A Business Model for Charging and Accounting of Internet Services

Helmut Kneer[1], Urs Zurfluh[1], Gabriel Dermler[2], Burkhard Stiller[3]

[1]Institut für Informatik IFI, University of Zürich, Switzerland, {kneer, uzurfluh}@ifi.unizh.ch
[2]IBM Research Laboratory, Zürich, Switzerland, dermler@zurich.ibm.com
[3]Computer Engineering and Networks Laboratory TIK,
Swiss Federal Institute of Technology, ETH Zürich, Switzerland, stiller@tik.ee.ethz.ch

Abstract. The fast growing development of electronic commerce within the past few years and the increasing number of multimedia applications for the Internet contribute to traffic congestion and bottlenecks on the Internet. Since bandwidth is and will remain a scarce network resource in the future, the demand for charging and accounting mechanisms for the Internet arose.

Besides from the implementation of charging and accounting mechanisms for Internet services, it is essential to define a business model for electronic business applications. This paper introduces a business model that defines and characterizes the business entities and describes their roles and functions within an e-commerce scenario. Based on these definitions, a technology model applies the different Internet technologies available today.

1 Introduction

Within the field of electronic commerce a variety of services and products is offered and purchased by means of an underlying network infrastructure that connects customer and merchant. Since the Internet was introduced in 1969 by a military project of the American Defense Department (ARPANET), it was developed to span most parts of the world today. Hence, the Internet as a global network infrastructure is ideal to do business electronically.

The critical point with Electronic Business is the delivery and the performance across the Internet since too much traffic traverses through the backbones of the ISPs' networks. Packet-switched Internet applications are very susceptible to congestion since their quality strongly depends on network parameters such as data throughput, bandwidth, and latency. Unfortunately, packet forwarding in today's Internet only works on a best-effort base which results in a poor quality for multiplexed audio or video streams, especially in times when the network is congested.

One solution to overcome this problem is to classify Internet traffic and charge the customers according to the transport service the ISPs provide. Usage-based pricing schemes consider charges for the actual amount of consumed network resources [17]. Charging and accounting of Internet services effects the Internet traffic in such a way that the classified traffic of the different applications is treated fairly according to their service level and price [19]. This means that the ISPs have to establish a basis for a multi-service-level Internet in which different service levels are offered and the desired Quality-of-Service (QoS) can be guaranteed. This would help (1) the ISPs to cover their operating costs by collecting usage fees from their customers and (2) the customers to receive the desired QoS that they chose and paid for. IP telephony, as an example for such a multi-level Internet application, was investigated in [20].

Apart from the implementation and realization of charging and accounting mechanisms, there has to be a business model that defines processes and phases from the initiation of the electronic business to the final payment for a product or service. Three levels of abstraction are proposed within this paper that are essential for charging and accounting of Internet services, (1) the business level, (2) the contract level, and (3) the network level. The business model describes on the business level what parties there are, what roles and functions they have, and what business relationships are essential for performing electronic business. Contracts and agreements are necessary between the parties to stipulate the business conditions and to declare the service performance. On the contract level, the parties are considered liable entities that are responsible for the delivery of the service according to the defined service conditions. Network components and architectures are explored on the network level forming the underlying network infrastructure and implementing designed business processes of the business model that were contractually defined on the contract level.

It is the cooperation of the many entities, the widespread business environment, and the competition among the parties that make the whole business model a challenge for charging and accounting on the Internet. There are trusted and un-trusted relationships among the business entities and some parties do not even know each other, e.g., the end-customer does not know all of the intermediate ISPs that are engaged in establishing the network connection to the ESP. There are security risks and threats for all involved parties, which are analyzed and identified in order to develop a security architecture [12] based on the business model [10] and the trust model [23]. The security architecture introduces a concept for secure business transactions based on five security-relevant aspects.

This paper is organized as follows. Starting on the business level, the business model is described in Section 2 for the e-commerce scenario between an end-customer, an ESP, and intermediate ISPs. Section 3 illustrates on the network level technical possibilities and characteristics of the Internet protocol architecture, which is important for the actual implementation of charging and accounting mechanisms. Section 4 introduces on the contract level the usage of Service Level Agreements (SLA) as a mechanism to ensure QoS based end-to-end communication, before a summary of results will conclude the paper in Section 5.

2 Business Model

A business model includes the course of business events and reflects the processes within an economic system. The economic system introduced in this paper consists of a simple business relationship between a customer and a merchant doing business over the Internet. In terms of electronic business systems, the merchant can be modelled by an Electronic Commerce Service Provider (ESP), and the customer by an end-customer, which simply means that this is an end point of a communication connection (cf. Figure 1).

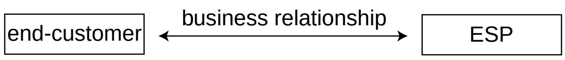

Figure 1: E-commerce business relationship.

The ESP offers products, contents, and services online via the World Wide Web (WWW) and represents the merchant or the seller of a good. Products offered by the ESP include physical goods that could also be purchased in stores, *e.g.*, books, cars, CDs, etc. Contents offered by the ESP comprise digital information as non-tangible goods in form of bits and bytes. Content includes for example the digital content of an online book or the digital version of a CD that can be downloaded onto the end-customer's local disk. Furthermore, there is a broad spectrum of multimedia contents and services (*e.g.*, audio-on-demand or video-on-demand) offered by the ESP online to the customers across the Internet. A CD online shop could offer the service to pre-listen to a CD before the customer decides to make the purchase. Another form of electronic commerce service offered by the ESP is the possibility for the end-customer to make data backups onto the ESP's storage medium.

The offers of the ESP can be retrieved by end-customers from all over the world using a web browser. Thus, the end-customer represents the buyer of a good. The end-customer may be a private individual or a commercial customer such as an enterprise or a university.

The basic idea of an economic system between an end-customer and an ESP is well understood and theoretically feasible. Problems arise with the introduction of performance and QoS for the delivery of Internet services. Internet Service Providers (ISP) provide the physical infrastructure for the Internet but also the technical network equipment like routers, switches, and network management software and form the basic foundation of all electronic commerce activities. ISPs provide the transport of data packets over the Internet between end-customer and ESP. However, traffic congestion and unreliable connections on the best effort Internet describe a problem for transmitting data, especially if high bandwidth and reliable throughput is required (*e.g.*, multimedia applications). ISPs could overcome this problem by charging the end-customers according to the transport service they request.

The focus of this paper is on the business model for realizing the business relationship between the ESP and the end-customer using ISPs for charging and accounting of transport services. Therefore, an e-commerce scenario was developed with an end-customer and an ESP on both ends, and several intermediate ISPs. Figure 2 shows the business relationships between all the involved parties of such a scenario. There might also be a payment provider involved in this scenario that is responsible for the financial clearing between the parties. The payment provider can be represented by a bank, a credit institution, or a Trusted Third Party (TTP).

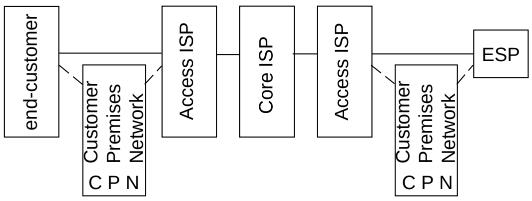

Figure 2: Relationships between involved parties of an e-commerce scenario.

The ISPs are divided into Access ISPs and Core ISPs according to their scope of duties. Access ISPs support Local Access Networks and provide Internet connections to the end-customer, be it directly or through a Customer Premises Network (CPN). Core ISPs increase the reach of Access ISPs to a global extent and form the backbone of the Internet. They perform the data transport service interconnecting Access ISPs. There may be more than one core ISP involved in a communication connection between end-customer and ESP, depending on the connectivity of ISPs and their local distances. In case there is only one ISP involved between end-customer and ESP, it acts as both Access ISP and Core ISP providing the data transport service. Thus, Access ISPs and Core ISPs may be physically similar.

It is possible that end-customers are affiliated to a Customer Premises Network (CPN), *e.g.*, a LAN of an enterprise or a university. A CPN represents a group of users in terms of a common policy and conceals the individual end-customer from the Access ISP. A CPN may offer additional private applications or extra conditions for data transport services to its end-customers. An end-customer can be connected either to a CPN or directly to the Access ISP, if the end-customer is a residential user. There may be multiple end-customers involved in one communication connection in case of a conference system.

2.1 Phases of the Business Model

In order to realize such an e-commerce scenario, as described in Figure 2, the involved parties need to cooperate and interact with each other. General rules and business conditions have to be set up between the end-customer, the ESP, and the involved ISPs. The business processes within the business model can be described by four coherent phases, an initial contracting phase, a reservation phase, a service phase, and a final clearing phase. Figure 3 (a) shows graphically the coherence of the different phases that are described in the following.

Contracting Phase. The contracting phase is the initial phase for the business entities before any products or contents are purchased or any services are performed. The different cooperating parties need to get to know each other, establish or maintain business connections to their partners, and arrange business conditions and contractual agreements for later businesses and services. Four different kinds of business relationships are investigated within this paper that need to be set up before any electronic business can be done, assuming that there is no CPN involved and the end-customer is a residential user connected directly to her Access ISP. There are relationships between (1) end-customer and ESP, (2) end-customer and Access ISP, (3) Access ISP and Core ISP, and (4) Access ISP and ESP (as can be seen in Figure 2).

The relationship between the end-customer and ESP does not require contracts or agreements. The end-customer simply logs in at the ESP's homepage to buy products or contents or to order a service. The log in process could be performed anonymously where the end-customer does not reveal any personal information about her identity. If the ESP offers the service to register its frequent customers, the end-customer can sign up declaring personal business related data. The ESP might have user profiles of its customers to keep track of their records in order to make individual offers of products or electronic commerce services tailored to the end-customers' needs.

The business relationship between the ESP and its Access ISP is very similar to the one between the end-customer and her Access ISP. The main difference is the fact that the ESP is always known to its Access ISP since both parties have fixed locations and anonymity would hinder the performance of the business.

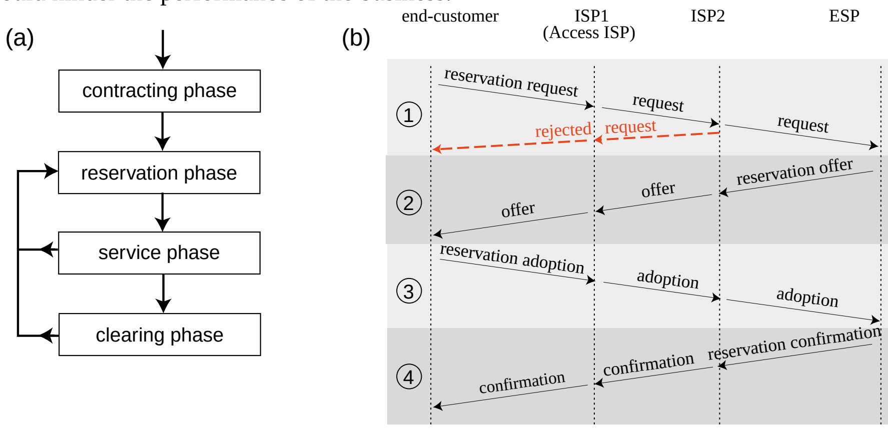

Figure 3: (a) Sequence of coherent phases of the business model.
(b) Reservation messages exchanged between the different parties.

The relationship between the end-customer and the Access ISP is established as soon as data needs to be transferred between the ESP and the end-customer. It is a special relationship since the Access ISP works as a broker or general contractor that offers the end-customer access to the Internet and represents a contact point for any data transport service to a sending or receiving ESP. Skeleton agreements are set up between the end-customer and the Access ISP that stipulate future business conditions and legal aspects. Within the contracting phase, the Access ISP creates user accounts for the end-customer which contain among others charging and billing information, monthly spending limits, or special types of reservations (*e.g.*, fixed weekly reservations of a data transport service). If the end-customer wants to remain anonymous and unknown to the Access ISP, the contracting phase may be skipped.

The relationship between Access ISP and Core ISP, or two general ISPs respectively, is formed by negotiations on traffic contracts and SLAs that regulate on the contract level the amount of incoming and outgoing traffic through the network of an ISP. The ISPs strive for maintaining a dense network infrastructure for their customers in order to deliver a data transport service with a guaranteed QoS. SLAs between the ISPs are made statically before any data transport service is performed but they can be changed and adjusted dynamically during operation if the network situation changes throughout the service delivery or the end-customer wants to change service parameters. Service Level Agreements for QoS-based communication is discussed in detail in Section 4 on the contract level.

Reservation Phase. The end-customer has decided to buy a content or service with the ESP and thus needs an Internet connection to exchange data with the ESP. Within the reservation phase, the end-customer reserves network resources with her Access ISP since it represents the starting point for subsequent network transport.

In a first sub-phase, the login phase, the end-customer chooses her Access ISP and requests a data transport service to the ESP. In a second sub-phase, the specific information phase, the Access ISP could provide its end-customer some interesting information about the short-term market situation, network traffic, or price comparisons with other ISPs. This information could vary for different registered or anonymous end-customers. Auctions could be used to find the best/cheapest/shortest connection between the end-customer and the ESP on a market where bandwidth and network resources are offered and sold [17]. The Access ISP, as a broker or a general contractor for the end-customer, is responsible for the execution and maintenance of the delivered service. Traffic contracts and SLAs were set up in the contracting phase already to organize traffic flows through the networks of the ISPs.

The actual reservation of a data transport service or a simple price query takes place within the negotiation phase, as a third sub-phase of the reservation phase. The reservation of network resources for single flows on the Internet is performed in four steps according to the RSVP protocol [25], [7].

In the first step, the end-customer sends out a signal to her Access ISP to indicate a reservation request (cf. Figure 3 (b)). This request contains information about the kind of reservation (*e.g.*, application, point of time, and duration), a scheme for charge sharing (end-customer pays only, ESP pays only, both pay), and parameters defining the QoS for the application. These parameters could include among others the bandwidth indicating the performance (bit/s), the transfer quality (max. fault-probability, encryption), routing parameters, or the maximum delay time. The flow specification, or flowspec, of RSVP contains the service requirements for the application and characteristics of the traffic stream or flow [3], [21], [14]. The request is sent across other (Core/Access) ISPs (in case there are other ISPs) to the ESP, which replies with a reservation offer, if it accepts the end-customer's request. If one ISP or the ESP does not agree upon the request, it may reject the request and send the denial back to the end-customer (dotted line in Figure 3 (b)). In this case, the negotiation starts anew.

In the second step of the negotiation process, the ESP replies to the reservation request with a reservation offer informing the end-customer on acceptance or denial of the reservation request. The reservation offer is sent the same way back to the end-customer, where different ISPs reserve the requested resources for that particular flow or they add pricing information in case it was only a price request [20]. If one party does not accept the reservation offer, the end-customer must re-send the reservation request with different parameters.

The third and fourth step of the negotiation phase (reservation adoption and confirmation) are optional and can be used to add sender or receiver provided electronic payments if, for example, reservation fees are required or payments have to be made in advance [7]. This four-step negotiation phase has to be run through for every newly requested service or if the end-customer wants to extend the currently performed service.

Service Phase. After the data transport service has been reserved within the previous reservation phase, the actual service can be performed within the service phase. The data transport service is delivered by the Access ISPs and eventually several intermediate Core ISPs according to the previously defined QoS parameters.

Of particular importance for ISPs and the payment of the service is the question, whether the end-customer is a known business partner, and thus can be trusted, or whether she decided earlier in the contracting phase to remain anonymous. In case the end-customer has registered with her Access ISP earlier in the contracting phase, charging, billing, and eventually payment for the service is done in the subsequent clearing phase (by post-paid payments).

If the end-customer is an anonymous business partner, she has to pay before the service is performed due to a lack of trust and security. Pre-paid payments could be made (1) from pre-paid accounts, (2) online with debit, credit, or some pre-paid money card, or (3) by some sort of electronic money that could eventually be included in the reservation adoption message (for end-customer payments) or the reservation confirmation message (for ESP payments). Depending on the charge sharing scheme, both end-customer and ESP pay their share for the service. If ISPs do not get money from the end-customer/ESP beforehand, they can immediately release the allocated resources for the reserved service. After the delivery of the service, the end-customer can go back to the reservation phase and request a new service or try to extend the currently performed service with similar QoS parameters.

Clearing Phase. The clearing phase contains the charging, billing, and the payment for the delivered data transport service. If the end-customer decided to remain anonymous, payments were made already within the service phase and thus the clearing phase can be skipped. Charging and billing includes the process of transforming the collected accounting records for the end-customer into monetary units and summarizing them on a bill [7]. The bill contains the used network resources and the delivered services within a certain period of time (e.g., one month) and the corresponding pecuniary value.

Different billing and payment schemes are possible and discussed in [9]. The best possibility is to perform billing and payment through the Access ISP who collects single bills from ISPs to combine them to one bill and redistributes the payment back to corresponding ISPs. Another possibility is that every involved ISP sends a separate bill to the end-customer and receives the payment for the service. There might even be a payment provider involved in the whole clearing process as a Trusted Third Party (TTP) assuming control over billing and payment. Payments can be made with debit or credit card, or simply by a transfer of funds.

The clearing phase is terminated when every ISP received its money. The end-customer can go back to the reservation phase and start reserving a new service again. An optional sub-phase for customer support could be appended which gives the ISP(s) the chance to receive feedback from the end-customer (*e.g.*, via questionnaire) regarding the performance of the service or special customer-related problems and suggestions (*e.g.*, special services for frequent customers).

2.2 Billing of the Content

Although this paper is focusing on charging and accounting mechanisms for transport services on the Internet, it is important to consider charges for the content combined with the transport. This subsection gives a quick description of relevant factors that decide, whether it is favorable to charge content together with the transport service or not. Figure 4 (a) shows the primary influencing factors, namely the relative value of content (defined as the value of content divided by the cost of transportation) and the required end-customer anonymity. Not the amount of bytes settle the value of the content, but the market price.

For low values of content it is more efficient to charge and bill the content together with the transport service either on one bill or through a pre-paid calling card. A separate bill for the content will be provided for registered end-customers, if the relative value of content is higher than the middle threshold (approximately around 10). If the end-customer likes to be anonymous and the relative value of content is between the low and the high threshold (approximately between 5 and 50), pre-paid money cards can be used to pay for the content. Extremely valuable content (above the high threshold) requires pre-paid accounts on the ESP's side with secret access codes. Both middle and high values of content should only be charged when the transfer was successful, *e.g.*, a big data file was received correctly.

Figure 4 (b) shows secondary influencing factors deciding, whether to bill separately for the content or together with the transport service. There is the number of logically independent content blocks per time and the value of content per time. Logically independent blocks of content are discrete and contain direct useful information. A feed of financial information (*e.g.*, stock rates) consists of many independent data packets which are billed continuously. A backup file or a high resolution image file can be seen as one big block of content. If only one intermediate packet gets lost, the transmission failed and no billing should be performed.

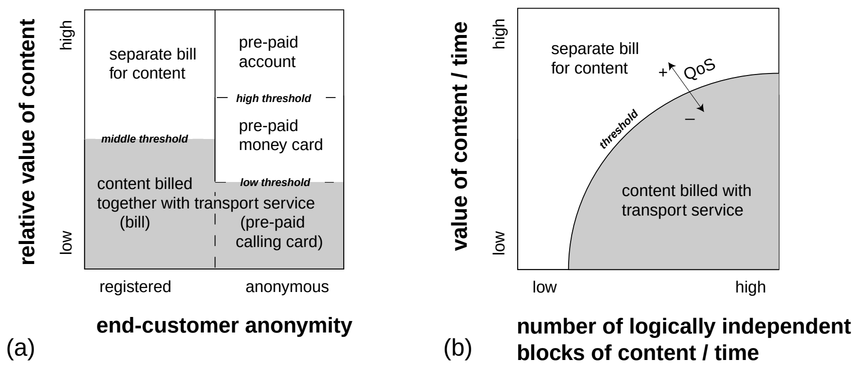

Figure 4: Influencing factors for a separate billing of the content.

Transmission of a high number of logically independent blocks of content with a low value of content (lower right corner of Figure 4 (b)) is billed together with the transport service, if the QoS for the transport service is below a certain threshold. If the QoS for the transport service increases (QoS threshold moves to upper left in Figure 4 (b)), the relative value of content decreases and billing of the content is performed with the transport service.

3 Internet Protocol Architectures

Providing charging and accounting services on the Internet requires a number of technical prerequisites. As of recently, the Internet has performed on a non-commercial basis and charging services have not been necessary or foreseen. However, with the commercialization of the Internet as a networking infrastructure and the Internet services offered, this point of view changes. In particular, once an end-customer has to choose from, say, two different service classes, a best-effort one and another one delivering some sort of bandwidth guarantees, a purely technical solution of providing these classes is not sufficient anymore. The reason is hidden in the greedy nature of almost every, certainly the majority of end-customers – they will choose the service class with the best QoS. Of course, if this is the case, the service class with less QoS will become obsolete, since it is not used. In turn, users encounter similar problems within the better class of service due to its heavily congested usage. This situation will remain unchanged as long as no financial incentives on choosing a service class, which is perfectly suited for the end-customer's needs, are provided by the Internet.

Today's Internet does not offer any service differentiation mechanisms, since the best-effort type of service still dominates. In addition, the basic protocol is defined by the Internet Protocol (IP) [16], which is currently used in its version 4 and does not provide any service class differentiation features besides the Type-of-Service (TOS) field. Nevertheless, this field is only optionally used within IPv4 on a broad scale. Some future enhancements, or better changes, to determine IPv6 is being prepared by the IETF [5], including a flow label field [15] and an extended TOS field. However, today two distinct Internet protocol architectures are used as approaches towards a service discriminated Internet.

3.1 Integrated Services Internet (IntServ)

IntServ defines a networking framework, which supports unidirectional end-to-end flows [4]. These flows may request a certain QoS and may use a Controlled Load Service [24] or a Guaranteed Service [18]. However, every flow needs to establish a context between the sender, the receiver, and intermediate nodes. Therefore, the Resource ReSerVation Protocol (RSVP) [3] has been defined as a protocol for reserving network resources for single flows of applications. RSVP allows sender and receiver to specify the desired traffic class in terms of a flow specification, mainly including bandwidth, delay, and loss characteristics. Furthermore, it propagates these information along a path in the Internet which will be set up for the subsequent data packets. This scheme relies on the existence of admission control, resource allocation, and packet forwarding mechanisms in each router to ensure that the requested QoS parameters can be guaranteed. In addition, RSVP assumes that a QoS-based routing protocol exists.

In summary, the IntServ model comprises a per-flow reservation of network resources for single Internet applications from one end-customer to another with a certain pre-defined set of QoS parameters. Advantageous features of the IntServ and RSVP approach encompass a receiver-driven QoS specification, the support of multicast traffic and its merging for reservations, and the soft state approach for maintaining the context data of a flow. However, the support on a per flow-basis is assumed to show scalability problems with respect to large number of flows and states to be kept in large backbone routers. The per-flow granularity imposes overhead which may not be necessary for a certain number of situations. Service classes in IntServ distinguish between best-effort and guaranteed services. The application of the business model as described in Section 2 has been performed for guaranteed services, which require a resource reservation based on RSVP.

3.2 Differentiated Services Internet (DiffServ)

Due to these assumed scalability and overhead problems in case of many single flows, a different framework was developed recently. Instead of treating a single flow as the entity of interest, the Differentiated Services Internet (DiffServ) handles Internet traffic based on the notion of aggregated flows and fixed numbers of service levels in terms of service profiles [2]. This approach minimizes the state to be kept in routers. In addition, this is supported by a domain-concept, where a group of routers implements a similar number of service levels and the appropriate set of policies. This DiffServ domain is defined by a fixed boundary, consisting of ingress and egress routers. However, traffic traversing such a DiffServ domain is required to be marked. This marking happens on a per IP packet-basis at the ingress routers and utilizes the DiffServ (DS) field in an IP packet [13]. This DS field replaces the TOS field from the IPv4 protocol and allows the definition of per-hop behaviors (PHB), which in turn determines the service level such a packet will be treated by. Once the DS Code Point as part of the DS field has been set, the packet “tunnels” through the DiffServ domain being treated equally in every interior router. Similarly as within IntServ, an aggregated DiffServ flow of flows is considered to be unidirectional.

Since single end-to-end flows are bundled to aggregated flows with a similar behavior within a DiffServ domain, the DiffServ approach requires less overhead. However, the need to mark IP packets at the DiffServ borders remains. In addition, a longer termed service contract may be required between different DiffServ domains, since a certain service level may be required. This type of flow aggregation in conjunction with service guarantees require some sort of admission control, since an over-utilization may lead to service degradations. Regularly, so-called Service Level Agreements (SLA) are set-up between interconnecting ISPs in order to maintain the desired service level for the aggregated flows. An initial SLA needs to be set up between interconnected ISPs before any service is performed. SLAs can also be adjusted dynamically. Basic details on SLAs are provided in Section 4.

3.3 Comparison and IntServ over DiffServ

As presented above, IntServ as well as DiffServ show a number of advantages and drawbacks. Based on six classification criteria, Table 1 depicts the main differences for IntServ, DiffServ, and best-effort traffic types in the Internet.

Table 1 Comparison of IntServ and DiffServ

	Best-effort	IntServ	DiffServ
QoS Guarantees	no	per data stream	aggregated
Configuration	none	per session (dynamic)	long-term (static)
Zone	entire network	end-to-end	domain-oriented
State Information	none	per data stream (in router)	(none, in BB, in edge router)
Protocols	none	signalling (RSVP)	bit field (DS Byte)
Status	operational	matured	worked on

The integration and combination of IntServ and DiffServ advantages is possible. Local Area Networks (LAN) tend to show an overprovisioning of bandwidth, which does not require a sophisticated resource management and signalling, if a certain topology and traffic considerations are taken into account. The Access Network, however, utilizes RSVP to signal flow requirements from senders to the Core Network. Edge routers perform a mapping of these requirements onto particular flow aggregation types available in the DiffServ core and represented by a dedicated SLA. Since core routers purely perform traffic forwarding based on PHBs, they are able to cope with many aggregated flows. Only edge routers need to keep the state of flows from their local domain.

4 Service Level Agreements for QoS-based Communication

Inter-provider agreements are employed between ISPs to capture the terms of service for exchanged traffic. As in a commercial environment, an offered service is an item which may be sold or purchased. An inter-provider agreement represents a contract-like relationship indicating the characteristics of the covered service and implied financial settlements.

Traditionally, within the Internet, ISPs engage in interconnection agreements in order to assure each other pervasive Internet connectivity. Such agreements allow engaging ISPs to exchange traffic at an interconnection point, either with or without financial settlements [1]. Traditionally, these agreements do not consider QoS related issues.

More recently, Service Level Agreements (SLAs) are used as a means to establish a provider/customer type of relationship between two providers. While the exact content of an SLA often depends on the specific business and technology context, the core of an SLA always includes a description of the traffic covered by the SLA and the service level which is to be applied. SLAs provide the contractual envelope for QoS based assurances.

4.1 DiffServ SLAs

The mentioned IntServ approach of the IETF does not refer to business related entities. In contrast, the DiffServ approach is based on the notion of network domains which can be operated by different ISPs. In a pure DiffServ world, both access and core domains would use DiffServ technology to transfer data within their domains as well as between domains. In order to indicate service commitments between domains, Service Level Agreements (SLAs) are employed.

DiffServ SLAs are defined on the contract level, abstracting from the type of underlying physical network connectivity, which is assumed to be given per se. Due to the ability of DiffServ to classify traffic, SLAs can be defined for various aggregations of flows, for instance based on the source or destination IP addresses. According to the envisaged support of DiffServ for multiple service classes, an SLA also specifies a selected service class. Thus, a DiffServ SLA includes:

- a description of the aggregated flow to which the SLA is applicable,
- the corresponding throughput, and
- the corresponding service class *(e.g.,* assured/premium service, determining QoS, delay or loss characteristics).

In principle, DiffServ SLAs can be defined at any granularity level, including the level of application flows. As DiffServ pursues the goal of high scalability, only a limited set of SLAs is likely to exist between any two domains, such that explicit support of individual flows through DiffServ SLAs will likely be an exception.

4.2 Commercial SLAs

SLAs have emerged in the commercial domain as a result of increasing customer demand to reliably understand what kind of service they actually can expect from an Internet Service Provider. Such SLAs are typically tied to the provision of a network service on a long-term basis. Service provision includes both the installation of physical equipment (*e.g.,* routers, access lines) as well as the provision of a data transport service (*e.g.,* based on IP protocols).

In order to capture QoS aspects, SLAs include similar parameters as considered in the DiffServ case. As an example, the SLA employed by UUNET [22] foresees the following:

- the throughput offered to customers (*e.g.,* T1),
- round-trip latency across the provider's network,
- the availability of the service,

- outage notification duration,
- duration between order and installation, and
- reimbursement procedure in case of non-compliance.

Both mentioned SLA approaches tend to serve the same purpose: establish *longer-term service* relationships between adjacent ISPs and provide assurances for *traffic aggregates* exchanged between them. These approaches avoid the overhead implied by a high number of concurrent SLAs and frequent changes in the terms of agreements.

4.3 Flow-based SLAs

In principle, the SLAs above may be applied at the service interface between Access ISPs and their end-customers. However, given that in the access domains per-flow handling of traffic is a viable option, a different approach is feasible which offers superior flexibility in applying charging schemes to QoS based traffic.

First, it is likely that QoS support within the Internet will, for a long time, not be pervasive. Whenever a QoS based flow is requested, the availability of such support has to be established depending on the source and destination end-points as well as the sequence of ISPs involved. Providing SLAs dynamically on a per-flow basis allows to take such dependencies into consideration and automatically adapt to increasing Internet support for QoS.

Second, considering QoS automatically implies a strong differentiation among Internet services. QoS can be provided at multiple levels (e.g. to support various audio qualities). In consequence, there is a need for differential, QoS dependent pricing in order to prevent users from making use of the best QoS level only. Similarly, applying differential pricing is required in order to reflect the communication path (i.e. ISPs) traversed including the source and destination locations. Both aspects lead to service prices which can be established only if the characteristics of a requested flow are known, i.e. on a per-flow basis.

Third, there is an ongoing discussion about pricing schemes that should be applied to Internet traffic. Dynamic pricing of offered services based on a current level of network usage was shown to improve the service characteristics a network can provide, for instance reduce congestion and smooth traffic [8], respectively. Assuming such an approach, both the implied QoS level and the provided price are to be established dynamically for each new flow.

Providing flow-based SLAs captures all the mentioned issues. Such SLAs are in strong contrast to the ones considered in the previous section. They directly concern application level flows and not aggregations of traffic. Furthermore, they are likely to be set up for the required duration of communication only.

Flow based SLAs are in line with the service model proposed by IntServ, as far as the end-customer's point of view is concerned. We find a significant difference in the motivation stated above. The motivation of considering single flows by IntServ is that customers want to have selectable QoS on a per-flow basis. In contrast, the arguments mentioned above are driven by economic considerations: ISPs want to provide incentives for end-customers to make use of more or less resources in the network and, in case of dynamic pricing, ISPs want to consider the availability of free resources when setting prices. These aspects are best considered in the context of individual demand units, i.e. flows.

5 Summary and Conclusions

Within this paper, a business model was introduced illustrating independent and coherent phases necessary for charging and accounting of Internet services. The business model was developed for an e-commerce scenario with an end-customer, an ESP, and several ISPs. Each of the players has a well-defined and integrated role within the scenario. Different levels of abstraction are necessary to implement the business model for charging and accounting of Internet services. On the top most level, the business level, there is the business model with interacting players doing business with each other. On a second level, the contract level, Service Level Agreements (SLA) are essential in order to provide reliable and accountable QoS-based end-to-end communication. The base of the entire business model is a technology model for the network infrastructure (covering IntServ and DiffServ models) carrying data across the Internet according to the contractual agreements of the contract level. The fact that multiple dynamic ISPs with heterogeneous resources are assembled for the delivery of a guaranteed service justifies the business model with its underlying network model for the communication infrastructure.

One of the key conclusions of this paper is the special role of the Access ISP within the entire business model. The fact that only the Access ISP is visible to the end-customer makes electronic business (1) easier and more trustworthy for the end-customer, since the Access ISP is a well-known point of reference, and (2) easier for the ISPs to charge and account the end-customer for the use of the transport service. If the end-customer decides to be anonymous to the Access ISP, she is invisible to the Access ISP and, thus, has to make payments in advance. The business model included two kinds of payments that have a big influence on the phases of the business model: (1) account-based post-paid payments or (2) immediate pre-paid payments. Depending on whether the end-customer wants to remain anonymous to the Access ISP or not, payments have to be made in advance or the end-customer will receive a bill with a list of the cost for the used services. Consequently, the payment scheme has an important influence on the business model.

Open issues in this area of work are the full end-to-end provisioning of services trespassing multiple core ISPs, which need to offer guaranteed services. These tasks will be an integral part of future investigations of Service Level Agreements and their definitions in a commercial and highly dynamic environment. Particularly, the different views of integrating the advantages of IntServ and DiffServ need to be studied with respect to signalling protocols and their efficient usage for e-commerce scenarios.

Acknowledgements. This work has been performed in the framework of the project Charging and Accounting Technology for the Internet – CATI (CAPIV 5003-054559/1 and MEDeB 5003-054560/1) which has been funded by the Swiss National Science Foundation, SNF, Berne, Switzerland. The authors like to acknowledge many discussions with project colleagues.

References

[1] J. P. Bailey: *The Economics of Internet Interconnection Agreements*, Internet Economics, MIT Press, Cambridge, 1997.

[2] S. Blake, D. Black, M. Carlson, E. Davies, Z. Wang, W. Weiss: *An Architecture for Differentiated Services;* IETF, Request for Comments RFC 2475, December 1998.

[3] R. Braden, S. Berson, S. Herzog, S. Jamin: *Resource ReSerVation Protocol (RSVP) – Version 1 Functional Specification;* IETF, Request for Comments RFC 2205, September 1997.

[4] R. Braden, D. Clark, S. Shenker: *Integrated Services in the Internet Architecture: An Overview;* IETF, Request for Comments RFC 1633, June 1994.

[5] S. Deering, R. Hinden: *Internet Protocol, Version 6 (IPv6) – Specification;* IETF, Request for Comments RFC 2460, December 1998

[6] G. Dermler, G. Fankhauser, B. Stiller, T. Braun, M. Günter, H. Kneer: *Integration of Charging and Accounting into Internet Models and VPN;* Public CATI Deliverable CATI-IBM-DN-P-004-1.5, June 30, 1999.

[7] G. Fankhauser, B. Stiller, C. Vögtli, B. Plattner: *Reservation-based Charging in an Integrated Services Network;* 4th INFORMS Telecommunications Conference, Bocca Raton, Florida, U.S.A., March 1998, Session MC-2.

[8] A. Gupta, D. Stahl, A. Whinston. *The economics of network management*, Communications of the ACM, Vol. 42, No. 9, 1999.

[9] U. Kaiser: *Sicherheits- und Kostenaspekte in elektronischen Zahlungssystemen für RSVP;* Institut für Technische Informatik und Kommunikationsnetze, TIK, ETH Zürich, Student's Thesis, August 1998.

[10] H. Kneer, U. Zurfluh, C. Matt: *Business Model (RSVP);* Public CATI Deliverable CATI-UZH-DN-P-001-2.0, March 12, 1999.

[11] H. Kneer, U. Zurfluh, C. Matt: *Business Roles and Relationships;* Public CATI Deliverable CATI-UZH-DN-P-003-2.0, June 30, 1999.

[12] H. Kneer, U. Zurfluh, C. Matt: *Security Architecture;* Public CATI Deliverable CATI-UZH-DN-P-002-3.0, June 30, 1999.

[13] K. Nichols, S. Blake, F. Baker, D. Black: *Definition of the Differentiated Services Field (DS Field) in the IPv4 and IPv6 Headers;* IETF, Request for Comments RFC 2474, December 1998

[14] C. Partridge: *A Proposed Flow Specification;* BBN, Request for Comments RFC 1363.

[15] C. Partridge: *Using the Flow Label Field in IPv6;* IETF, Request for Comments RFC 1809, June 1995.

[16] J. Postel (Edt.): *Internet Protocol;* IETF, Request for Comments RFC 791, September 1981.

[17] P. Reichl, S. Leinen, B. Stiller: *A Practical Review of Pricing and Cost Recovery for Internet Services;* 2nd Berlin Internet Economics Workshop (IEW'99), Berlin, Germany, May 28–29, 1999. Session "Survey I".

[18] S. Shenker, C. Partridge, R. Guerin: *Specification of Guaranteed Quality of Service;* IETF, Request for Comments RFC 2212, September 1997.

[19] B. Stiller, T. Braun, M. Günter, B. Plattner: *The CATI Project – Charging and Accounting Technology for the Internet;* Springer Verlag, Berlin, Germany, Lecture Notes in Computer Science, Vol. 1629, 5th European Conference on Multimedia Applications, Services, and Techniques (ECMAST'99), Madrid, Spain, May 26–28, 1999, pp 281–296.

[20] B. Stiller, G. Fankhauser, G. Joller, P. Reichl, N. Weiler: *Open Charging and QoS Interfaces for IP Telephony;* The Internet Summit (INET'99), San Jose, California, U.S.A., June 22–25, 1999, Track 4 "Technology", Session "IP Audio".

[21] C. Topolcic: *Experimental Internet Stream Protocol, Version 2 (ST-II);* CIP Working Group, Request for Comments RFC 1190.

[22] *UUNET Service Level Agreements;* available at URL: www.uunet.com/customer/sla

[23] N. Weiler, G, Joller, B. Stiller, G. Fankhauser: *Trust Model;* Public CATI Deliverable CATI-TIK-DN-P-010-2.0, May 28, 1999.

[24] J. Wroclawski: *Specification of the Controlled-Load Network Element Service;* IETF, Request for Comments RFC 2211, September 1997.

[25] L. Zhang, S. Deering, D. Estrin, S. Shenker, D. Zappala, *RSVP: A New Resource Reservation Protocol,* IEEE Networks Magazine, Vol. 31, No. 9, pp 8-18, September 1993.t