22. ISP Platforms Under a Heavy Peer-to-Peer Workload

Gerhard Haßlinger (T-Systems, Technologiezentrum Darmstadt)

22.1 Introduction

Peer-to-peer (P2P) applications presently contribute the main part of the traffic volume on Internet access platforms in Europe and North America. Distributed file sharing systems first emerged as a widely used application supported by a number of protocols, where the size of most popular networks counts millions of nodes being involved in mutual online data exchange. In addition, large voice over IP networks are using Peer-to-Peer technology and more applications based on a Peer-to-Peer overlay structure are expected to become popular. This has ambivalent consequences on the business models of service and network providers. Peer-to-peer networking is a driving demand of broadband access, motivating many users to subscribe to ADSL access with 1 - 6 Mbit/s line speed as presently offered by Deutsche Telekom and other Internet service providers (ISPs). At the same time Peer-to-Peer overlays also open the market for many additional services as a technology to make them globally present for a small investment in protocol development.

In this context, an efficient transport of data in Peer-to-Peer networks is crucial for the success of Peer-to-Peer applications and is also a main concern of ISPs and especially the network management on the IP layer. Some early versions of Peer-to-Peer protocols were subject to a large messaging overhead. Measurement in the Gnutella network [517] attributed a minimum overhead of 6 kbit/s per connection. In addition, search queries got flooded through the Gnutella network and thus limited its scalability. The current version of Gnutella as well as other popular file sharing networks like eDonkey, BitTorrent or FastTrack have reduced the routing overhead and become scalable with millions of interacting peers. Nevertheless, a significant messaging overhead often remains; see chapter 24 on traffic characteristics and performance evaluation.

The data search and selection of sources for downloading establish routing mechanisms on the application layer, which are usually independent of the routing on the IP network layer. The distribution of data over the Peer-to-Peer network and the (mis-)match of routing strategies on different layers is essential for an efficient bandwidth utilization and data throughput on IP platforms. On the other hand, the diversity and variability of Peer-to-Peer

R. Steinmetz and K. Wehrle (Eds.): P2P Systems and Applications, LNCS 3485, pp. 369-381, 2005.

protocols make it difficult to monitor them or even to control and optimize their performance on the management level of IP platforms.

In this chapter we will have a closer look at the implications of Peer-to-Peer applications on traffic and network management for IP platforms with broadband access. Besides a brief discussion of monitoring and traffic analysis of Peer-to-Peer applications, the main focus is on the efficiency of resource usage and consequences of a dominant Peer-to-Peer workload on the quality of service in IP networks, which support an expanding variety of services.

22.2 Peer-to-Peer Traffic Characteristics

22.2.1 Traffic Mix on IP Platforms

From 1999 Napster has offered a platform for file sharing, which generated a considerable portion of traffic ($> 20\%$) on IP networks in the USA within a few months. Despite the shut-down of Napster due to copyright infringements and persisting problems of illegal content distribution, file sharing traffic has continuously increased until it became the dominant source of traffic [555].

Table 22.1 shows some representative measurement results for the components of the Internet traffic mix in Europe in 2003-2004. Based on the evaluation of TCP ports, more than half of the traffic is attributed to Peer-to-Peer applications. In addition, the Peer-to-Peer traffic portion becomes even larger when observed at application layer [475] e.g. from 50% to almost 80% as reported in [40]. In recent time, most of the FastTrack protocol has been replaced by BitTorrent activity in Deutsche Telekom's and France Telecom's traffic statistics, while eDonkey is still dominant [492].

22.2.2 Daily Traffic Profile

The application mix on IP networks also varies according to the time of day: web browsing (HTTP) oscillates between a peak in heavy traffic hours in the evening and almost no activity for some hours after midnight. Figure 22.1 illustrates the main traffic portions of Peer-to-Peer, web browsing (HTTP) and other applications, which again have been distinguished via TCP standard ports.

Peer-to-peer applications again dominate the traffic volume and at the same time show an overall smoothing effect on traffic profiles as compared with client-server architectures for several reasons:

– Traffic variability over time:
  The daily traffic profiles in broadband access platforms typically show high activity during the day time or in the evening. For Peer-to-Peer traffic, the

Application Mix from TCP Port Measurement	Deutsche Telekom 1st Half 2004	France Telecom 2003	France Telecom Sept. 2004	CacheLogic at an EU Tier-1 Provider in June 2004
eDonkey	60 %	38 %	∼ 54.5 %	∼ 20 %
FastTrack	6 %	8 %	∼ 1 %	∼ 10 %
BitTorrent	-	-	∼ 3.5 %	∼ 16 %
Other Peer-to-Peer	4 %	4 %	∼ 1 %	∼ 10 %
All Peer-to-Peer	70 %	50 %	∼ 60 %	∼ 56 %
HTTP	10 %	15 %	-	∼ 12 %
Other (non-Peer-to-Peer / unknown)	20 %	35 %	-	∼ 32 %

Table 22.1: Port measurement of Peer-to-Peer traffic in Europe

ratio of the peak to the mean rate is usually smaller than 1.5 due to background transfers which often last throughout the night. Web browsing and many other applications have a ratio of 2 or higher. The ongoing Peer-to-Peer data transfers through the night time are initiated by long-lasting background downloads of large video files with sizes often in the Gigabyte range. When peers are connected via ADSL access lines, the throughput of Peer-to-Peer transmission is limited by the upstream speed of the peers. Thus it requires hours or days for a peer to download a file of Gigabyte size at a rate of about 100 kbit/s when peers are staying continuously online.

– Traffic variability over network topology:
The popularity of many Internet servers is often changing dynamically. Traffic sources may spontaneously arise and vanish at different locations in the network topology, where servers are attached. On the contrary, the nodes in large Peer-to-Peer networks are more or less uniformly distributed over the access area. In search phases the Peer-to-Peer protocols often involve supernodes which are similar to servers as traffic sources. Downloads on the other hand run completely distributed among the peer nodes. While spontaneous accesses from many clients to a server can lead to bottlenecks, frequently referenced data is soon replicated and afterwards downloaded from many nodes in a Peer-to-Peer network. Hence, Peer-to-Peer applications lead to a more uniform distribution of traffic sources over the network independent of sudden changes in the popularity of material on the Internet and the locations of originating nodes.

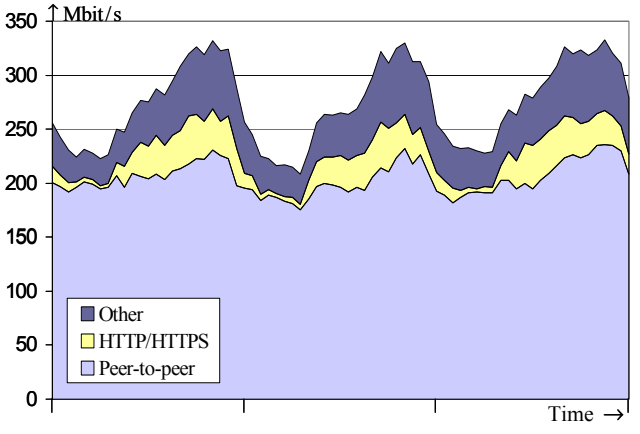

Fig. 22.1: Usual profile of Peer-to-Peer, HTTP and other traffic on Deutsche Telekom's IP platform measured over three days in the 2. half of 2003

– Variability due to different access speeds:
  Most of the Peer-to-Peer source traffic originates from subscribers with a limited access rate especially for uploads. The upload speed is presently in the range of 128 - 512kbit/s even in most broadband access lines. Owing to upload speed limitation and a preference for splitting the transfer of large files into many data chunks, which can be transmitted in parallel TCP connections to and from the peers, the traffic of Peer-to-Peer applications is subdivided into a large number of small flows. Thus the preconditions for the smoothing effect of statistical multiplexing are strengthened due to the increased multiplexing degree for Peer-to-Peer applications. As a consequence, the burstiness of aggregated traffic on network and especially backbone links is significantly reduced.

When the user population and thus the number of parallel sessions is constant and transfer activities in the sessions are independent of each other, then the central limit theorem is applicable to the rate statistics of aggregated traffic being composed of a large number of small and independent flows. Remote access routers at the boundary of IP platforms already aggregate the traffic of hundreds or several thousand users.

As a consequence, the coefficient of variation, i.e. the ratio of the variance to the square of the mean, becomes smaller with a higher level of aggregation and the traffic rate approaches a Gaussian distribution [40, 280]. Simple link dimensioning rules are available for Gaussian traffic with regard to quality of service aspects [278, 281].

22.2.3 Traffic Growth and Prognosis

In recent years, the deployment of broadband access lines for the mass market together with extensive use of file sharing applications pushed the traffic volume [459]. From 2000 to 2003 Deutsche Telekom supplied several million homes with ADSL access and the traffic on the Internet platform increased by a factor of more than 100 over this 3 year period, coming close to the scalability limits of mature carrier-grade routing equipment available on the market. Meanwhile the broadband access penetration is continuously expanding with higher access speeds being offered, while the traffic growth rate is flattening. In general, there are still undoubted requirements for larger bandwidths in telecommunication networks, whereas the future development of Peer-to-Peer traffic is difficult to predict.

Most video files are currently using MPEG compression in order to adapt to limited transmission capacities on account of reduced quality. Television studios presently demand high resolution for their video transmissions in high definition television (HDTV) quality together with coding schemes without loss of information. The corresponding transmission rate amounts to several Gbit/s for a single video stream. Nowadays IP backbone and access speeds would have to be increased about 1000-fold for a widespread transport of video in HDTV quality. Although many people may nowadays be satisfied with low video quality, improving quality for video and further emerging broadband applications will continue to increase demand for even larger bandwidths for at least the next decade.

With regard to Peer-to-Peer applications, the illegitimate use of copyright protected content and the future effect of countermeasures are unknown factors of influence. In addition, legal downloads and video streaming are currently being offered on client-server architectures via Internet under acceptable conditions for the mass market, which may then partially satisfy the requirements of present Peer-to-Peer users. Moreover, scalability and many security aspects including resistance against denial-of-service attacks seem to be handled without much care by current Peer-to-Peer protocols.

On the other hand, the superior efficiency of the Peer-to-Peer principle for fast widespread distribution of large amounts of data is attractive for software distribution, especially when updates are frequently required, e.g., for virus scanning programs. There is a high potential for supporting various upcoming applications by Peer-to-Peer networks including online gaming, radio and TV over IP etc., with some candidates for drivers of future traffic growth among them.

22.2.4 Asymmetrical Versus Symmetrical Access Lines

On the whole, Peer-to-Peer traffic is symmetrical in up- and downstream directions, since an upstream flow from a sources always corresponds to a

downstream flow to the destination when no data is lost during transmission. Since many users are at first interested in downloading files, Peer-to-Peer protocols have to take care of a balanced transfer in both directions. Nowadays, protocols in widespread use start to upload data in parallel to a download as soon as the first data chunks of a downloaded file have been received. Moreover, they enforce a give and take policy for each participant such that users with a higher upload volume are preferred when they are requesting downloads.

On the other hand, broadband access lines are usually asymmetrical with smaller upstream rate, e.g., 128 kbit/s versus 1 Mbit/s. Therefore, the upstream capacity in broadband access platforms becomes a bottleneck for Peer-to-Peer applications. In measurement of upstream traffic the Peer-to-Peer portion is very high, since most other transmissions go downstream.

Symmetrical access lines would be more appropriate for Peer-to-Peer data exchange, but if service providers were to replace ADSL lines with symmetrical DSL at comparable speed, then Peer-to-Peer traffic could be expected to increase with the upstream speed. This shows that enforcement of a close relation between up- and download volume of each user is essential to achieve a high throughput of Peer-to-Peer traffic on broadband IP platforms. An adaptation of Peer-to-Peer video streaming to asymmetrical access is studied in [646].

There are other symmetrical broadband applications e.g. video telephony and conferencing, which demand symmetrical access or at least a more balanced ratio of up- to downstream capacity.

22.3 Cross Layer Aspects

22.3.1 Routing on Application and IP Layer

Peer-to-peer networks establish their own routing on the application layer when searching and selecting data sources. They are different from the routing principles on the network layer with consequences for the efficiency and scalability. Earlier versions of the Gnutella network simply forwarded broadcast search requests to all neighbours within a limited distance measured by the hop count, i.e. the number of intermediate nodes. In this way, a large amount of messages were exchanged for each search, which imposed restrictions on the scalability of the Gnutella network [517].

Presently, Peer-to-Peer protocols often introduce hierarchically structured search nodes to reduce the overhead together with functions known from IP routing like hello messages to confirm connectivity to the peers. Nevertheless, a significant messaging overhead can be observed in large Peer-to-Peer networks. The implementation of two independent routing layers causes inefficiency to the extent of the mismatch between different underlying network structures.

22.3.2 Network and Transport Layer Analysis

Monitoring and analysis of Peer-to-Peer traffic is essential for Internet service providers

- to determine the components of the traffic mix, which may indicate shifts and trends in the usage of applications,
- to estimate the overhead in Peer-to-Peer messaging and
- to analyze the Peer-to-Peer network structure and transmission paths to be compared with the routing on the network layer.

The components of traffic can be analyzed by evaluating information in the IP and TCP header, i.e. IP addresses, TCP ports and information about the TCP signaling and connection status. Peer-to-peer protocols are often associated with a standard port or a range of specific ports, which allows them to be recognized on the transport layer. When the search phase relies on a small set of servers then the corresponding requests can be traced by an analysis of IP addresses. The analysis on this layer is supported by sampling mechanisms provided by the router equipment and can be performed at line speed even on backbone links.

However, most widely used Peer-to-Peer protocols can be configured to use any arbitrary TCP port and thus an essential part of Peer-to-Peer traffic cannot be categorized on this layer, which runs over non-standard TCP ports or even standard ports of other applications. Dynamically changing ports are another variant. Recently, the portion of unknown TCP ports is increasing in transport layer statistics, as can be observed e.g. in the Internet2 NetFlow statistics [318].

22.3.3 Application Layer Pattern

To get better insight into Peer-to-Peer traffic components, application level analysis is necessary which also inspects the payload of transmitted packets. Depending on the protocol, patterns can be recognized, which indicate data transmitted in Peer-to-Peer search phases or the initiation of data exchanges. Afterwards they can be traced on the transport layer until the termination of a corresponding TCP connection [285, 160]. Table 22.2 shows examples of patterns observable in current versions of popular file sharing protocols [336].

In general, the analysis on this layer is complex and cannot be performed permanently for the complete traffic on high speed links. Recognition patterns have to be identified for each relevant Peer-to-Peer protocol and, in addition, the analysis has to be updated whenever new protocols or new variants of existing protocols emerge. Although there are approaches for the analysis of several currently used protocols, a complete classification of Peer-to-Peer traffic remains difficult since some protocols have developed techniques to

P2P Protocol	eDonkey2000	FastTrack	BitTorrent
Transport	TCP & UDP	TCP & UDP	TCP
Standard Ports	4661 – 4665	1214	6881 – 6889
Block Sizes	10,240 Byte	65,536 Byte 2,048 Byte	16,384 Byte 32,768 Byte

Characteristic Pattern in Packets	eDonkey: `0xe3`
	eMule: `0xc5`
		TCP:
		`GET /.hash;`
		`GIVE;`
			`GET /announce?info.hash;`
			`GET /torrents/;`
			`GET TrackPak;`
			`0x13 BitTorrent;`
			`0x00000005;`
			`0x0000000d;`
			`0x00004009`

Table 22.2: Characteristic protocol pattern in IP packets and their payload

disguise their purpose also on the application layer. Anonymous Peer-to-Peer actions are addressed in the chapter on security. Nevertheless, several manufacturers of measurement equipment offer tools for monitoring traffic including Peer-to-Peer application layer analysis to some extent.

Besides the analysis of the components of complete data flows, another approach of application layer analysis has been carried out by crawling into Peer-to-Peer protocols [534]. To do this, a node is inserted into a Peer-to-Peer network, which can collect data about the connectivity and structure of the network as well as the status of other nodes.

22.3.4 Distribution of Sources for eDonkey File-Sharing

Peer-to-peer networks have no knowledge of the infrastructure of the underlying IP network structure. Thus, protocols on the application layer can be developed independently of the lower layers. Nevertheless, layered protocols should be coordinated in order to make the transport efficient. One aspect is the length of paths from source to destination in Peer-to-Peer downloads. Figure 22.2 shows an example of source locations, which have been selected for downloading a popular file in the eDonkey network to a destination in Darmstadt, Germany.

On the whole, the sources of the download are partly concentrated in the near of the destination, most of them in Europe, but nevertheless some are spread over the globe and only a minority is located on the IP platform of the same service provider. Thus, most of the traffic for the download originates from a number of remote autonomous systems and is routed as off-net traffic

through peering points and the backbone of the IP platform to which the destination is attached.

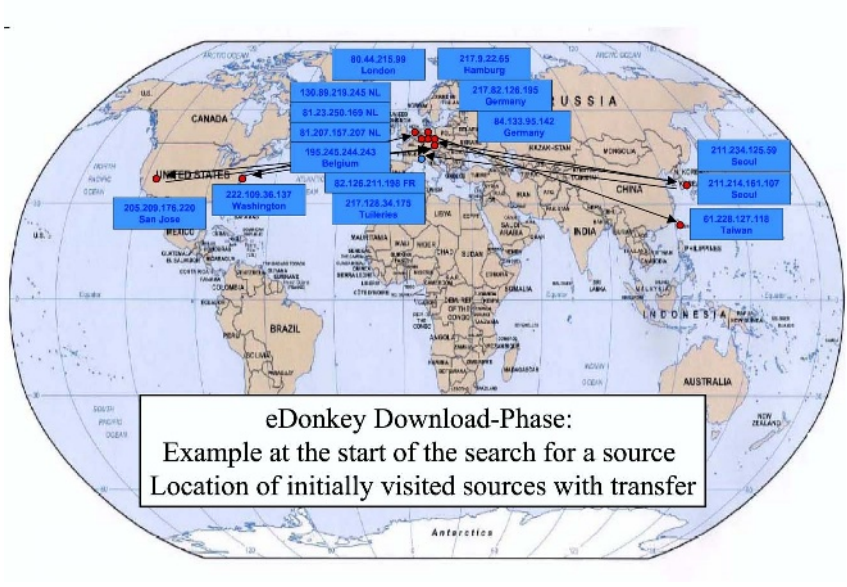

Fig. 22.2: Globally Distributed Sources for Downloading a File with eDonkey

On the other hand, Figure 22.3 illustrates the usual access and backbone structure of broadband access providers. Tree-shaped access areas are attached to the backbone at points of presence (PoPs), where remote access control routers handle the registration and sessions being set up for the users. For large provider networks serving millions of subscribers, it can be expected that a majority of the data of global file sharing systems can already be found to be replicated at some sources on the same ISP platform and often even in the same access area. This especially holds for the most popular and referenced data, since it is observed that the major portion of downloads comes from a small set of very popular files. Thus the source distribution of figure 22.2 indicates unnecessarily long transmission paths increasing the traffic load between autonomous systems and in backbone areas. This leaves potential for more efficient data exchange when a better match of network structures on the application and IP layer could be achieved.

In the considered example, a Linux software file was downloaded. The situation may be different when most audio and video data is transmitted by people in some country in their own language. France Telecom observed that a major part of the file sharing traffic in their Internet platform is locally to France, as can be expected by a differentiation of communities by languages [196].

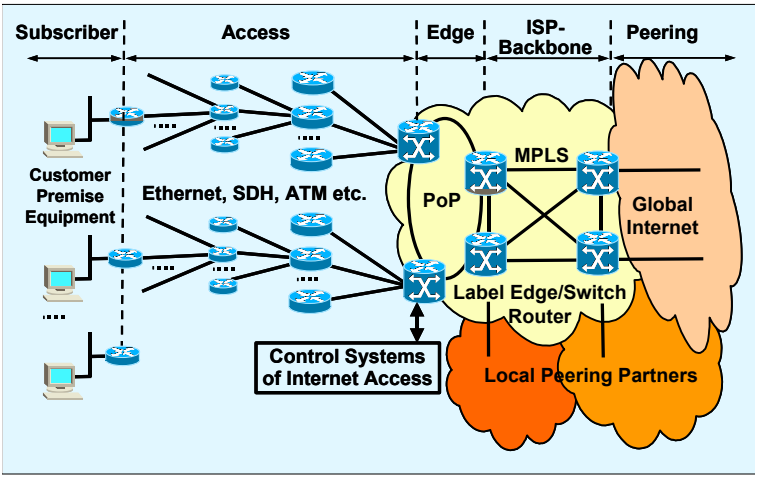

Fig. 22.3: Basic structure of ISP Platforms for Internet Access

22.3.5 Caches for Peer-to-Peer Data

Web caches provide an opportunity to optimize traffic flows. Usual web caches do not apply to Peer-to-Peer traffic and have become inefficient. On the other hand, caches can be set up specially for Peer-to-Peer traffic. Therefore, they act as a proxy node in a Peer-to-Peer network, which stores a large amount of data. A major problem of usual caching is that data in the cache is often expired, while it has already been updated on the corresponding web site. Peer-to-peer file sharing systems are not subject to expired data since data is referenced via unique hash identifiers.

Web caches are not intended to play an active role in Peer-to-Peer networks. They should be transparent and used only to shorten transmission paths to distant nodes, but should not be directly addressed as a source in the Peer-to-Peer network. When a download is requested for some data chunk available in the cache, then the cache can respond instead of a source, which has been selected in the search phase. For transparency reasons, the data should be transferred as if it originated from the selected source with regard to

- the IP addresses,
- the upstream access bandwidth of the source and
- possible changes in the status of the source, e.g. accounting for balanced up- and download volume of Peer-to-Peer network nodes.

However, caches cannot be made completely transparent. A data transfer from the cache will usually have a shorter transmission delay and a cache will

not be able to match the available upstream rate of the original source, including time-varying bottlenecks on the transmission path beyond the cache. But at least the online availability and the access speed of the original source should be taken into account. In fact, the upload capacity of the caches will substitute a part of the upload capacity of nodes in the Peer-to-Peer network with consequences for the total data throughput. The efficiency of caches depends on the source selection by the Peer-to-Peer protocol. In principle, unnecessary load on backbone and expensive international links can be avoided.

By this method, caches for Peer-to-Peer traffic have to be adapted to discover data for the most popular protocols in use. They do not reduce the messaging overhead in the search phase. An alternative approach has been taken by eDonkey, where caches of service providers can be directly included as a configuration option of the Peer-to-Peer protocol. An open issue for caching again lies in its partial use for illegal content, which was already a problem before Peer-to-Peer became popular, but is becoming more serious with file sharing.

22.4 Implications for QoS in Multi-service IP Networks

The Internet has developed from data transfer applications to a service integrating platform with steadily increasing variety of service types including file transfer, email, web browsing, voice over IP, Peer-to-Peer data exchange etc. Each service type has its specific quality of service (QoS) demands regarding bandwidth, transmission time, e.g., real time constraints, as well as tolerance for transmission errors and failure situations. Peer-to-peer data exchange is usually of the best effort service type without strict QoS demands. Downloads often run for several hours or days in the background.

Although shorter transfers or even real time transmissions would be desirable for some Peer-to-Peer applications, users are aware that economic tariffs in a mass market impose access bandwidth limitations such that broadband transfers require considerable time even with increasing access speeds.

On the other hand, the impact of Peer-to-Peer traffic on other services has to be taken into account. The present traffic profile in IP networks with a dominant Peer-to-Peer traffic portion of the best effort type suggests that the differentiated services architecture [102, 618, 76] is sufficient as a simple scheme to support QoS by introducing traffic classes to be handled at different priorities.

Since presently less than 20% of the traffic in ISP networks seems to have strict QoS requirements including voice over IP and virtual private networks (VPN), sufficient QoS could be guaranteed for those traffic types by a strict forwarding priority. Even applications like web browsing and email are

included in a 20% portion of traffic with the most demanding QoS requirements.

When the network is dimensioned for an essentially larger traffic volume as is generated only by the preferred traffic classes, then they will not suffer from bottlenecks and queueing delay in normal operation. Vice versa, the impact of preferring a small premium traffic class on a much larger portion of Peer-to-Peer traffic is moderate.

The delivery for premium traffic classes can even be assured in some failure situations, e.g. for single link breakdowns, provided that restoration links are available in an appropriate network design. Since overload may occur on those links, the best effort traffic will then often be affected.

Nowadays Peer-to-Peer protocols can cope with temporary disconnections on the application layer and recover transmission from the last current state afterwards. When file transfers via FTP or HTTP protocols are interrupted, an essential part of the transmission is often lost and a complete restart of the transfer may be required. Segmentation and reassembly of large data files into small chunks improves reliability and efficiency of Peer-to-Peer transfers, which is essential for non-assured QoS of best effort transmission.

An obstacle for the application of differentiated services is the difficulty to classify Peer-to-Peer traffic. A treatment with lower priority based on TCP port numbers will increase the tendency to disguise Peer-to-Peer applications by using randomly chosen ports for unknown protocols or by transporting Peer-to-Peer data exchange e.g. over the HTTP port for web browsing. Therefore, the only efficient way to classify traffic seems to be through a declaration and marking of the complete premium type traffic by the users themselves or by the originating servers, combined with a corresponding differentiated tariff scheme. But even then unresolved problems remain for supporting QoS for inter-domain traffic and for QoS-sensitive traffic which is transferred over peering points into a service provider platform.

22.5 Conclusion

Despite of increasing the traffic volume on the Internet, Peer-to-Peer traffic has a smoothing effect on the variability of the traffic rate and the daily profiles on broadband platforms. The geographical distribution of sources becomes more uniform as compared to most other traffic types. Potential for more efficient transport can be seen in unnecessarily long transmission paths due to the source selection in popular file sharing protocols.

These properties partly facilitate the network dimensioning and planning for service providers, but the monitoring, analysis and control of the Peer-to-Peer components and the prediction of their future development remains difficult.

The present traffic mix with a dominant portion of best effort type data exchanges in the background has implications for the quality of service concept, suggesting that differentiated services [1] are sufficient to support QoS-sensitive traffic types. Presently such traffic types generate less traffic volume, even including applications like web browsing. A comprehensive and appropriate classification of service types is still subject to many unresolved issues.