their region. For instance, we could have picked the only cable ISP in a country, where all other ISPs only offer DSL access.

Protocol class definition as used in this study

Protocol Class	Definition	Examples
P2P	P2P file sharing	Ares, BitTorrent, eDonkey, Gnutella
Web	Web pages incl. file hosting, excl. streaming	all Web sites, RapidShare, Megaupload
Streaming	Audio and video streaming	Flash, QuickTime, Real Media, RTSP
VoIP	Voice over IP - Internet telephony	IAX, H.323, MGCP, SIP, Skinny, Skype
IM	Instant messaging	IRC, Gadu-Gadu, XMPP (Jabber, Google Talk), MSN, Oscar, Paltalk, Yahoo
Tunnel	Encrypted and unencrypted tunneling protocols	GRE, IPsec, OpenVPN, SSL, Tor
Standard	Legacy Internet protocols (without Web)	BGP, DNS, FTP, ICMP, IGMP, IMAP, NTP, POP, SMTP, Telnet, Usenet
Gaming	Multiplayer and network games	Battlefield, Half-Life 2, Lively, Quake, Second Life, Steam, World of Warcraft, Xbox
Unknown	Non-classified traffic	

What Is New? What Is Different?

Key facts

- 8 regions: Northern Africa, Southern Africa, South America, Middle East, Eastern Europe, Southern Europe, Southwestern Europe, Germany
- 1.3 petabytes of user traffic monitored
- 1.1 million users represented
- At 8 ISPs and 3 universities
- DPI and behavioral analysis of about 100 protocols
- Packet size analysis for all protocols

The regional coverage has been extended to include eight regions of the world. Last year's study covered the five regions Australia, Eastern Europe, Germany, the Middle East and Southern Europe. The additional regions are Northern and Southern Africa, South America and Southern Europe. The countries covered under the Southern Europe region last year are now Southwestern Europe. We decided to distinguish between these two regions because the new countries added in this study showed different results that would have vanished in an aggregation. Just as in previous years, regional differences are significant.

Measurement sites

Region	Monitored Users	Evaluated Traffic (TB)	Network Type
Northern Africa	250.000	70	ISP
Southern Africa	50.000	83	ISP, satellite uplink
South America	50.000	38	ISP
Middle East	200.000	60	ISP
Eastern Europe	100.000	200	ISP
Southern Europe	20.000	170	ISP
Southwestern Europe	80.000	100	ISP
Germany	100.000	560	1 ISP, 3 universities

¹ ipoque's P2P Survey 2006 and Internet Study 2007 are available at <http://www.ipoque.com/resources/internet-studies>.

The time period of the measurements was shortened from an average of four weeks to two weeks. While two weeks are long enough to get a snapshot of the user behavior, even four weeks would have been too short to discern broader trends. So this has no impact on the validity of the results, but reduces the data collection and analysis effort.

Due to the ongoing improvement of ipoque's protocol support, more different protocols and applications are covered in this study, including encrypted and unencrypted tunneling, streaming and Internet TV.¹

For the first time, we have recorded the average packet size for all protocols – separately for downstream and upstream traffic. This analysis is only available for South America.

Content Type Classification

The content type analysis, that looks at what kind of files are shared in P2P networks, is limited to only a few of the measurement sites – mostly due to privacy concerns – and covers two regions: Germany and Southern Europe. The same methodology as in the previous studies is used.

In our 2006 P2P study¹, we introduced this semi-automatic content type classification procedure for files shared through the BitTorrent and eDonkey networks. For eDonkey, the classification is mostly based on file name keywords combined with a manual classification of the most popular titles. The BitTorrent classification mostly relied on meta data available on various large BitTorrent tracker or repository sites such as The Pirate Bay² and Mininova³. Over the last two years, the importance has gradually shifted towards BitTorrent. Unfortunately, the automatic classification has become more complicated. While the same approach still generally works, it requires a far greater effort due to an exploding number of tracker sites – many of them offering local-language content making the classification more difficult. With the tracker sites included for our 2007 study we were able to classify 55 percent of all transferred files representing 70 percent of all BitTorrent traffic. This time we were only able to classify about 30 percent of all files and close to 50 percent of the traffic volume using the same tracker sites.

Protocol and Application Class Distribution

Key Facts

- P2P generates most traffic in all regions
- Lower percentage of P2P than in 2007
- Higher percentage of Web traffic mainly due to file hosting services

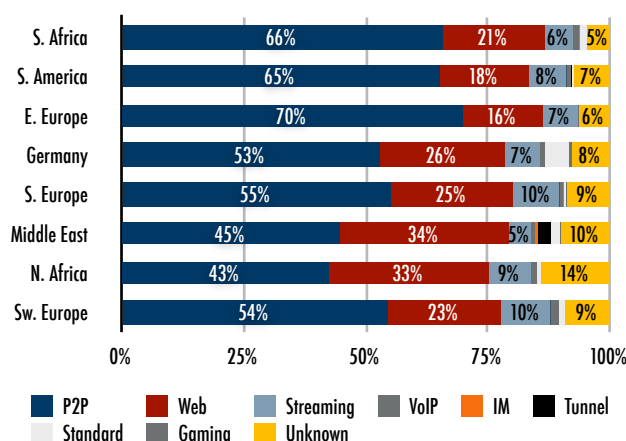
The following table list the proportion of the overall network traffic generated by the respective protocol class. A good quality indicator for these results is the amount of unclassified traffic, which is below 10 percent for all but one region. In Northern Africa, tunnel protocol support was not licensed by the customer and thus disabled in the PRX Traffic Manager – thus the high value of 14 percent.

Peer-to-peer file sharing (P2P) still generates by far the most traffic in all monitored regions – ranging from 43 percent in Northern Africa to 70 percent in Eastern Europe. The regional differences can probably be attributed to varying subscriber access bandwidth, availability of localized content, and cultural habits. Interestingly, the two Arab regions, Northern Africa and the Middle East, share a consistently lower P2P ratio. This is compensated by a comparably higher proportion of Web traffic.

Protocol class proportions 2008/2009

Protocol Class	Southern Africa	South America	Eastern Europe	Northern Africa	Germany	Southern Europe	Middle East	South-western Europe
P2P	65,77%	65,21%	69,95%	42,51%	52,79%	55,12%	44,77%	54,46%
Web	20,93%	18,17%	16,23%	32,65%	25,78%	25,11%	34,49%	23,29%
Streaming	5,83%	7,81%	7,34%	8,72%	7,17%	9,55%	4,64%	10,14%
VoIP	1,21%	0,84%	0,03%	1,12%	0,86%	0,67%	0,79%	1,67%
IM	0,04%	0,06%	0,00%	0,02%	0,16%	0,03%	0,50%	0,08%
Tunnel	0,16%	0,10%	-	-	-	0,09%	2,74%	-
Standard	1,31%	0,49%	-	0,89%	4,89%	0,52%	1,83%	1,23%
Gaming	-	0,04%	-	-	0,52%	0,05%	0,15%	-
Unknown	4,76%	7,29%	6,45%	14,09%	7,84%	8,86%	10,09%	9,13%

Distribution of protocol classes 2008/2009



¹ A complete list of the supported protocols is available on our Web site at <http://www.ipoque.com/products/protocol-support>.

² The Pirate Bay: <http://thepiratebay.org>

³ Mininova: <http://www.mininova.org>

Trends and Regional Differences

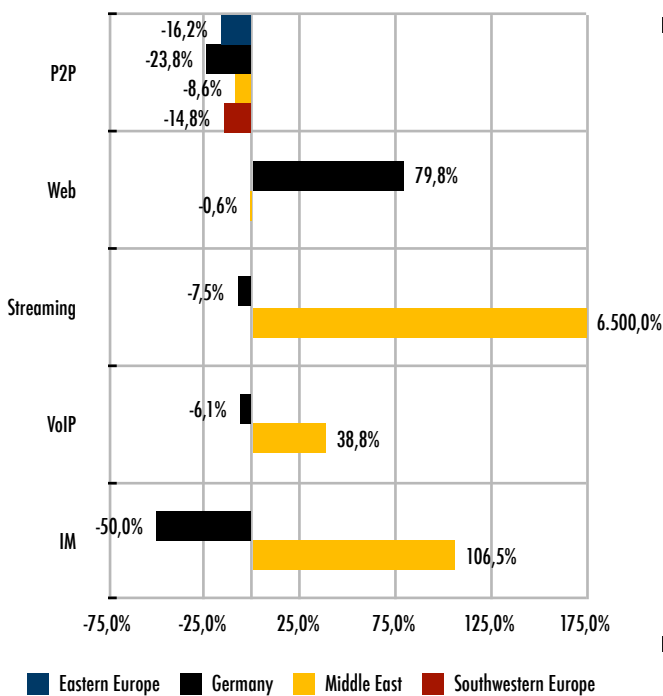
The proportion of P2P has decreased in all regions. It is important to note, though, that this does not mean there is less P2P traffic than one year ago, but only that P2P has grown slower than other traffic. Unfortunately, most of the measurement points participating in this study changed compared to 2006 and 2007 making a comparison with previous results only possible for Germany and the Middle East.

Protocol class proportions 2008/2009 versus 2007

Protocol Class	Eastern Europe		Germany		Middle East		Southwestern Europe	
	2008/2009	2007	2008/2009	2007	2008/2009	2007	2008/2009	2007
P2P	69,95%	83,46%	52,79%	69,25%	44,77%	48,97%	54,46%	63,94%
Web	16,23%	-	25,78%	14,34%	34,49%	34,71%	23,29%	-
Streaming	7,34%	-	7,17%	7,75%	4,64%	0,07%	10,14%	-
VoIP	0,029%	-	0,86%	0,92%	0,79%	0,57%	1,67%	-
IM	0,002%	-	0,16%	0,32%	0,50%	0,24%	0,08%	-

In Germany, Web traffic experienced the fastest and most significant growth in absolute terms – from 14 to 26 percent – mainly due to file hosting sites – also known as direct download links (DDL) such as RapidShare and Megaupload – but also due to the increasing media-richness of Web pages. This, however, does not include streaming video, which has slightly decreased in proportion.

Protocol type proportion changes relative to 2007



The Middle East maintained its comparably high level of Web traffic of 35 percent. Media streaming has jumped from virtually non-existent to nearly 5 percent. Most likely this can be attributed to higher access speeds which have made this application possible only during the last year. The same is true, to a lesser degree, about VoIP (plus 40 percent) and IM, which doubled. In Germany, the proportion of these two applications has decreased, which could be caused by stagnating user numbers keeping traffic levels constant.

Most Popular Protocols

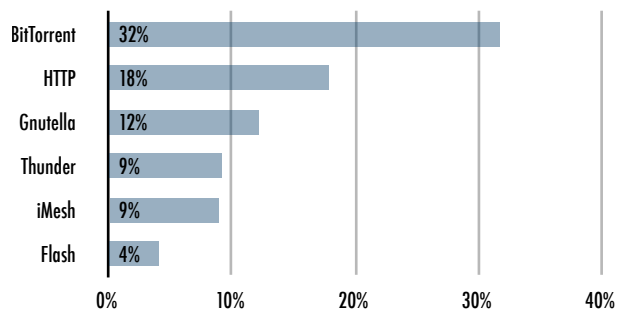
Key Facts

- BitTorrent is number one
- HTTP second in most regions
- At least two P2P protocols among top-5 in each region

In all regions apart from South America, BitTorrent is the dominating protocol followed by HTTP. In South America, Ares, another P2P protocol, leads the pack before BitTorrent and eDonkey. Between two and four P2P protocols make it into the top-five list. The varying popularity of the different P2P protocols can often be explained by local content. See the section on P2P for more details.

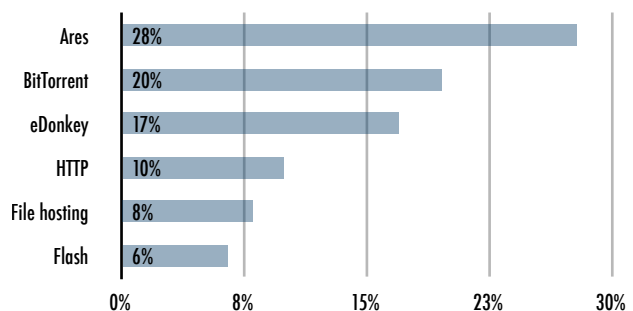
Southern Africa

has an unusually high amount of the iMesh and Thunder P2P protocols

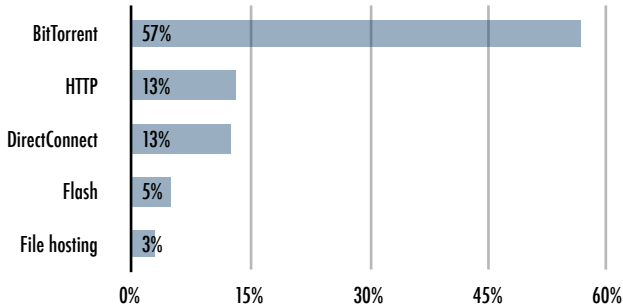


South America

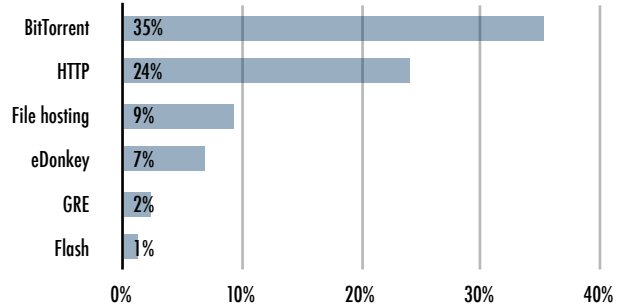
is the only region where the Ares P2P protocol produces most traffic



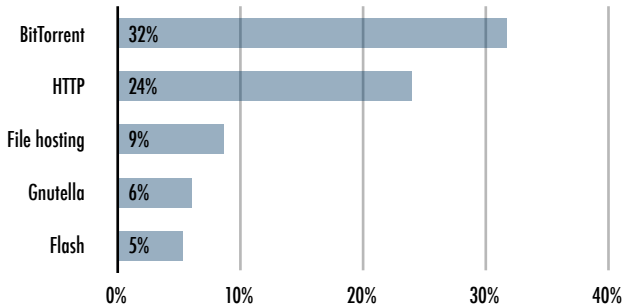
Eastern Europe
has the most BitTorrent traffic
of all regions



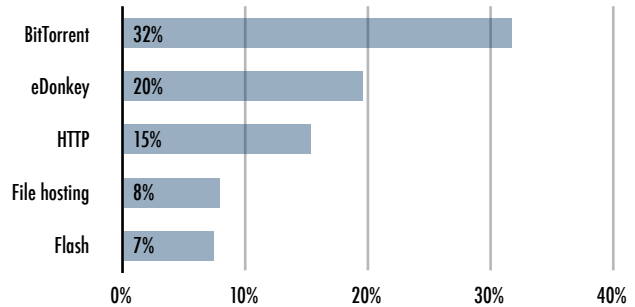
The Middle East
shows an unusual high GRE usage
because OpenVPN is blocked



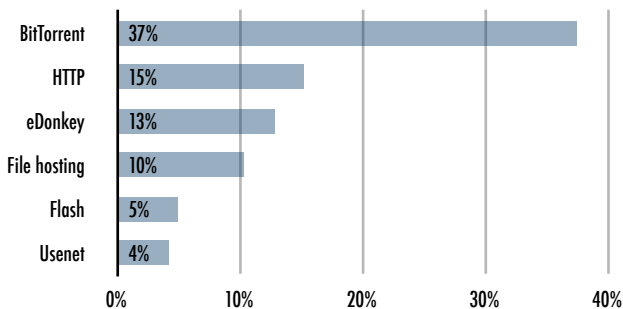
Northern Africa



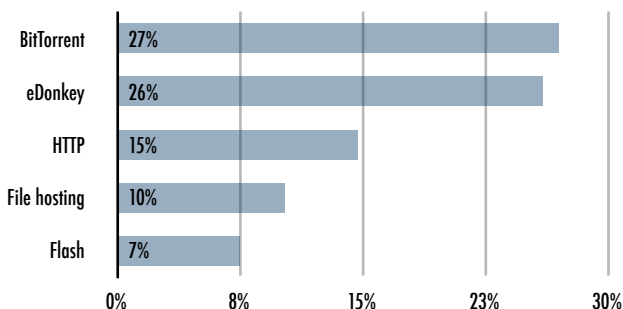
Southwestern Europe



Germany
is the only region with significant
Usenet traffic



Southern Europe
has the highest proportion of
eDonkey traffic



User Number Statistics

Key Facts

- HTTP used by over 90% of all network nodes
- Very high number of P2P users in South America
- Strong concentration on a few applications among university users

The collection of user number statistics needs to be explicitly turned on in the PRX Traffic Manager. Many network operators do not enable this feature for privacy reasons, or they do not make these data available to anyone outside their organization. For this reasons, this study only includes user number statistics for South America and a German university. Unsurprisingly, the results show a tremendous difference in user behavior between the university users and the South American ISP subscribers.

The following table lists the proportion of unique network users per protocol. Users are identified by their IP address and thus not only include client programs operated by a human being, but also application servers. Dynamic IP addresses are automatically mapped to a fixed user identifier so that a network node is only counted once even it uses dynamic IP addresses.

In both regions, HTTP is the most-used protocol. The higher number in South America can be explained by a lower relative number of server machines run by the mostly resi-

dential ISP subscribers compared to the university. At the university, only three protocols are used by more than 10 percent of the network nodes – HTTP, Skype and BitTorrent – while at the commercial ISP, 27 protocols get beyond this threshold. Overall, the variety of applications is considerably higher for the IPS users where we have detected 78 different protocols versus 50 at the university.

It is striking how popular P2P file sharing is in South America. While we recorded only about 20 percent P2P users in our previous study, a number supported by other studies as well, it is much higher at the South American ISP. There are four P2P protocols used by more than 10 percent of the subscribers – Ares, BitTorrent, eDonkey and Gnutella. The worldwide most popular protocols BitTorrent and eDonkey are both used by nearly 40 percent. Ares has an unusually high popularity in this region with a 64 percent user base. One possible reason could be the availability of Spanish and Portuguese language content, but this is no sufficient explanation because there is no similar Ares user base in Southwestern Europe.

*Relative user numbers per protocol
for South America and one German university*

#	South America		Germany, University		#	South America		Germany, University	
	Protocol	Users	Protocol	Users		Protocol	Users	Protocol	Users
1	HTTP	96,66%	HTTP	91,08%	40	Joost	0,92%	Soulseek	0,04%
2	ICMP	91,01%	Skype	26,77%	41	Pando	0,87%	TVAnts	0,04%
3	Flash	88,47%	BitTorrent	14,13%	42	iMesh	0,75%	H323	0,02%
4	MPEG	80,45%	MSN	9,71%	43	DHCP	0,64%	SMTP	0,02%
5	Wndws Media	79,00%	Thunder	7,59%	44	QQLive	0,63%	HamachiVPN	0,01%
6	NTP	70,74%	DNS	7,55%	45	SopCast	0,50%	IMAP	0,01%
7	File hosting	65,62%	Flash	5,14%	46	TVAnts	0,47%	POP	0,01%
8	Ares	63,61%	eDonkey	4,92%	47	Steam	0,42%	SopCast	0,01%
9	DNS	47,51%	ICMP	4,09%	48	Usenet	0,42%	TVUPlayer	0,01%
10	RTSP	42,31%	Oscar	3,83%	49	RDP	0,40%	Usenet	0,01%
11	FTP	41,90%	SSL	2,80%	50	TVUPlayer	0,37%	iMesh	0,01%
12	BitTorrent	38,76%	File hosting	1,76%	51	DirectConnect	0,27%		
13	eDonkey	36,93%	QQ	1,63%	52	PPStream	0,15%		
14	QuickTime	32,20%	Yahoo	1,47%	53	QQ	0,14%		
15	MMS	27,58%	Gnutella	0,93%	54	Freenet	0,13%		
16	Skype	22,92%	IRC	0,63%	55	KaZaa	0,11%		
17	MSN	21,34%	PPStream	0,61%	56	Telnet	0,07%		
18	SIP	20,87%	MPEG	0,52%	57	UUSee	0,07%		
19	SMTP	20,40%	RTSP	0,48%	58	H323	0,07%		
20	Yahoo	18,64%	SIP	0,44%	59	NFS	0,05%		
21	POP	16,21%	XMPP	0,37%	60	Wrl d. Warc.	0,05%		
22	SSL	12,05%	Zattoo	0,37%	61	SSDP	0,04%		
23	RTP	11,83%	Wndws Media	0,34%	62	IAX	0,03%		
24	IRC	11,74%	GRE	0,33%	63	GRE	0,02%		
25	AVI	11,50%	Ares	0,32%	64	SSH	0,02%		
26	Gnutella	11,46%	DirectConnect	0,31%	65	Winny	0,02%		
27	Realmedia	11,35%	MMS	0,31%	66	Feidian	0,02%		
28	XMPP	9,52%	FTP	0,29%	67	HalLife2	0,01%		
29	PPLive	6,83%	OpenVPN	0,28%	68	Zattoo	0,01%		
30	Manolito	6,21%	QQLive	0,22%	69	Slingbox	0,01%		
31	OpenVPN	5,04%	AVI	0,18%	70	Tor	0,01%		
32	ORB	4,09%	RTP	0,18%	71	IGMP	0,01%		
33	Thunder	2,38%	NTP	0,16%	72	Citrix	0,01%		
34	Oscar	1,76%	QuickTime	0,12%	73	Filetopia	0,01%		
35	IPsec	1,65%	SSH	0,09%	74	EGP	0,01%		
36	VNC	1,25%	IPsec	0,05%	75	WinMX	0,01%		
37	OGG	1,19%	Realmedia	0,05%	76	Gadu-Gadu	0,00%		
38	Soulseek	1,03%	Gadu-Gadu	0,04%	77	OSPF	0,00%		
39	IMAP	1,01%	PPLive	0,04%	78	Xbox	0,00%		

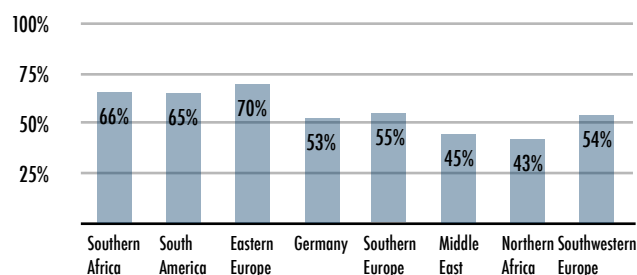
Another noteworthy observation for South America is the high number of hosts that send DNS messages. Depending on where PRX Traffic Manager is installed in the network, it should see either close to 100 percent of hosts using DNS (if the DNS server is on the external side of the system) or close to zero percent (the more common case where hosts and DNS server reside both on the internal side). This is what we observe at the university network, where PRX has seen DNS messages for 7.55 percent of all hosts – which apparently use an external DNS server. At the ISP, PRX sees DNS traffic for nearly half of all hosts. Most likely, the ISP operates several load-balancing DNS servers on both sides of PRX Traffic Manager.

P2P Protocols

Key Facts

- Lower proportion of P2P across all regions
- BitTorrent by far the most popular P2P protocol
- eDonkey still second, but much lower than last year
- Significant amount of Gnutella, iMesh, Thunder and DirectConnect traffic in some regions

Relative P2P traffic volume



The overall proportion of P2P traffic has declined throughout all monitored regions. As noted before, this does not constitute a decline in the absolute amount of P2P traffic – other protocols have simply experienced a greater increase.

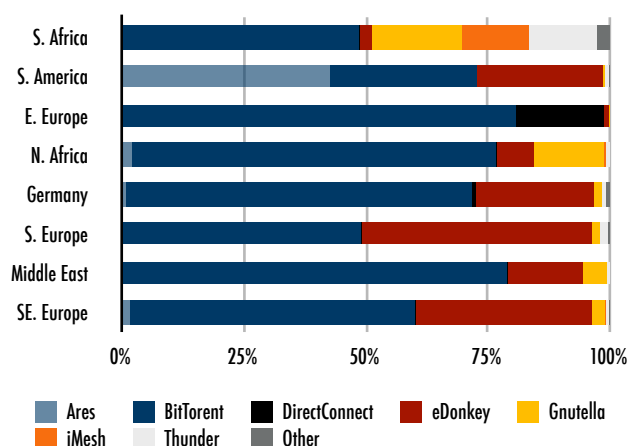
The most significant decrease was observed in Germany where the P2P share has dropped by 24 percent. This can have several reasons. Some ISPs have reportedly started to throttle or de-prioritize P2P traffic. This does not only affect the subscribers of the bandwidth-managing ISP, but all other peers trying to download from those throttled peers, thus reducing the overall download capacity in the P2P network.

In addition, German file sharers have a significant risk of being prosecuted if they share copyright-protected material. Current copyright infringement investigations mainly target the BitTorrent and eDonkey networks due to their overwhelming popularity in Germany. At the same time, commercial download services such as iTunes offer a legal alternative with growing repositories of high-quality content. Some file sharers have also turned to other, free and unregulated, alternatives, mainly file hosting services (e.g.

RapidShare, Megaupload) and the Usenet (often using commercial services such as UseNeXT). A reliable substitution analysis (i.e. who is moving away from a service to one or several new ones, and for what reasons) would only be possible with a representative poll and is beyond the scope of this study.

Most popular P2P protocols

Protocol	Southern Africa	South America	Eastern Europe	Northern Africa	Germany	Southern Europe	Middle East	South-western Europe
All P2P	65,77%	65,21%	69,95%	42,51%	52,79%	55,12%	44,77%	54,46%
Ares	0,29%	42,63%	0,00%	2,24%	0,84%	0,16%	0,11%	1,80%
BitTorrent	48,34%	30,02%	80,83%	74,51%	70,77%	48,94%	78,85%	58,20%
DirectConnect	0,01%	0,00%	17,87%	0,08%	0,85%	0,00%	0,12%	0,30%
eDonkey	2,48%	25,99%	1,16%	7,70%	24,22%	47,17%	15,37%	35,99%
Gnutella	18,60%	0,36%	0,14%	14,21%	1,75%	1,66%	5,00%	2,75%
iMesh	13,60%	0,02%	0,00%	0,47%	0,00%	0,03%	0,00%	0,14%
Thunder	14,04%	0,80%	0,00%	0,69%	0,77%	1,64%	0,52%	0,62%
Other	2,64%	0,19%	0,00%	0,10%	0,68%	0,41%	0,03%	0,21%



Just as in previous years, BitTorrent is the dominating P2P protocol in all regions but South America. This is most likely a result of the higher performance – important for large file exchanges such as DVD images – and better content quality of BitTorrent compared to other P2P networks. Fakes are more often found in eDonkey, where they were used by the content industry as a countermeasure against illegal downloads of copyrighted material¹, than in BitTorrent. Quite frequently, fakes are also distributed by P2P users themselves – for no obvious reasons. eDonkey has maintained its second place in five out of the eight regions, but with a significantly decreased proportion of all P2P.

Another possible reason for the increase of BitTorrent and the decrease of eDonkey could be that copyright investigators are more active in the eDonkey network because investigations are technically slightly easier than in BitTorrent. This may not only drive eDonkey users to BitTorrent, but

also to alternative file sharing systems, mainly Web-based file hosting sites and the Usenet.

Trends and Regional Differences

BitTorrent has the biggest share in Eastern Europe and the Middle East with about 80 percent. It has its lowest popularity in South America with only 30 percent. In all other regions, at least half of all P2P traffic is BitTorrent. Four regions significantly deviate from the usual BitTorrent/eDonkey dominance. Northern Africa has only 8 percent eDonkey, but over 14 percent Gnutella traffic. In Eastern Europe, it is DirectConnect with 18 percent, that replaces eDonkey. South America has an unusually high share of Ares traffic of 43 percent. In all other regions, Ares is less than 2 percent of all P2P. In Southern Africa, we have observed a nearly equal share of around 15 percent for Gnutella, iMesh and Thunder. Thunder is a Chinese P2P network that is apparently used by Chinese immigrants living in South America.

Changes in P2P proportion 2007 to 2008 Germany

Protocol	2008	2007	Change
Other	3,77%	0,99%	280,54%
BitTorrent	70,77%	66,70%	6,11%
eDonkey	24,22%	28,59%	-15,27%
Gnutella	1,75%	3,72%	-53,06%

Changes in P2P proportion 2007 to 2008 Southwestern Europe

Protocol	2008	2007	Change
Other	3,07%	0,63%	386,85%
BitTorrent	58,20%	40,09%	45,17%
Gnutella	2,75%	2,23%	23,21%
eDonkey	35,99%	57,05%	-36,92%

Changes in P2P proportion 2007 to 2008 Eastern Europe

Protocol	2008	2007	Change
Other	1,30%	5,57%	-76,62%
BitTorrent	80,83%	65,71%	23,01%
Direct Connect	17,87%	28,72%	-37,78%

Changes in P2P proportion 2007 to 2008 the Middle East

Protocol	2008	2007	Change
Other	0,78%	2,36%	-67,10%
BitTorrent	78,85%	56,21%	40,28%
eDonkey	15,37%	38,51%	-60,09%
Gnutella	5,00%	3,10%	61,41%

¹ See ipoque's white paper "Copyright Protection in the Internet" for a detailed comparison of countermeasures against the sharing of copyright-protected material in P2P networks available at <http://www.ipoque.com/resources/white-papers>.

Encryption and Obfuscation in P2P Networks

Many modern P2P clients offer various forms of obfuscation and encryption. The common claim that this is meant to improve the privacy of P2P users is plain dishonest. Even if encryption is enabled, files are still shared with the general public, so for everybody to download and store unencrypted. This is also why encryption does not provide any protection against copyright investigations in P2P networks, where investigators use normal P2P clients to participate in the network and download files from potential infringers. The only sensible reason for encryption is the attempt to circumvent bandwidth limitations imposed for P2P transfers by the ISP. However, with modern traffic management systems, which are able to reliably detect obfuscated and encrypted P2P traffic, this measure is totally ineffective.

Proportion of encrypted and unencrypted BitTorrent and eDonkey traffic in Germany and Southern Europe

Region	BitTorrent		eDonkey	
	encrypted	unencrypted	encrypted	unencrypted
Germany	22,81%	77,19%	16,08%	83,92%
Southern Europe	26,21%	73,79%	7,03%	92,97%

Comparative numbers are only available for Germany, where the relative amount of encrypted P2P traffic remains nearly constant for eDonkey (15 percent in 2007, 16 percent in 2008) and has slightly increased for BitTorrent (18 percent in 2007, 23 percent in 2008). In Southern Europe, the disparity in encryption usage between these two most popular networks is even greater.

To understand the higher encryption prevalence for BitTorrent, it is worth looking at the client development over the last years. eMule, the most popular eDonkey client, has only seen one release and one update since May 2007. Development is driven by volunteers, and eMule has only two developers, aMule ten. For Vuze, formerly known as Azureus, the most popular BitTorrent client, there were one major release, one minor release and twelve updates during the same period. Vuze, like some other BitTorrent clients, is a commercial venture offering users the possibility to publish and monetize their original content.

P2P Content Type Analysis

Key Facts

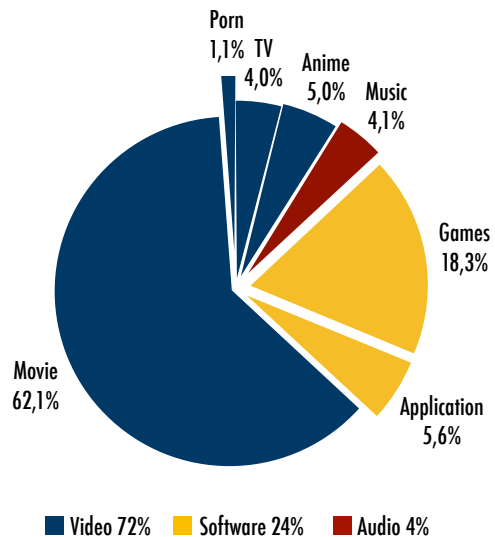
- Video is most popular content by volume and by number of files
- Software is second in BitTorrent with about one third of the volume
- Audio files are second in eDonkey
- Less video, particularly porn, and more software downloads with BitTorrent compared to 2007

As in 2006 and 2007, we have gathered lists of all files downloaded with eDonkey and BitTorrent in Germany and

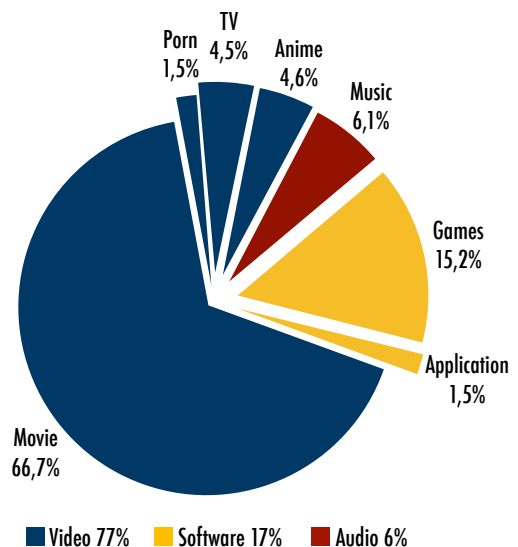
Southern Europe. The German data set was gathered at a large university, the other at an ISPs.

The popularity of video downloads is consistent with last year's results. Between 58 percent (BitTorrent, Germany) and 86 percent (eDonkey, Southern Europe) of all P2P traffic, and between 44 percent (eDonkey, Germany) and 77 percent (BitTorrent, Southern Europe) of all transferred files are videos. Drilling down to subcategories, movies are most popular ranging from 16 (BitTorrent, Germany) to 62 (BitTorrent, Southern Europe) percent by volume and 6 (eDonkey, Germany) and 67 (BitTorrent, Southern Europe) percent by number of files.

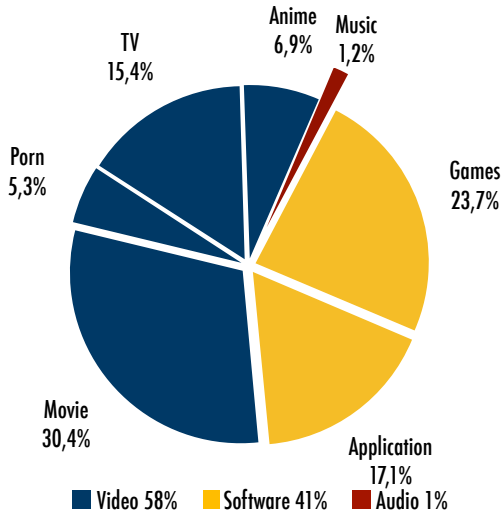
*BitTorrent content type distribution
Southern Europe
Traffic volume*



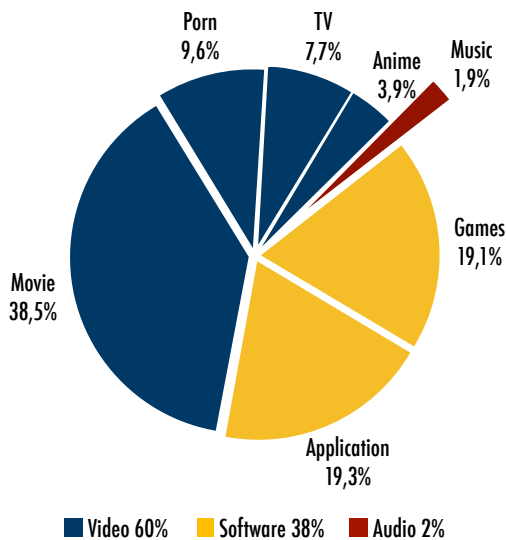
*BitTorrent content type distribution
Southern Europe
Number of unique files*



BitTorrent content type distribution
Germany (university)
Traffic volume



BitTorrent content type distribution
Germany (university)
Number of unique files

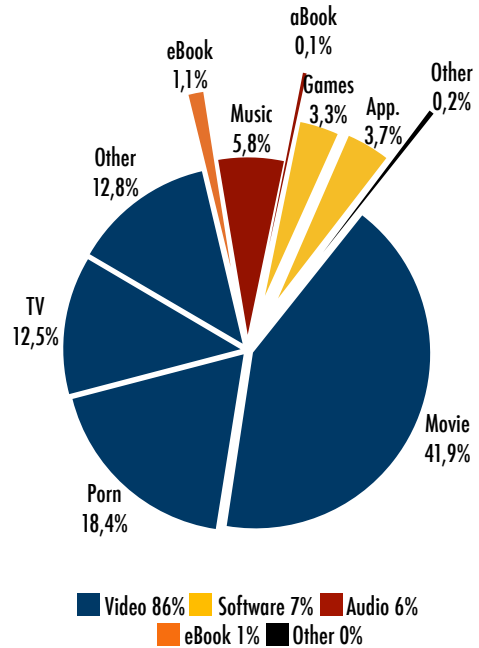


The ratio between volume and number of files for video content is close to one for BitTorrent and about 1.5 for eDonkey. This does not mean that eDonkey video files are on average larger than BitTorrent ones (in fact, BitTorrent files tend to be larger), but that fewer unique files are downloaded more often via BitTorrent. This is due to the higher administrative overhead involved in sharing a file through BitTorrent.

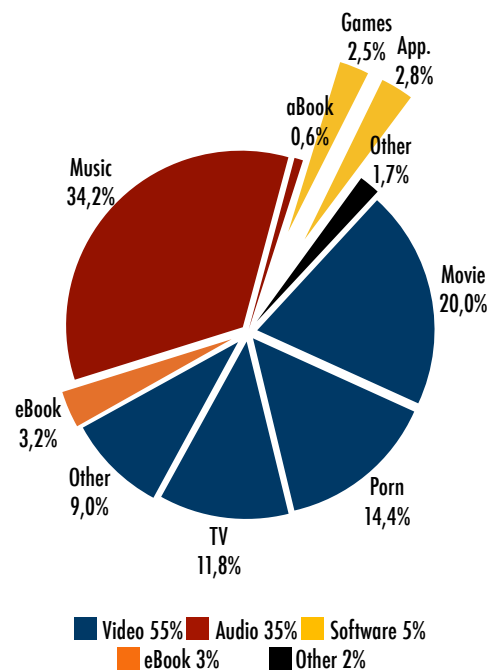
For the 'movies' subcategory, the BitTorrent ratio is again close to one, while for eDonkey it is even greater, with 2 in Southern Europe and 4.5 at the German university, again hinting at larger or higher-quality files.

The second-most popular category for BitTorrent is software ranging from 24 (Southern Europe) to 41 (Germany) percent traffic volume and 17 (Southern Africa) to 38 (Germany) percent of all files. The by far most popular subcategory is games. For eDonkey, software is below about 10 percent of the traffic volume.

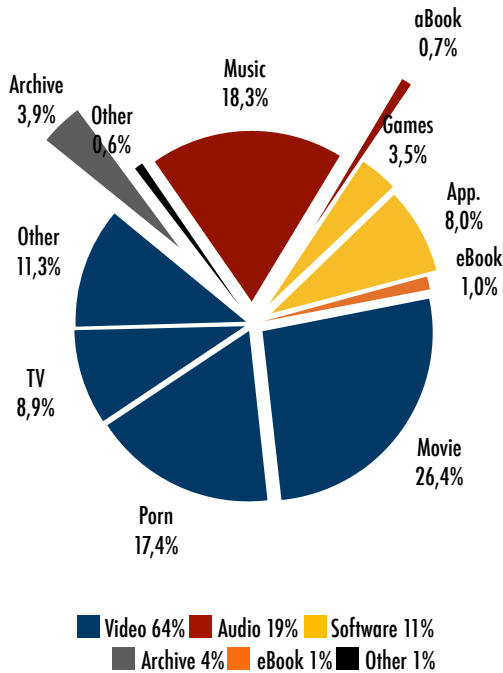
eDonkey content type distribution
Southern Europe
Traffic volume



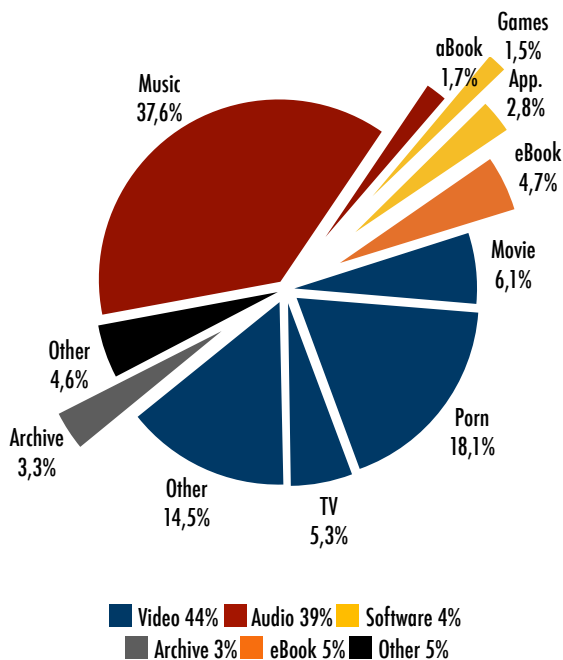
eDonkey content type distribution
Southern Europe
Number of unique files



eDonkey content type distribution
Germany (university)
Traffic volume



eDonkey content type distribution
Germany (university)
Number of unique files



The second most popular category for eDonkey is audio. It is 6 percent in Southern Europe and 19 percent in Germany by volume, and 35 and 39 percent, respectively, by number of files. This ratio is caused by the small file size for audio content.

Trends

A comparison to 2007 is only possible for Germany. For BitTorrent, the amount of video content has declined from 79 to 58 percent, while software downloads have increased from 11 to 41 percent. Of all video downloads, the subcategory 'porn' has disproportionately decreased both in volume and in the number of unique files. Most likely this trend will continue with the growing popularity of Web streaming portals offering similar content without the need to download large files first before watching them. Interestingly, audio downloads dropped from 9 percent volume and 22 percent of all files to a mere 1 respectively 2 percent. eDonkey has seen now similarly pronounced changes.

Voice over IP and Skype

Key Facts

- SIP generates over 50% of all VoIP traffic in most regions
- Skype is number one in Eastern Europe and the Middle East
- SIP/RTP initiated by IM applications very popular with about 20-30% of VoIP everywhere but in the Middle East
- Virtually no IAX/IAX2 traffic anywhere

Internet telephony – or Voice over IP (VoIP) – has seen considerable movement in the past year. In our last study, Skype was by far the most popular application due to its unique combination of functionality and ease of use. It offers voice and video calls, instant messaging and file transfer, and is notorious for its ability to pass through corporate firewalls, network address translation (NAT) gateways and Web proxies. Particularly these capabilities have caused some headaches for IT management staff responsible for network security.

Now, other applications, such as the Microsoft and Yahoo Messengers, that have previously been used only for instant messaging (IM), come with VoIP functionality as well. Different from Skype, these applications use the standards-based SIP protocol or slightly modified versions thereof for establishing RTP voice calls.

While there had been many SIP/RTP VoIP clients and operators offering computer-to-computer and computer-to-phone calls (e.g. Vonage, Sipgate) before, they all lacked Skype's zero configuration properties. This has changed with these new, VoIP-enhanced IM applications that allow to place a phone call at least to all IM contacts, but quite often – through cooperations with phone companies – also to ordinary phones without any additional configuration. This clearly shows in this year's VoIP statistics.

We have added a new category 'IM-VoIP' covering the RTP voice and SIP call set-up traffic triggered by these IM appli-

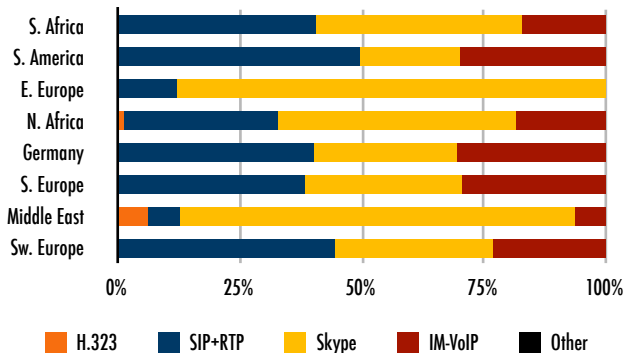
cations. This category was not supported in last year's study, so we need to compare the sum of SIP and IM-VoIP against last year's SIP numbers. Comparative values are available for Germany and the Middle East. In Germany, there was less than 2 percent SIP traffic in 2007, compared to over 70 percent in 2008. VoIP numbers in the Middle East have remained fairly constant with a negligible decline from 14 to 13 percent.

Skype is the most popular VoIP protocol in Eastern Europe and the Middle East with bandwidth shares larger than 80 percent. Skype appears to be particularly popular in places with lower Internet access speed due to its dynamic audio codec that adapts to varying bandwidth availability.

Unchanged since 2007, the IP PBX interchange protocols IAX and IAX2, included in 'Other' in the table below, have no practical significance. H.323 only shows up in Northern Africa with 1 percent and in the Middle East with 6 percent.

Traffic distribution of Internet telephony protocols¹

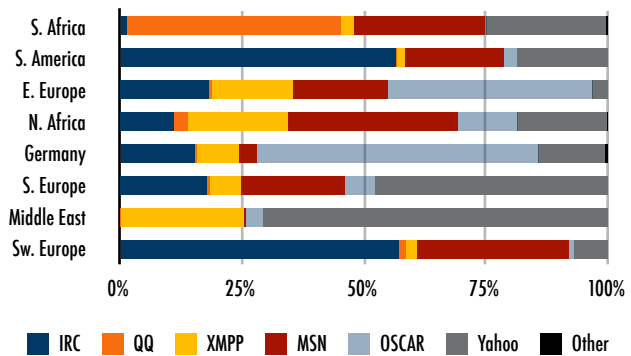
Protocol	Southern Africa	South America	Eastern Europe	Northern Africa	Germany	Southern Europe	Middle East	South-western Europe
All VoIP	1,21%	0,84%	0,03%	1,12%	0,86%	0,67%	0,79%	1,67%
H.323	0,03%	0,00%	0,00%	1,19%	0,09%	0,23%	6,20%	0,12%
SIP+RTP	40,45%	49,55%	12,02%	31,54%	39,97%	38,11%	6,39%	44,40%
Skype	42,35%	20,51%	87,98%	48,84%	29,33%	32,08%	80,97%	32,23%
IM-VoIP	17,16%	29,94%	-	18,40%	30,60%	29,58%	6,42%	23,24%
Other	0,01%	0,00%	0,00%	0,03%	0,01%	0,00%	0,01%	0,00%



across all regions as for instance for P2P. Even an exotic messaging service like the Chinese QQ generates 44 percent of all IM in Southern Africa, most likely due to a large Chinese minority living and working there.

Traffic distribution of instant messaging protocols

Protocol	Southern Africa	South America	Eastern Europe	Northern Africa	Germany	Southern Europe	Middle East	South-western Europe
All IM	0,04%	0,06%	0,00%	0,02%	0,16%	0,03%	0,50%	0,08%
IRC	1,50%	56,57%	18,32%	11,16%	15,36%	17,90%	0,04%	57,21%
QQ	43,88%	0,27%	0,45%	2,87%	0,44%	0,60%	0,14%	1,44%
XMPP	2,53%	1,54%	16,63%	20,29%	8,64%	6,35%	25,31%	2,15%
MSN	26,84%	20,29%	19,65%	34,87%	3,61%	21,22%	0,24%	31,23%
OSCAR	0,48%	2,78%	41,80%	12,36%	57,80%	6,30%	3,65%	1,09%
Yahoo	24,47%	18,56%	3,15%	18,41%	13,64%	47,62%	70,61%	6,85%
Other	0,30%	0,00%	0,00%	0,04%	0,52%	0,00%	0,00%	0,01%



Most IM clients offer the capability to exchange files. The actual use of these capabilities differs widely between the different IM clients and distorts the volume-based statistics. This distortion can be eliminated by looking at user numbers instead. Due to privacy reasons, user data for this analysis is only available from South America. The table below shows the percentage of subscribers using a certain IM service next to the generated data volume relative to all IM traffic. The last column shows the ratio of the two values.

Both the Microsoft (MSN) and Yahoo Messengers have a ratio of close to 1. OSCAR, that is used by AOL's instant messaging systems ICQ and AIM, is known to be used for file transfers and thus shows a slightly larger ratio of 1.58. For the XMPP protocol, also known as Jabber and used, for instance, by Google Talk, it is only 0.16, meaning that there is much less traffic per users. At the time of the measurement, file transfers with Google Talk and other XMPP clients were nearly impossible, which explains the low per-user data volume. IRC is often used for file sharing as an alternative to P2P. Unsurprisingly, IRC exhibits a much higher per-user traffic than all other IM protocols with a ratio close to 5.

Instant Messaging

Key Facts

- No single most popular protocol
- Higher local popularity variance as for other protocols
- High per-user volume for IRC due to file exchange

Each of the six major instant messaging (IM) protocols show a significant popularity (more than 20 percent of all IM traffic) in at least one region. There is no clear leader

¹ The PRX Traffic Manager in Eastern Europe was running an outdated firmware version, thus this value is not available there.

Relative IM traffic volume, user numbers and ratio of these two values

Protocol	South America Volume	South America Users	Volume/User Number Ratio
IRC	56,57%	11,74%	4,82
XMPP	1,54%	9,52%	0,16
MSN	20,29%	21,34%	0,95
Oscar	2,78%	1,76%	1,58
Yahoo	18,56%	18,64%	1,00

Media Streaming

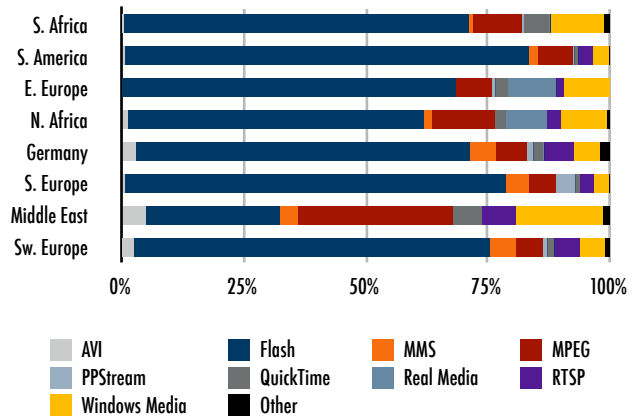
Key Facts

- Flash comprises over 60% of all streaming
- Only MPEG and Windows Media Stream have significant share in all regions
- Streaming is taking over P2P users for video content

The growing popularity of audio and video streaming is undeniable. Adobe's Flash platform, that is used to embed multimedia content into Web pages, is the uncontested number one. It makes up between 60 and 83 percent of all streaming traffic in all but one regions. Video portals such as YouTube, currently ranked as the third most popular Web site by Alexa¹, use Flash to distribute commercial and user-generated video content. Only in the Middle East, Flash remains below 30 percent, most likely because access to some sites whose content is deemed inappropriate by governmental regulators is blocked by ISPs. This is probably the reason for the unusually high share of MPEG and Windows Media streaming data.

Traffic distribution of media streaming protocols

Protocol	Southern Africa	South America	Eastern Europe	Northern Africa	Germany	Southern Europe	Middle East	South-western Europe
All streaming	5,83%	7,81%	7,34%	8,72%	7,17%	9,55%	4,64%	10,14%
AVI	0,52%	0,61%	0,00%	1,25%	2,92%	0,61%	4,87%	2,53%
Flash	70,49%	82,95%	68,52%	60,77%	68,43%	78,11%	27,56%	72,98%
MMS	0,90%	1,77%	0,00%	1,59%	5,23%	4,65%	3,57%	5,22%
MPEG	10,06%	7,08%	7,23%	12,74%	6,53%	5,65%	31,93%	5,53%
PPStream	0,50%	0,48%	0,91%	0,05%	1,29%	3,96%	0,00%	1,06%
QuickTime	5,04%	0,63%	2,51%	2,32%	1,91%	0,87%	5,74%	1,31%
Real Media	0,40%	0,01%	9,91%	8,43%	0,43%	0,01%	0,00%	0,05%
RTSP	0,11%	3,02%	1,49%	2,76%	5,98%	2,91%	7,17%	5,23%
Windows Media	10,83%	3,22%	9,43%	9,62%	5,15%	3,01%	17,75%	5,11%
Other	1,15%	0,23%	0,00%	0,46%	2,14%	0,22%	1,42%	0,98%



The growth of streaming traffic will probably continue in the near future. It provides a user-friendly alternative to get access to audio and video content without the hassle of having to download it first. Specifically P2P will be displaced further as it requires the installation of client programs, which involves the risk of spyware or other malware. Streaming is simply easier to use for the average Web surfer.

Web Browsing

Key Facts

- Web usage is second behind P2P across all regions
- File hosting (DDL) has increased to up to 45% of all Web traffic
- HTTPS usage below 3.5% in all regions

Web traffic is the second largest traffic contributor after P2P worldwide. Over 90 percent of all network nodes use the HTTP protocol (see section 'User Number Statistics'). HTTP, the protocol used for accessing Web pages – excluding any embedded audio or video streaming content (as for instance YouTube Flash movies), that is covered in 'Media Streaming' – accounts for the majority of Web traffic, ranging from 55 percent in South America to 86 percent in Southern Africa.

The share of Web traffic is higher than last year. This indicates a higher growth in popularity than other services such as P2P have experienced. One reason certainly is the surging popularity of social networking sites such as MySpace, Facebook and LinkedIn. Independent from popularity, the average size of Web pages is constantly growing. For example, a typical news portal page such as CNN.com or NYTimes.com has a size of 1-2 Mbytes. HTTPS², the encrypted version of the HTTP protocol used for online banking and software-as-a-service (SaaS) sites such as

¹ Alexa Internet is a company gathering Web site usage data to generate a popularity ranking. <http://www.alexa.com>.

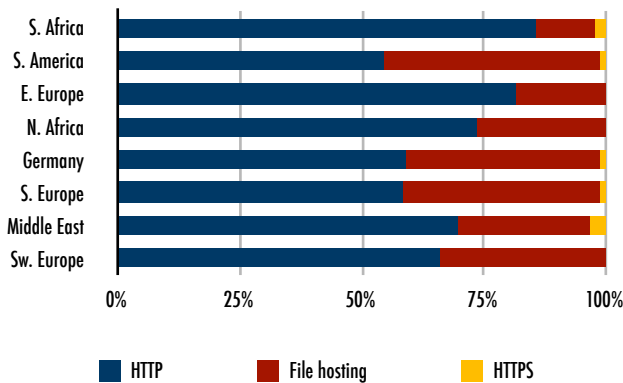
² Note that HTTPS detection was only activated in five out of the eight monitored regions.

Salesforce.com, only generates little traffic – between 1 and 3.2 percent.

File hosting sites such as RapidShare and Megaupload generate a substantial amount of Web traffic – between 12 percent in Southern Africa and 44 percent in South America – contributing up to 10 percent to the overall Internet traffic. File hosting services are mainly used to share and download large files. Due to privacy concerns, this study does not include the ranking of the various file hosting services as in the 2007 study. Assuming a similar traffic distribution among file hosting sites as last year, it is fair to say that RapidShare alone is responsible for 5 percent of the worldwide Internet traffic.

Traffic distribution of Web usage

Protocol	Southern Africa	South America	Eastern Europe	Northern Africa	Germany	Southern Europe	Middle East	South-western Europe
All Web	20,93%	18,17%	16,23%	32,65%	25,78%	25,11%	34,49%	23,29%
HTTP	85,52%	54,55%	81,57%	73,47%	58,90%	58,32%	69,70%	65,99%
File hosting	12,11%	44,26%	18,43%	26,53%	39,78%	40,54%	27,08%	34,01%
HTTPS	2,37%	1,19%	-	-	1,33%	1,14%	3,23%	-



Tunnel Protocols

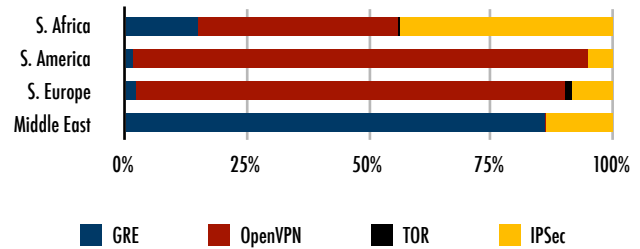
Key Facts

- OpenVPN most popular
- Nearly no Tor traffic
- GRE popular as VoIP tunnel in the Middle East

Tunnel protocol data are only available from four regions. OpenVPN is by far the most used tunnel protocol in South America and Southern Europe with shares of around 90 percent. Southern Africa shows similar amounts of IPsec and OpenVPN traffic – slightly above 40 percent. In the part of the network we monitored in the Middle East, OpenVPN has been blocked during the measurement period. GRE is used there as an alternative tunneling protocol, most likely by small commercial SIP operators to tunnel IP voice calls.

Traffic distribution of tunneling protocols

Protocol	Southern Africa	South America	Southern Europe	Middle East
All tunnels	0,16%	0,10%	0,09%	2,74%
GRE	15,01%	1,71%	2,23%	86,01%
OpenVPN	40,92%	93,20%	88,03%	0,28%
Tor	0,55%	0,00%	1,27%	0,03%
IPsec	43,52%	5,09%	8,47%	13,67%



Standard protocols

Key Facts

- E-mail protocols >40% of standard protocols
- E-mail is about 0.5% of all Internet traffic
- High Usenet data volume in Germany
- FTP and SMTP are the two busiest protocols across all regions

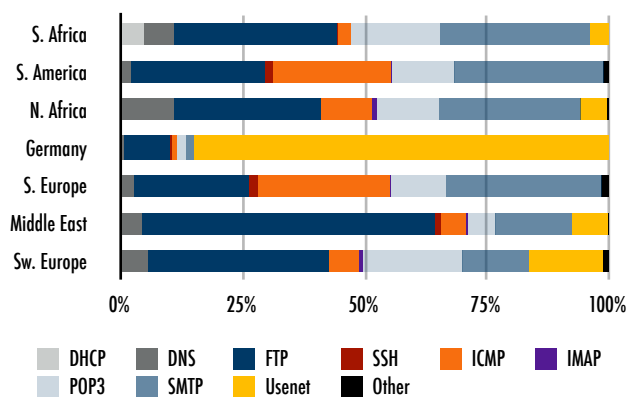
The standard e-mail protocols (SMTP, IMAP, POP3) combined make up about 0.5 percent of a country's Internet traffic and over 40 percent of the standard protocol traffic in four out of seven monitored regions. In Germany, an unusually high Usenet volume of 85 percent – equivalent to over 4 percent of all German Internet traffic – strongly distorts the other values. The Usenet is a distributed Internet discussion system based on the NNTP protocol. In the early days of the Internet it was used to exchange e-mail-style discussion posts through Usenet servers. Later, the capability to exchange binary data – just like e-mail attachments – was added. This functionality is now utilized by commercial Usenet access providers, most prominently UseNeXT¹, that position themselves as alternatives to P2P and file hosting services for large downloads. At least in Germany, this message seems to get through to file sharers.

Apart from this artifact in Germany, FTP is the single most busy standard protocol throughout all regions with a peak of 60 percent in the Middle East. While this sounds a lot, it is only 1 percent of all traffic. Also notable is the high proportion of ICMP traffic in South America (24 percent) and Southern Europe (27 percent). Most likely this is caused by customer premises equipment (CPE), such as DSL and cable modems, using ICMP messages for monitoring or troubleshooting purposes.

¹ UseNeXT: <http://www.usenext.com>

Traffic distribution of standard protocols

Protocol	Southern Africa	South America	Northern Africa	Germany	Southern Europe	Middle East	Southwestern Europe
All standard protocols	1,31%	0,49%	0,89%	4,89%	0,52%	1,83%	1,23%
DHCP	4,62%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%
DNS	6,20%	2,01%	10,95%	0,65%	2,60%	4,31%	5,46%
FTP	33,52%	27,53%	29,96%	9,34%	23,68%	59,94%	37,12%
SSH	0,00%	1,55%	-	0,52%	1,72%	1,43%	-
ICMP	2,73%	24,25%	10,55%	0,86%	27,10%	5,05%	6,29%
IMAP	0,15%	0,05%	1,03%	0,05%	0,18%	0,37%	0,81%
POP3	18,22%	12,95%	12,69%	2,04%	11,29%	5,65%	20,32%
SMTP	30,68%	30,60%	29,05%	1,60%	31,93%	15,72%	13,61%
Usenet	3,74%	0,00%	5,45%	84,88%	0,00%	7,32%	15,15%
Other	0,14%	1,04%	0,33%	0,06%	1,51%	0,20%	1,23%



Average Packet Sizes

Key Facts

- Average packet size in South America is 556 bytes – 789 bytes down-, 281 bytes upstream
- Largest downstream packets (> 1,450 bytes): file hosting, Flash and QuickTime
- Largest upstream packets (1,034 bytes): QQLive; followed by OpenVPN (494 bytes)
- Smallest downstream packets (<61 bytes): ICMP, Oscar, QQLive
- Smallest upstream packets (<50 bytes): MMS, QuickTime, Steam
- Largest difference downstream-upstream (>1,400 bytes): QuickTime, file hosting, Flash

For the first time, we have recorded the average packet size for all encountered protocols, separately for downstream and upstream traffic. This statistic is only available for South America.

Many results are as expected. The average downstream packet size is considerably larger than the upstream packet size. File downloads and video streaming have large downstream packets often close to the Internet-typical MTU of 1,500 bytes. Instant messaging and VoIP protocols have smaller and more symmetric average packet sizes. Only the IM protocols IRC and XMPP exhibit a pronounced asymmetry hinting at file downloads from clients on the egress side of the monitored network. Classic client-server protocols

(FTP, HTTP, POP3) have large downstream and small upstream packets.

But some results appear to be more unusual. QQLive, a P2P streaming protocol, has extremely large average upstream packets of 1,034 bytes on average. No other protocol comes even close to this. Second is OpenVPN with only 494 bytes. Apparently, an unusually high amount of data – even for P2P – is leaving the participating peers. A likely explanation is that many foreign users watch streaming content provided by QQLive peers in the monitored ISP's network.

Average packet size per protocol and direction

Protocol	Downstream Packet Size (Bytes)	Upstream Packet Size (Bytes)	Average Packet Size (Bytes)	Protocol	Downstream Packet Size (Bytes)	Upstream Packet Size (Bytes)	Average Packet Size (Bytes)
All	789	281	556	Pando	931	359	688
Ares	718	454	589	POP	1032	72	672
AVI	1415	108	1015	PPLive	1067	325	691
BitTorrent	771	370	587	PPStream	602	387	496
DirectConnect	672	266	479	QQ	440	109	308
DNS	155	64	107	QQLive	56	1034	417
eDonkey	793	413	610	QuickTime	1474	48	1015
File hosting	1468	54	1014	RDP	149	94	124
Flash	1454	50	974	Realmedia	1441	51	966
Freenet	228	136	182	RTP	825	125	723
FTP	1303	148	902	RTSP	1211	54	824
Gnutella	636	241	452	SIP	156	144	150
GRE	405	218	324	Skype	214	196	205
H323	79	101	89	SMTP	103	169	138
HalfLife2	188	86	138	SopCast	602	214	423
HTTP	1153	187	758	Soulseek	759	536	654
ICMP	61	55	58	SSH	667	83	410
IMAP	656	221	374	SSL	740	229	512
iMesh	459	471	465	Steam	1438	48	1013
IPSEC	432	166	297	Telnet	202	62	129
IRC	299	58	177	Thunder	623	320	476
Manolito	1008	343	852	Tor	848	360	649
MMS	1207	48	788	TVAnts	677	194	455
MPEG	1393	51	940	TVUPlayer	516	131	346
MSN	348	278	312	UUSee	776	96	471
NTP	76	76	76	VNC	113	190	150
OGG	1159	53	802	Windows Media	1355	51	886
OpenVPN	635	494	560	World of Warcraft	248	56	161
ORB	179	57	103	XMPP	357	96	224
Oscar	59	70	65	Yahoo	181	201	189



Contact

ipoque GmbH
Mozartstr. 3
D-04107 Leipzig
Germany

Tel.: +49 (341) 59 40 30
Fax: +49 (341) 59 40 30 19

E-Mail: pr@ipoque.com
URL: www.ipoque.com

About ipoque

ipoque is the leading European provider of deep packet inspection (DPI) solutions for Internet traffic management and analysis. Designed for Internet service providers, enterprises and educational institutions, ipoque's PRX Traffic Manager allows to effectively monitor, shape and optimize network applications. These include the most critical and hard-to-detect protocols used for peer-to-peer file sharing (P2P), instant messaging (IM), Voice over IP (VoIP), tunneling and media streaming, but also many legacy applications.

For further information, please visit www.ipoque.com.

Copyright Notice

Copyright by ipoque 2009.

All images and data can be used for publication free of charge. An attribution of the source is required.

A copy of any print publication and a link to www.ipoque.com or to www.ipoque.com/resources/internet-studies for any online publication is appreciated.