Chapter 6
— Pricing and QoS —

Burkhard Stiller, Pere Barlet-Ros, John Cushnie, Jordi Domingo-Pascual, David Hutchison, Rui Lopes, Andreas Mauthe, Mihai Popa, Jim Roberts, Josep Solé-Pareta, Denis Trcek, Carlos Veciana, Lars Wolf

Abstract. In this chapter the state of the art of pricing for Internet services and its relation to Quality-of-Service (QoS) is addressed. Essential economic and technology basics, covering terms, accounting, and security are followed by a user-centered view, a content-based scheme, and a cost sharing approach.
Keyword: Pricing, QoS, Accounting, Content, Customer and Provider Views.

1 Introduction

To offer different Quality-of-Service (QoS) levels within a network implies different prices to be paid for these levels, while allowing users to choose and control what best meets their QoS requirements and payment possibilities. The QoS actually achieved depends on how much network capacity is provided to meet the expressed demand, on the current traffic load, or on QoS mechanisms provided within a network to guarantee requirements. Therefore, the tasks of charging, pricing, and billing for Internet services, including the mobile Internet, have been identified as an important research area over recent years. The global provision of information and services in a private and commercial fashion is a great motivator for industry and academia alike, but to make the most out of this opportunity, efficient methods for charging and billing need to be proposed, developed, and deployed [1], [2], [6], [34], and [15].

For the commercial Internet use, which is managed by network providers and Internet Service Providers (ISP), users may be charged with a specific tariff. These tariffs may be based on traditional schemes, such as time-based or leased-line charges, or on new concepts related to IP networks, such as volume or other traffic characteristics. Moreover, additional scenarios for applying charging are driven by the content offered by an Internet service on top of the transport network, *e.g.*, including web access, e-commerce, video conferencing. Charging for these services involves accurate end-to-end tracing of the service offered, even more in some cases QoS guarantees, inter-network provider, and inter-ISP charging agreements.

Besides (a) the technical service differentiation perspective, two major economic motivations for Internet charging exist: (b) the recovery of the cost of investment to provide the service and (c) the generation of profit for companies providing these services. Finally, (d) the operational view point includes the need of a provider to provide congestion control in the Internet, which is possible besides traditional technical means through differentiating service levels according to price and congestion pricing. With respect to the relation between technology and economic points of views, pricing and provisioning are obviously related since the former must generate sufficient revenue to cover the cost of the latter. In addition, two particularly

different scenarios for providers can be distinguished: the commercial Internet and National Research Network (NRN) providers. Therefore, major influencing factors for pricing differentiated services, QoS-enabled services, or different networks include the following ones:

- **Technical:** QoS guarantees, accounting, and security mechanisms.
- **Economic:** Incentive-compatible pricing schemes, tariff models, and charges.
- **Organizational:** Congestion control mechanisms, customer care, and billing.
- **User-oriented:** Ease-of-use, transparency, and application-independence.

In addition, single or multi-provider differentiation is important to distinguish between inter- and intra-domain activities to take place for providing charged services. The customer (or user) and provider perspective are critical with respect to the transparency and ease-of-use of pricing models and resulting charge calculations. On one hand, a provider is interested to know about QoS, user demands, and prices users are prepared to pay for. On the other hand, providers should offer a range of QoS, associated to different pricing schemes, allowing and asking users to choose what best meets their demands and financial budget. Aspects on user-driven price control and new provider-user relationships will enhance the view of Internet pricing models.

1.1 Terminology

To avoid any misconceptions on terms utilized within this chapter, the following list of definitions is used for key terms in the area of pricing for QoS [39]:

- **Metering** determines the particular usage of resources within end-systems (hosts) or intermediate systems (routers) on a technical level, including Quality-of-Service (QoS), management, and networking parameters.
- **Accounting** defines the summarized information (accounting records) in relation to a customer's service utilization. It is expressed in metered resource consumption, *e.g.*, for the end-system, applications, calls, or any type of connections.
- **Charge Calculation** covers the calculation of a price for a given accounting record and its consolidation into a charging record, while mapping technical values into monetary units. Charge calculation applies a given tariff to the data accounted for.
- **Charging** is an overall term, depicting all tasks required to calculate the finalized content of a bill at a higher layer of abstraction.
- **Pricing** covers the specification and setting of prices for goods, in this chapter specifically networking resources and services in an open market situation. This process may combine technical considerations, *e.g.*, resource consumption, and economic ones, such as applying tariffing theory (prices calculated on a cost/profit base) or marketing.
- **Tariff** defines the algorithm used to determine a charge for a service usage. It is applied in the charge calculation for a given customer and service he utilizes. Tariffs may contain, *e.g.*, discount strategies, rebate schemes, or marketing information.
- **Billing** defines the collection of charging records, summarizing their charging content, and delivering a bill/invoice including an optional list of detailed charges.

2 Economic Basics

To help understand Internet charging and billing issues, the wide range of proposed and existing pricing schemes is discussed. The essential pricing function of return on investment is considered, which is clearly a critical consideration for a commercial operator. Cost sharing is a particularly important issue, when institutions share a common infrastructure, as in the case of an NRN (National Research Network). A brief overview of current projects related to these issues concludes this section.

2.1 Pricing Schemes

While most ISPs currently practice flat rate pricing, at least for residential users, it is widely accepted that the introduction of advanced services with differentiated QoS will lead to the development of more sophisticated schemes [40]. Briefly, the advantages and drawbacks of flat rate pricing are outlined, before discussing respective principles of usage-based pricing, congestion pricing, and service-related pricing. Another overview on pricing may be obtained from [17].

2.1.1 Flat Rate Pricing

Flat rate pricing has the great advantage of simplicity. This facilitates implementation and avoids obvious sources of contention about the bill between customer and supplier. The bill is perfectly predictable and this brings considerable comfort to users who have no need to keep one eye on the clock or byte counter to avoid an unpleasant surprise at the end of the month. Experiments in moving away from the flat rate scheme to some form of usage-based charging have always been met with considerable consumer resistance.

A principal disadvantage of flat rate is the absence of any direct relation between price and cost. There is no disincentive to prevent users generating an excessive amount of traffic thus requiring increased investment in network infrastructure. Flat rate pricing is unfair when users with vastly different usage, *e.g.*, the peer to peer hacker and the Web surfer, must pay the same price while incurring quite different costs. Finally, flat rate is closely associated with best effort service as there is no mechanism to allow any user who so wishes to pay more to avoid the negative effects of congestion.

2.1.2 Usage-based Pricing

The main goal of the usage-based pricing model is charging the usage of the network resources. This usage knowledge is acquired by detailed traffic monitoring. One of the most important traffic characteristic to take into account is the traffic volume. Nevertheless, other traffic characteristics can be also considered to obtain a more detailed charging scheme. Some of these characteristics could be the packet origins and destinations, applications, information contents of the packets, etc. One special type of usage-based pricing is the content-based pricing. This model has several difficulties, such as the impossibility of processing encrypted packets, and the legal restrictions. Furthermore, the number of resources needed for content analysis grows drastically in high-speed links.

Usage-based pricing could be desirable in order to make users aware of the implication of their actions on the network. On the other hand, this is also the main argument against usage-based pricing, since users reduce their network usage when they are charged by a usage-based scheme. This could be positive in NRN environments, but the effect is not clear in commercial networks.

2.1.3 Congestion Pricing

Pricing may be used to ensure that a scarce resource is used to produce maximum utility when shared among a population of contending users having different valuations. This is the principle of congestion pricing.

Many Internet pricing schemes have been put forward to achieve such optimal sharing, a notable example being the "smart market" proposal [30]. This is more an economic ideal than a practical scheme, however. More pragmatic congestion pricing principles are included in the DiffServ architecture. [37] suggested that users could express the value of their packets by choosing between a small number of classes with each class being priced depending on its relative quality level. Users would be obliged to choose expensive classes in times of congestion, if their utility justified the cost, but could revert to cheaper classes in slack periods. The notion of expected capacity filter with "in-profile" and "out-of-profile" packets was introduced [10]. Pricing is based on the parameters of the filter, which can be chosen by the user to modulate the proportion of in-profile packets and thus determine realized QoS in times of congestion. An alternative approach to congestion pricing has been proposed recently [19]. In this scheme, packets are marked when they contribute to a congestion situation and pricing is related to the number of marked packets. Users can modulate their charge by changing their packet rate in response to the current level of congestion.

2.1.4 Service-related Pricing

An alternative pricing principle in a multiservice network is to relate tariffs to required quality levels. Price is related to user-declared traffic characteristics and performance requirements and is logically chosen to reflect the cost of meeting the respective demands. To be credible, such a scheme requires Quality-of-Service differences that are consistent and measurable. Unfortunately, this proves difficult since performance depends in a quite complex way on the statistical characteristics of demand and the amount of resources provisioned to meet that demand. It is relatively easy to ensure excellent Quality-of-Service for all (by over provisioning) but practically impossible to meet precise intermediate levels of performance between excellent and bad. *E.g.*, meeting a packet loss rate exactly between 0.001 and 0.01 without knowing a mathematical relation between demand, capacity, and performance, is not possible.

2.2 Cost Recovery

Installed infrastructure and network operation constitute significant cost items which must be recovered by pricing [41]. A potentially useful model for cost recovery is the charging scheme of the telephone network. Traditional telephone networks are generally uncongested by design so that all revenue comes from non-congestion related pricing, i.e., flat rate or usage-based pricing. Telephone prices are set at a level

such that expressed demand (accounting for price elasticity) is within the limits of available capacity and generated revenue is sufficient to cover the cost of that capacity. Time of day price modulation is used to smooth demand somewhat by giving incentives to use the network in off-peak periods. In general, competition drives providers to operate their networks efficiently with just the right degree of over-provisioning to avoid congestion. To follow the example of the telephone network would require ensuring that sufficient capacity is available in the Internet to meet demand, save in exceptional circumstances. Pricing would then logically be essentially usage-based. Alternative mechanisms, such as flow-based admission control, as discussed in Chapter 1, would be necessary to preserve Quality-of-Service in rare cases of overload.

2.3 Cost Sharing

Volume-based charging and billing schemes have been applied mainly in NRN, addressing cost sharing among their universities and research institutions.

The Case of New Zealand: One of the first examples was the case of New Zealand's NRN [7]. This experience has served as a springboard for many other cost-sharing proposals in other regions. Since 1990 the University of Waikato has operated a single Internet gateway to the United States, charging users by volume to recover the costs. This was very successful, allowing the steady growth of this link speed. Each organization predicted the amount of traffic that it would move during one month. Then a fixed price per unit was paid for that traffic. Also, the real traffic moved was accounted for, and an extra charge per unit was applied for those beyond the predicted traffic. Also, the latest versions of the charging scheme propose different charges for different network services, which can be charged for at a different price.

The JANET Model: JANET, the NRN of the United Kingdom, charged its members by volume, but only for traffic to links in the United States [34], [12]. Charging was applied during high-activity hours of the day in an attempt to reduce traffic and to redistribute traffic to low-activity hours. Traffic classification was based on local Network IP addresses, and every organization charged was able to receive detailed information from some groups of its servers. The system operated until 2000 when it was decided to revise the charging scheme. A combination of fixed and usage-based charges was proposed in order to make the budget process more predictable.

The SWITCH Model: The SWITCH network provides connectivity to universities and research institutions in Switzerland, and it recovers its costs from member institutions [34], [41]. Each institution pays a fixed rate for the connection, only one third of the cost. Moreover, the other two thirds of the cost are charged for based on volume. Therefore, charging is mainly volume-driven. In the case of SWITCH, the effects of volume-based charging have been observed in the behavior of the users.

2.4 Projects Related to Charging

QoS-based charging support, of course in combination with accounting, for premium IP services have also been researched. The SUSIE project [42] resulted in valuable results and observations including user trials and charging schemes appropriate for

DiffServ, IntServ, and ATM networks. It supports an accounting and charging platform suitable for audio-visual services in heterogeneous scenarios with wholesale and retail service providers. For the inter-operation issues, SUSIE has outlined a trade accounting architecture. While integrating and validating accountable IP services, a model for convergence charging has been established based on a usage charged ATM network delivering Premium IP with ATM related QoS and charge prioritized ATM streams selected via a QoS/Price trader. The M3I project [6] developed a Charging and Accounting System (CAS) for Internet services, which utilizes metered data originating from network elements of interest. These data are accounted for and depending on a communicated pricing model and tariff the charges are calculated for the service usage. The MobyDick project [32] applies Diameter to account for mobile service data, which are used in turn to calculate charges.

Finally, the special focus on content-based charging reveals that there are several ongoing projects relevant to this research. Various IETF WGs have a focus on content delivery and content charging. The CDI WG aims to define protocols to allow the inter-operation of separately administered content networks, with several Internet Drafts, *e.g.*, [14] or [13], already released but no RFCs to date. The AAA Working Group [2] is driving the Diameter protocol standard for AAA across peering networks, which is also being considered by the CDI WG as the AAA protocol of choice. The WEBI WG is working on defining the Resource Update Protocol (RUP), which may be adopted by the CDI WG for managing the distribution of content within a CDI network. Within the MPEG standards group [38] the emerging MPEG-7 and MPEG-21 standards complement and contribute to the efforts ongoing with CDI networks. The IST CADENUS project [8] is researching the implementation of SLAs across Premium IP networks, which may also be applied to CDI networks.

3 Technical Basics and Services

The basis for a charging approach is given by a suitable accounting infrastructure, today embedded in existing AAA architecture and protocol proposals, where some of which are available as commercial platforms. In addition, security-relevant services are addressed in combination to ensure that measured data are kept accordingly.

3.1 AAA and Beyond

The IETF and IRTF both have on-going research on an AAA architecture [2] to meet the short- and long-term requirements for the Internet as gathered from the NASREQ (Network Access Server Requirements), MOBILE IP, and ROAMOPS (Roaming Operations) Working Groups. These are currently tending towards Diameter and the earlier RADIUS as the preferred protocols. The need for service requires, in many models, Authentication, to verify a claimed identity, Authorization, to determine if a particular right can be granted to the presenter of a particular credential, and Accounting, to collecting information on resource usage for the purpose of trend analysis, auditing, billing, or cost allocation. Regardless how one function of the three AAA leads to or derives from others, there is common agreement that they are closely interdependent.

Basically, the AAA architecture includes local and home agents and AAA servers that establish secure channels for the purposes of exchanging sensitive (access) information. An agent that attends to the client's request is likely to require that the client provides some credentials that can be authenticated before access to the resources is authorized. Then, accounting information is interchanged. That architecture requires transport-independent AAA protocols meeting requirements on security, scalability, reliability, as well as inter-domain access control. They also have to provide — including a clear documentation — an accounting operations model for each type of network access, the support for IPv6, an explicit proxy support, a data model separated from the protocol, a MIB support, a RADIUS backward compatibility, as well as a full coverage of operational problems by a set of error messages.

Since current AAA architectures, protocols, and implementations do not cope fully with heterogeneous application scenarios and many requirements for various services, ranging from connectivity to content, are not supported, this lack of a generic approach drove the A^x development [35]. It distinguishes between support services and user services and integrates a policy-based management architecture by separating decision points from enforcement points on a per-service basis. So-called A^x services can be offered by a specialized A^x system. A^x services, apart from metering, can be offered from one provider to another because of their future separation based on A^x, enabling providers to build systems driven by their specific business requirements.

3.2 Accounting Platforms

High-speed links that deal with high volumes of traffic belonging to a high number of users from a wide variety of profiles need tools for traffic analysis that can gather traffic information with a high degree of detail. Currently, the most known accounting tools are the Cisco NetFlow technology [9] and the CAIDA's CoralReef Suite [33].

3.2.1 Cisco NetFlow

Cisco IOS NetFlow is a part of the Cisco IOS software for routers and switches [9]. The main functionality of NetFlow is to export the IP information that the router equipment may possess, aggregated in flow records. The basic architecture of the NetFlow has three components: the NetFlow Data Export, the NetFlow Flow Collector (NFC), and the NetFlow Data Analyzer (NDA).

NetFlow's main component is the NetFlow Data Export. It operates inside the routing equipment. It captures and stores traffic under flow records. Periodically, it exports this information to the NFC, which collects the data. The NetFlow Data Export can aggregate the flow information under programmable flow definitions in order to reduce the amount of data to be transmitted. Among other information, it accounts for the number of packets and the number of bytes that belong to the same flow. The IP addresses, transport protocol, and port numbers define the basic flow key-identifier. As flow detection is performed inside the router, additional information belonging to the routing process is able to be stored, *e.g.*, logical interface number, next-hop router, or AS information. Once a flow expires, it will be packed with other flows, into a UDP packet to be transmitted to the NFC. UDP is faster than TCP but can lead to a loss of

data in congested networks. There is, however, overhead traffic on the network that could affect the traffic under monitoring. Also, the router should reserve more resources to store NetFlow information and must send additional packets. This makes NetFlow a possible bottleneck for monitoring with high traffic load. The remaining two components in the NetFlow Architecture are the NetFlow Flow Collector and the NetFlow Data Analyzer, which collect and analyze the data exported by one or more pieces of equipment running NetFlow.

Concerning the potential performance degradation by applying NetFlow in an operational system, this is measured usually by comparing the “no drop rate” of a network with and without NetFlow being enabled. According to the documentation the no drop packet rate is degraded by about 44% in a test of switching 64 byte packets, when NetFlow is enabled on a RSP2 with 128 MB of RAM.

3.2.2 CoralReef

CoralReef is a set of tools developed by the CAIDA group (Co-operative Association for Internet Data Analysis) [33]. Its main functions are traffic capture in high-speed networks, data storage and IP traffic analysis. CoralReef allows traffic capture at different layers, *e.g.*, cells/frames, packets, flows.

Passive traffic monitoring is the main difference between NetFlow and CoralReef. CoralReef performs passive traffic monitoring using optical splitters. No router or switch is performing the capture and analysis. Instead, another computer, which receives a copy of the traffic does all the work. Consequently, no additional data have to be sent by the router and no additional resources need to be reserved inside the network equipment. CoralReef can report information on layer two (cells/frames), layer three (IP packets) and layer four (IP flows). IP flow information is the one that reaches the highest rate of capture and data reduction in full traffic analysis. CoralReef accounts for data from each flow register it has seen in the network. The IP addresses, transport protocol, port numbers, and timestamp for the first and last packet identify a flow register. The information accounted for is the number of bytes and the number of packets belonging to that flow. One can define a maximum time between packets to detect flow expirations, or force flow expiration whenever one deems it necessary.

3.3 Security

QoS inherently includes security issues. As security services are more and more applied to Internet services, they require certain resources and induce costs. In order to set up appropriate pricing for them, tangible means have to exist. This section decomposes existing security services into intrinsic components that can be used for the quantitative management of quality of security services. This forms the technical basis for services to be charged for.

3.3.1 Security Services Taxonomy

Within the context of QoS, security related issues became a topic of research only recently. Assuring security requires proper management of resources, which consequently results in certain costs. To address these issues, appropriate concepts, i.e.

taxonomy for quality of security services (QoSS) is needed. Next, these services have to be decomposed into intrinsic components and cost metrics has to be associated with each of these components. Such metrics form the basis for effective utilization of resources, pricing policies, and charging. Firstly, all variety of security services is decomposed into intrinsic components (i.e. the QoSS taxonomy). Secondly, with each of these components, appropriate cost metrics are associated to enable practical implementations. A similar approach is found in [23], but extended here by an explicit treatment of coupled issues. It should be mentioned that a Public Key Infrastructure (PKI) is considered as well, which is a problem especially for the wireless world owing to extensive computation for certificates and revocation lists [31].

With regards to taxonomy, it is essential to refer to security services using a well-established framework, which is the ISO framework [25]. It defines the following security services: authentication, confidentiality, integrity, non-repudiation, access control, and auditing/alarms. These services are implemented with various security mechanisms that can be grouped as follows:

- Cryptographic primitives (sym. and asym. algorithms, one-way hash functions);
- Access control mechanisms;
- Auditing and alarm mechanisms.

Although quite extensive, these groups are not sufficient. Experiences have shown that the availability of communication systems is very important, especially in the light of growing denial of service attacks [20]. Next, traffic flow confidentiality is also important, but it is almost completely left out from the above standards. Finally, physical security should be included; all equipment and cryptographic keys have to be properly stored and physically protected. Thus, basic groups are given in the relation as of Table 1.

The meaning of fields is as follows: *s-id* is a unique identifier of a security service, *sec-srvc* is the category of security service, and *s-name* is a descriptive name of security service. Other attributes are related to the cryptographic protocol overhead, which are: *ent-id* for identification of entities, *t-nonces* for time-stamps or nonces, *cont* for a content, *cert* for certificates and *CRL* for certificate revocation list. These attributes, measured in bytes, are building blocks of cryptographic protocols. The last attribute is measured in seconds. It is needed to calculate bandwidth usage that is an important issue when PKI related operations for mobile handheld devices are considered. Non-repudiation is left out as it is a compound service, which consists of authentication and integrity.

Table 1: Services Relation with Cost-related Attributes

s-id	**sec-srvc**	**s-name**	**ent-id**	**t-nonces**	**cont**	**cert**	**CRL**	**time**
at#	authentication	...	...	...	...	...	...	...
cf#	confidentiality							
it#	integrity							
tf#	traffic flow							
ac#	access control							

Table 1: Services Relation with Cost-related Attributes

s-id	sec-srvc	s-name	ent-id	t-nonces	cont	cert	CRL	time
av#	availability							
ad#	audit/intrusion detection							
ps#	physical security							

3.3.2 Cost Metrics

Having taxonomy of security services, we can divide them further down to intrinsic components. This first requires identification of mechanisms, which can be later decomposed into generic operations. The definition of mechanisms is given in the relation of Table 2, where *m-id* stands for a unique identifier of a mechanism, *mechanism* for a category of a mechanism, and *m-name* for the name of a mechanism:

Table 2: Mechanism Relations

s-id	m-id	mechanisms	m-name
...	sa#	symmetric algorithm	...
	aa#	asymmetric algorithm	
	hf#	hash function	
	dp#	dummy packets	
	rr#	rerouting	
	lu#	matrix /look-up	
	cs#	code-scanning	
	sc#	software-correctness	
	bu#	bandwidth use	
	la#	og-analysis	
	pm#	physical mechanisms	

Access control mechanisms should not include only access lists and matrixes look-up operations, but also complex firewall operations (payload scanning). Similarly, auditing and alarm mechanisms should not include only pure event logging and related man-power, but also subsequent analysis (deployment of specialized intrusion detection techniques). From security point of view, a whole communication system should be divided into processing (active) components and transmission media (passive) components. This decomposition differs from the one described in [9], which distinguishes between intermediate nodes, end systems, wires and the total area subnet. We don' t find such division of services area useful - location as described in the above mentioned work is of a little relevance to costs. Services have to be provided end to end. If there are many segments, it makes no sense to apply, *e.g.*, authentication only to end system without having authenticated the rest of the systems along the path.

Services have to be identified exactly along the communication path. It is a handy approach to relate QoSS component to IP address of a device, where a service takes place. Through IP numbers all CPUs, memory elements (RAM, disks,...) can be uniquely identified and the same holds true for transmission media, be it a wire or wireless.

There is certainly a problem of measuring QoSS with regards to availability. Although not mentioned in ISO standards, it became a serious concern in recent years, as availability is mostly degraded by (distributed) denial of service attacks [20]. The reason for these kinds of attacks is increasingly complex software. It is hard to quantify and directly relate software correctness to QoSS taxonomy and pricing. A possible objective metrics for this purpose is to use Common Criteria [24], which address this issue through EAL leveling - the more rigorous tests of equipment, the higher the price. So, denial of service is implicitly addressed through EALs. Finally, it is possible to identify basic cost components and their measures. These are as follows:

- CPU usage, measured in cycles for performing load / store, arithmetic, logical / shift, control and byte manipulation operations;
- Memory allocation, measured in bytes;
- Bandwidth usage measured in bytes per second;
- Software correctness measured in EAL levels;
- Manpower measured in man-years;
- Costs of physical protection of equipment in currency units per year.

CPU and memory usage are obtained using SNMP measurements. Bandwidth usage is obtained through amount of transferred data bits (bytes), divided by the upper time limit for a protocol to finish. Manpower and costs of physical protection are obtained as statistical aggregates during a particular period; they are not calculated repeatedly for each QoSS request.

Table 3: Cost Relation with m-id as a Foreign Key

		cost elements					
ip-id	**m-id**	**cpu**	**mem**	**sw**	**mp**	**phys**	**bw**
ip_1	sa_1	$c_{1,1}$	$c_{1,2}$	...		$c_{1,5}$	$c_{1,6}$
...							
ip_n	pm_h	$c_{n,1}$	$c_{n,2}$	...		$c_{n,5}$	$c_{n,6}$

In the above table *ip-id* presents a unique identifier, an IP address of a device (in case that a device more than one IP number, the one with the lowest value is taken). Further, *cpu* stands for CPU usage, *mem* for memory usage, *bw* for bandwidth usage, *sw* for software correctness, *mp* for manpower and *phys* for costs of physical security. It should be noted that $n \geq k$. As the amount of data should cover arbitrarily large messages, it is necessary to normalize cost elements. Except bandwidth that is computed based on SERVICES relation, and can be treated as a fixed cost for a given protocol, all other elements are thus normalized per byte of payload. To calculate the total cost *C* for a required set of security services, a weighted sum of cost elements $c_{i,j}$ is used and multiplied by the total amount of data *D* in bytes, for which security service is required:

$$C = D * (\sum_{i=1}^{n}(\sum_{j=1}^{5} c_{i,j} * w_{i,j}) + c_{i,6}), \quad c_{i,j}, w_{i,j} \in \mathbb{R}$$

Fig. 1. Calculation of Total Cost

3.3.3 Example

The following example is based on a hybrid authentication and key distribution protocol [36]. *A* and *B* stand for names of communicating parties, *E* for an encryption with a public key, while E^{-1} for decryption, and *K* for a session key:

- Bob sends to Alice his public key.
- Alice's system generates random session key, encrypts it with Bob's public key and sends it to Bob, i.e. $E_B(K)$.
- Bob decrypts by using his private key and recovers session key, i.e. $E_B^{-1}(E_B(K))=K$.

In order for the above protocol to function properly, it is necessary to verify public keys and it is assumed that public key infrastructure is available. Suppose that Bob wants to establish a secure session with Alice at IP address 128.231.3.4 that is on the same network segment as Bob's computer. He decides for strong authentication using RSA with 512 bit long keys.

Table 4: Example of Services Relation

s-id	sec-srvc	s-name	ent-id	t-nonces	cont	cert	CRL	time
at#	authentication	hybrid with key exchange	100	0	16	3000	30000	2
...	...							

Table 5: Example of Mechanism Relation

s-id	m-id	mechanisms	m-name
at_3	aa_4	asymmetric algorithm	512-bit RSA
...	...		

Table 6: Cost Relation Example

		cost elements					
ip-id	m-id	cpu	mem	sw	mp	phys	bw
128.231.3.4	aa_4	20	100	0.005	0.001	0.002	16558
...							

Only the first two steps of the above protocol are relevant for QoSS, as the third step is done by Bob's system. Values in the relation above are obtained as follows: entity identification for Bob and Alice requires 100 Byte, there are no nonces, payload is a session key of 16 Byte, X.509 version 3 certificate has on average 3000 Byte, CRL is expected to be 30,000 Byte. Upper-bound time limit is two seconds. The value for CPU usage is 20 cycles per Byte. Further, memory usage is 100 per Byte, bandwidth usage is 33,116/2 Byte/s. EAL for the appropriate encryption module is 0.005 per Byte, human resources costs are 0.001 man years per Byte, and physical security is 0.002 currency units per Byte. Except for bandwidth usage, all calculations are related to 16 Byte payload processing, which represents secure session key. The absolute value for all weights is taken to be 1, while their units are obtained as reciprocal values of corresponding cost elements. Thus, the total cost *C* for establishment of a secure session is 18,478.128.

4 User-centered Charging

Further aspects of the way users perceive a given service should be investigated, to really keep people as the focal point of service provisioning. With the convergence of voice and data services in IP networks, together with future services and applications, the number of service providers greatly increases. Each end-to-end Internet session or transaction may involve many service providers, each one claiming a kind of payment to regenerate the consumed resources (including investments and revenue). Moreover, users may ask services from application, content, and service providers, Web servers and pages, access networks, and alike. The collection and processing of the charging data from the Internet soon becomes unmanageable for the service providers and the Internet users alike. The number of vendors and service providers operating in the market offer the average user too much choice, without any of the required help and assistance to act as a guide to what is available in the Internet. To manage this, the user is likely to want to deal with only one home ISP, which acts as a broker and mediator between the various service providers involved. Single unified billing for the services and applications consumed in the Internet is one route to improve the global QoS the users and ISPs experience. This is where the Unique-Kiosk Approach (UKA) is positioned.

The UKA bases on the Internet AAA architecture. It provides both the Internet user and service provider with a single Point-of-Charging (PoC) that allows all Internet access charges to be made from one service provider. The UKA is not tied to specific charging models or tariffs that service providers set for the service provision. To capture and process the charging data required by the UKA, an open Charging Platform (CP) architecture is designed. The UKA and CP aim to meet a basic set of requirements for charging and accounting for Internet usage:

- Provide a call detail record for all charges incurred and requiring settlement between the various commercial providers in the loop;
- Allow end users control over the charges being made by the provider;
- Allow itemized billing for all services charged to each subscription, including voice-based and data-containing phone calls, and covering services and content;
- Allow billing consolidation and convergence;
- Provide fraud detection and prevention in the network.

Fig. 2 shows possible charging points that may be used by an ISP.

4.1 Unique Kiosk Approach

Typically, between a user and an ISP there is a communication path belonging to a network provider (NP). Thus, a user should deal, based on separate contracts, with at least two, usually different, providers, and take care of all those contracts. Moreover, in case of Internet link failure, using multiple vendors usually results in finger pointing, delays, and widespread frustration.

To address and prevent these kinds of problems, we now consider the model of the UKA. Although a service provision mechanism, UKA supports and improves the

charging and billing for Internet services. In UKA, a user is to deal with only one provider by means of a two-part Virtual Kiosk (VK). The user part is running on his terminal - computer or mobile device - while the ISP part is running on ISP. The VK belongs to the ISP, which is the parent ISP (PISP) of that VK and a preferred ISP for the user. Relying on the UKA, also an ISP may become “user” for another ISP. To provide services, a PISP should open a VK on the user's terminal. Whenever the user asks a PISP for a service, the appropriate local VK obtains the necessary information from the host user, “finds and hires” the intermediate actors able to connect the user's terminal to the PISP, and further supports the requested service, while maintaining LOG files. A level of QoS is negotiated. The PISP assumes all the costs for the requested services. At the agreed interval, the PISP issues a bill containing all previous collected costs for the services the user requested and consumed. The UKA implies and supports rather than enforces charging/billing/payment policies, thus allowing freedom for the ISPs in this area.

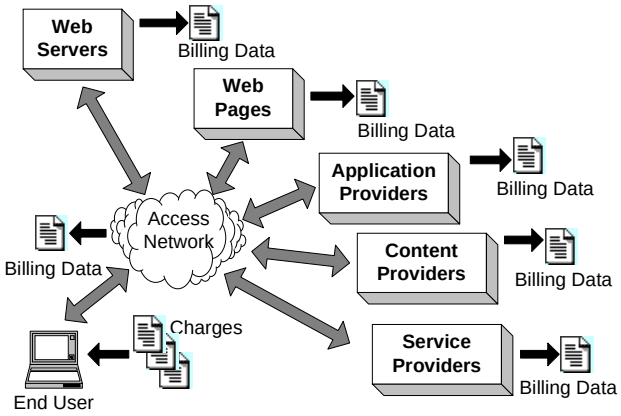

Fig. 2. Charging Points for Internet Usage

The UKA is independent of the access network. The VK knows about the communications interfaces available on the user’s terminal since installation or a later re-configuration, and spawns the appropriate communication agents. Before or during an Internet session, and according to user indications, the VK may choose between NPs based on the required quality of transfer, the physical link availability and quality, tariffs at that time, in addition to other metrics.

At the installation time, the VK asks and assists the user to provide a valid client ID (CID), including *username*, *password*, and *network access identifier*. Also, a VK comes with a Kiosk ID (KID) unique for its PISP. In the two-step log in, the VK checks off-line with the user provided CID and, if it passed, connects to the PISP and sends the pair (CID, KID) for further authentication, authorization, and accounting. Except link problems, more on-line log in attempts may indicate a fraud attempt. Besides CID and KID, a VK encapsulates and hides from other parties the PoPs of its PISP and the strategy to access them (*e.g.*, using metrics like distance and time-of-day) so that the fraud is severely limited.

4.1.1 Points of Charging and Charging and Billing Information Flow

An important component of UKA is the Point-of-Charging (PoC), which is the point where the charging data, bills, or both go either for on-line or latter payment. Although a user is required to specify only one PoC, he has the freedom to specify more permanent or temporary PoCs, instructing the VK how to use them, *e.g.*, a credit card company for local payments in mobile access, the called party, in a VoIP session, or the PISP for Web access. Whenever the user accesses the Internet, the VK presents a PoC as credential indication.

Based on SLAs it has with its neighbor service and network providers, an ISP will be charged by its neighbors for every ISP's user access they support. Each provider, whose network the user traffic passed over, registers a per-flow bill entry and passes the bill backward to the appropriate neighbor provider, at the agreed interval or cycle. This way, the bills accumulate in a scalable manner at the preferred ISP. The ISP then pays to its neighbors the received bills and issues user bills to the indicated PoCs, based on the LOG files maintained by the corresponding VKs.

4.2 Mobility and Roaming

The development of mobile terminals with multiple physical or software-defined interfaces is expected to allow users to seamlessly switch between a wide variety of wired and wireless access technologies, often with overlapping areas of coverage and dramatically different cell sizes. A need has been generated to allow users to negotiate a Point-of-Attachment (PoA) to any domain convenient to their current location [1]. The VK may change the current PoA and interface used to connect to an ISP for a competing NP in its area, based on dynamic VKs.

An (access) NP may host VKs of several ISPs willing to sell their services in that area. The NP sinks/delivers packets from/to users of a hosted ISP and issues a bill to that ISP for the provided service. An ISP may have both active VKs, where their user-part is already running on user terminals, and inactive VKs, where their user-part, hosted by that ISP, is waiting to migrate to a user terminal. For mobility purposes, the user-part of a VK has the capability to host the VK user-part of a foreign ISP. After the user terminal finds a PoA that best meets the user’s needs, the selected ISP becomes temporarily the preferred ISP of the user, and a VK user-part migrates to and runs as guest VK user-part on the user terminal. The guest VK asks the hosting VK or the user for ID data and PoC, performs all required registrations, and supports the service.

4.3 Charging Platform Architecture

The (CP) enables the UKA to be prototyped and evaluated in live and experimental networks. Using a layered approach, the CP is concerned with the collection of traffic data for the purposes of charging, cost allocation, billing, capacity and trend analysis, and auditing. It covers the functionality of the PoCs, meeting the requirements of the AAA model. The main CP’s components are:

- *Multi-User Environment and SQL Database Storage* to process the captured data into billing information, while applying multiple tariff structures and charging models, and as a tool to implement and evaluate charging models and tariffs;
- *Web Server Interface* that uses dynamically produced web pages, and allows to easily navigate, download or forward to additional back-end systems, for further processing or report generation, the CP's functionality, data and information.
- *Scripting Language Support* for rapid prototyping of new ideas and concepts with low development overhead.

Similar architectures have been proposed in [42] and [16]. Suitably extended, the CP carries out also authorization and admission control functionality, based on [1].

4.3.1 Charging Platform

On a first level, the CP acts as a mediation device, as in [27]. The added value comes in terms of the services and applications that can be used for the complex manipulation and modelling of the captured data. The CP has a flexible architecture that offers many benefits and functionality through the on-line and off-line modelling and processing of the captured data. The functionality includes:

- *Charging Data Capture* from many sources, such as: IP level traffic data, http protocol data, GSM voice and data traffic, as well as non-standard application level data. Important elements of the CP are the use of standard data formats, for compatibility with other platforms in use, and the granularity of the data captured, as the level to which network traffic can be metered and charged.
- *Charging for Content* that requires the capture of sufficient data to enable detailed analysis of what content has passed over the network. Instructed by PISP, a VK provides the CP with all relevant data for the user's degree of satisfaction.
- *Charging Models and Tariffs*: Providing a flexible framework for implementation of various charging models, the CP can capture, archive, and process charging data of a wide range of formats. The CP is also designed to model new charging models on archived charging data. Such modelling and analysis will hopefully result in more cost-effective and innovative charging schemes.
- *Charging* for QoS, as a necessary feature of most networks that provide guaranteed service levels, and as a metric that users are willing to pay a premium for. The CP supports differential charging, using suitable charging models.
- *Scalability and Growth:* The CP can run on a single host or on host systems distributed in the network, to reduce the total throughput and capacity requirement on each host, and the traffic carrying the metering information. The number of distributed CPs is likely to increase much more slowly than the number of users.
- *Meeting the Internet Charging Requirements*, as summarized in Table 1.

4.4 Conclusions

Internet access is usually based on services and equipment from different providers. Many vendors usually result in finger pointing and general frustration. To avoid this, the UKA performs on behalf of a preferred ISP all the activities the Internet connection

Table 7: Internet Charging Requirements

Charging Requirements	UKA and Charging Platform Architecture Solution
Call and session records	Produced using the captured and processed network data
User control over charging	UKA allows the user choice and control over the costs via selecting suitable ISP, PoC, and network resources
Itemized billing	Produced using the captured and processed network data
Billing convergence	'One-stop shop' approach of the UKA and the data processing carried out by the charging platform
Fraud detection	Implemented in combination with standard AAA architectures; UKA hides the sensitive data of an ISP from unknown users

requires. This approach gives the user great flexibility in choosing where to procure Internet services from, while keeping a centralized payment arrangement. The use of the CP architecture to capture and process the network data required by the UKA aims at unified Internet charging [11], on metrics over and above the usual access and duration measurements, and simplicity in the charging function.

5 Content Charging

The use of Internet technology for content distribution in a closed or controlled Intranet is an evolving business model for content providers. This allows the use of Internet protocols for the delivery, exchange and distribution of the content and also the implementation of mechanisms for improved or configurable QoS. Charging for the content and the associated delivery is required to enable the various content and network providers involved to generate revenue. Total charge for content is made up of the transport cost plus the cost or charge for the actual content. The cost of transport is made up of the Internet access charges plus access charges when accessing or traversing peering core/access networks plus the charge for the QoS required for the connection. Subscription models are very suited to charging for content delivery from both a consumer and provider viewpoint and may be implemented with current protocols and technology.

5.1 Content Distribution and Exchange

Within the last five years content and its electronic distribution has become more and more important for the communications and media industry. This trend is driven by a number of developments. On the consumer side it is the possibility to receive digital content for instance via the Web or Digital Video Broadcasting (DVB). In the production domain it is the increase of digital platforms and production formats in conjunction with the demand for an ever-increasing amount of quality information that has to be delivered quicker than before.

5.1.1 What is Content?

According to the SMPTE (Society of Motion Picture and Television Engineers) and EBU (European Broadcasting Union) task force definition [38] content consists of essence and metadata. Essence is the raw programme material itself and represents

picture, sound, text, or video. Metadata describes the actual essence and comprises all kinds of data related to the actual essence. Metadata includes content related metadata that describes the actual content; material related metadata that describes the available formats and location related metadata describing the location, number of copies and condition of carrier. Depending on the application area the two different constituents of content are of varying relevance. Metadata is needed to describe and find content. The actual essence is consumed and operated upon.

5.1.2 Content Distribution Requirements

Applications require support for content location, communication and charging. On top of this there might be a number of special services that support for instance the distribution of processes, media analysis and indexing, localized branding. In the context of this section the focus is on communication and charging aspects.

5.1.3 Communications Requirements

Content has by definition multiple parts, viz. metadata and essence. The metadata is conventionally composed of discrete media elements such as text and images, i.e. key-frames. However, metadata segments can have a time relationship, especially when they are referring to segments within a piece of continuous media. The essence of continuous media is clearly time dependent. Very often it is composed of multiple components, *e.g.*, a video track and a number of audio tracks, that have to be combined and displayed at the same time. Content is usually delivered either by streaming or file transfer. A mixture of the two delivery mechanisms that involves caching is also possible. Media streaming is especially in the context of live events required. In order to assure a continuous data flow QoS mechanisms have to be in place. If content is delivered via file transfer very large files have to be handled. These files contain not only essence but also metadata.

5.1.4 Charging Requirements

Only very few goods can be exchanged electronically and content is one of the major products in this area. This implies that there are two chargeable elements in the exchange and delivery of content, viz. the communication costs and the costs for the content itself, i.e. the rights to consume, publish, use. Because of the high bandwidth and QoS requirements the cost of communication is substantial. Content costs depend on the kind of usage. Content distribution or xCast might be charged in a similar way to cable and satellite TV. Video-on-Demand charges are usually modelled similar to video rental prices. Costs for the acquisition of content IPRs (Intellectual Property Rights) highly depend on the kind of content, how many rights holders there are and the kind of rights that will be acquired. There can be a rather complex negotiation process involved. Both parts of content charging can be done entirely separately or charging models can be developed that combine communication and content costs.

5.2 Charging for Content Delivery within CDNs

For charging and billing to take place within a CDN suitable accounting systems need to be in place that can monitor and record network events associated with request-

routing, distribution, and delivery of the digital content. As these network events are normally in the upper OSI layers (4-7) it may be preferable to use accounting protocols and methods that are also hosted in those layers. This would abstract the accounting tasks away from the low level TCP/IP transport of the underlying Internet, and should reduce the amount of accounting data collected for typical content distribution transactions. Charging and billing across a CDN requires infrastructure to authenticate users, create usage records, mediation and aggregation of the usage data, pricing of the usage and content, and consolidation and settlement of charges across various peering networks. Accounting systems must also be able to scale to size, reach and complexity of the peering CDN and not add heavyweight performance overheads to networks.

5.2.1 Service Level Agreements (SLAs) for CDNs

SLAs need to be in place with all negotiated relationships that the CDN operator maintains to be able to offer different levels of QoS to content providers and users alike. A framework for such SLAs has been proposed for Premium IP networks [16], which may also be applied to a CDN overlaid onto the Internet. When SLAs are in place then effective network measurement should also be implemented to ensure that all parties concerned are adhering to the SLAs and charging can be adjusted accordingly. The measurements required should include as a minimum congestion monitoring, failed content requests and end-to-end delays within the CDN.

5.2.2 The Role of MPEG Standards in Billing and Charging for CDNs

There are several relationships that can be identified between the IETF CDI (Content Distribution Internetworking) [21] and ISO/IEC MPEG initiatives. Possibly the most obvious refers to the usage of the MPEG standards for the representation of the coded multimedia content that can be exchanged between, and delivered by, CDNs. The mutual relevance between these initiatives extends much beyond content representation, as exampled by the most recent ISO/IEC MPEG initiative with the Multimedia Framework or MPEG-21 [29], [26]. The MPEG-21 initiative aims at assuring interoperability in the interactions that can occur in the entire multimedia chain (creation/ distribution/usage/consumption) [5]. The seven technical areas that are considered as relevant for this interoperability correspond to the six key core qualifiers that can be applied to user pair interactions plus the reporting of events within these interactions [4], [18]. Both initiatives consider at the same time charging and billing systems as out of their standardization scope and identify accounting (event reporting) as the cornerstone that has to be standardized in order to implement these systems [5]. Furthermore, both initiatives recognize that the definition of a standardized interface and set of metrics are the key components of interoperable accounting systems. CDN accounting system metrics are expressed and conveyed via *Content Detail Records* (CDRs). The payload of a CDR is composed of several fields that qualify in a quantitative/unambiguous manner the activity that is being accounted. The MPEG-21 standard also aims to address different technical requirements that are associated with the representation and distribution of CDRs.

5.2.3 Subscription Charging for Content Delivery

Using subscription services the content delivery provider becomes a broker for the content being distributed and delivered from the content providers. Content delivery/ distribution providers may be able to set up subscription services to the content they host that covers the basic connectivity charges as well as content charging, and act as an ISP. Alternatively a content delivery only subscription service is possible without the added requirements of ISP provision. For subscription charging to work a layered model is required to allow revenue generation to be achieved across a wide customer base. Such layered models have already been successfully implemented in the cable and satellite TV markets, as well as the printed media market for some years. Subscription models have so far proved to be sustainable and successful business models for companies such as B-Sky-B, AOL-Time-Warner, and Kingston Communications. Internet provision is a richer media than cable and satellite TV so the metrics that may be charged for in a subscription are also more diverse. The Internet is capable of carrying and delivering a wide range of services and content and this is reflected in the layered subscription model. QoS provision based on priority scheduling or on bandwidth division or resource reservation is possible based on and dependent upon the level of subscription purchased by the end-user. Increased requested priority results in higher subscription charge for the content being delivered, or possibly varying the quality of the delivered source content, or possibly the scheduling of time-sensitive content, *e.g.*, live/recorded football matches. Other value added content and or meta-data may be chargeable through this model, *e.g.*, MPEG-4 type object content, varying camera angles on streaming video for live sport events etc. A layered subscription model may include the following elements:

- Flat-rate subscription for Internet access, and basic services such as e-mail and web browsing. This may be purchased direct from the ISP or content delivery provider as a package. Un-metered Internet access is becoming the standard business model in this space, and makes sense for customers and providers alike, providing the take up rate is sufficient to cover the overheads.
- Internet access purchased may be based on one or multiple access technologies including dial-in modem, ISDN, xDSL, WLAN, and GSM/GPRS/UMTS.
- Alternatively the Internet access may be purchased from other access providers.
- Subscription packages for valued added services that can be customized to the customers' requirements. Such services may include video, audio, and event content, such as MTV, music/video on demand and news channels.
- Subscription packages for pay-per-view content, such as movies on demand, sporting events and news articles. The emphasis here is on higher value content or better quality content that may be charged for reasonably by the provider.
- Subscription packages for QoS on demand, which allow the customer to have a budget for priority and bandwidth allocation requests over the network for Internet connection and content delivery. This would also allow the customer to purchase more QoS budget if/when required to compensate for expected network usage and utility.

Past research [3], [28] has shown that Internet users like to have predictable charges for their Internet access. Subscription models fulfil this user requirement since the level of spending can be controlled by the level of subscriptions taken out. The layered subscription model also provides the content providers and the network providers with value added revenue streams that subscribers can choose to take up or not. This allows the baseline subscription revenue to be supplemented by the package add-ons and extras. Once subscribers are signed up to the content delivery service converged payments may be used to pay for all the services purchased through one vendor, thereby simplifying the invoicing, charging and payment for both subscriber and provider alike. The charges and subscription rates imposed for the various layers of the subscription model are set by the provider according to their business model, current market trends and the cost of any sourced content and delivery infrastructure.

5.3 Conclusions

Content distribution and exchange is the key to the future development of the Internet. Infrastructure and methods for the auditing, accounting and charging for the interchange are necessary to enable content delivery and exchange to be efficiently charged for and to enable the required revenue generation and collection.

The total charge for the distributed content needs to be a reasonable charge for the content being delivered based on the value of the content and the delivery mechanism employed. Subscription based charging reduces the accounting and billing overhead for the content providers and network providers alike. It also has the advantage of prepayment for content being delivered or consumed. A layered subscription model as proposed provides more choices for the consumer and also the possibility of increased revenue streams for the content and network providers. Both consumers and providers mutually benefit by offering configurable QoS as a user selectable option or parameter in a pricing plan that also has an associated charge. This enables content providers and network providers to generate the revenues required to sustain, maintain and grow the Internet and the services offered.

6 Case Study on Cost Sharing

The case study discussed here is based on the experience acquired in the MIRA project [43]. This project was focused on developing a simple, but effective, charging scheme based on volume and traffic classification for high-speed Internet links. The system bases byte accounting and byte classification on four characteristics of each byte: its origin, destination, application and direction. This charging scheme is more complex than simple byte accounting, and may result in fairer billing and provides additional information about user profiles.

The objective of the traffic classification method for billing is to use heuristics in order to discern user behavior (group of applications, origins, and destinations) from layer 3 and layer 4 header information. Once each byte is accounted for under a class of traffic, the user can be charged for it, taking into account the potential academic use, which includes applications for research, teaching, and development purposes.

The work is oriented towards traffic charging and billing based on network resource usage for private groups (NRNs, corporate networks) that currently offer free Internet access or flat rate access and wish to continue offering services in a fairer way. Network resource usage is characterized by the volume of the traffic and the potential academic profile. The academic profile is derived from traffic classification combinations (group of applications and origins and destinations).

In the particular case of the Spanish NRN (RedIRIS) the topology shows characteristics which made it easy to apply a charging scheme. The topology is a star, with a central node, managed by RedIRIS, and one link for each political administrative region/nationality. Moreover, three additional links connect the RedIRIS Backbone to the Internet. One link connects RedIRIS to the European R&D network Géant. Another link connects RedIRIS to the Spanish Internet Exchange Point (Espanix) where commercial Network Providers are also connected. Finally, a third link connects RedIRIS to the United States for the traffic directed there or for those routes that do not have a path through the Espanix or Géant Networks. Traffic going to the United States may be either American or default-route traffic.

As there are no loops in this topology, monitoring it link by link provides the proper information (no duplications) of the traffic in the Spanish NRN. The regional access points do not inject transit traffic from third-party networks. Consequently, all the traffic present in the regional links belongs to RedIRIS-connected institutions. In order to distinguish the local chargeable institutions, each is assigned a network prefix. To characterize the traffic by what link that traffic uses to gain access to a institution or to go to a destination, external entities are characterized by an Autonomous System group number. Therefore, we can characterize the destination/origin of each traffic flow with a 4x17 matrix, as there are seventeen institutions, one representing each link; and 4 destinations, one for each group of external networks. Each flow detected in a link is also characterized by its application port. Sets of ports join together applications with the same characteristics. Seventeen groups of applications are currently defined.

Finally, a charging matrix can be programmed in order to charge combinations of origin/destinations, applications, and traffic direction. Each origin/destination can be charged differently depending on the potential academic use of that link. Every application group can be charged relative to the potential academic use of applications.

6.1 Traffic Accounting Platform

The basic requirement is that the capture platform must be passive. This requirement allows the parallel development of a traffic capture and analysis system without affecting network performance and administration.

At time the of this study the high-speed transport technology in the Spanish NRN was ATM155. A pair of optical splitters supplied the passive copy of the traffic. Each splitter gives a copy of the incoming and the outgoing traffic of the link under study. The accounting platform can be deployed in the central node, accounting entirely all the regional networks, or it can be deployed at the regional access point. In this case, it can account for smaller institutions. A combination of both configurations was used in order to account for the central node and the regional node for Catalonia separately.

A full-rate traffic accounting platform was deployed using CoralReef flow-based accounting tools, which reports flows instead of packets, thus reducing the amount of information to be analyzed for traffic classification. Now, traffic classes are based on volume, origin, destination, traffic direction, and application (determined from the TCP/UDP ports in the transport header). Later, the academic nature of each traffic flow is set by the traffic classification processes based on some heuristics.

6.2 Traffic Classifications

The IP flows are translated into charging flows by adding class attributes. Subsequently, they are logged under the same flow record, taking into account only the new class attributes. The number of flows is reduced drastically. The source and destination values are reduced to MxN, where M is the number of institutions under consideration, and N is the number of sets of destinations defined for the traffic flows. Also, the number of applications was reduced from the maximum TCP/UDP port number to the number of defined application classes.

These processes are executed once every charging period. The charging period depends on the minimum time for which one wishes to apply a charging scheme. Four class attributes have been selected: the traffic direction, local institution, application group and destination group. Except for the direction class, all class attributes can be configured by adding or removing attributes from configuration files. In this way, the platform is flexible and can be deployed easily in other networks.

Traffic direction: Each traffic flow is classified independently, by taking into account its direction. By maintaining different records for incoming and outgoing traffic, different charging schemes can be applied to each direction of the traffic.

Institution: The second class attribute is the local chargeable institution to which the flow belongs. An identifier represents each institution. There is a database where identifiers are related to network prefixes. As such, institutions having several network prefixes with the same label are charged under the same charging scheme. Also, you can define highly-detailed institutions by defining long network prefixes. The main rule to determining which institution a flow belongs to is by finding which one has the longest prefix match. Each flow is characterized by the institution attribute after analyzing the source IP address (for outgoing traffic) or the destination IP address (for incoming traffic). In the regional node for Catalonia, 40 institutions have been selected for charging. For a centralized charging scheme 17 institutions were defined, one for each regional network. In this case, the IP source or destination addresses are not analyzed, but instead an institution is assigned to each flow based on its VPI/VCI label. As each regional node has a unique VPI/VCI identifier, this method allows for greater speed in the traffic analysis process.

Location: The location class identifies the external origin or destination of the traffic (from the point of view of the Spanish NRN). In order to reduce the number of choices and to relate the information to an academic criterion, four destination attributes are defined. Each attribute represents a set of Autonomous Systems. All Autonomous Systems in a group are connected to a different external link in the core NRN. From a NRN point of view, each set represents a different type of network. The

first group is the Spanish NRN group itself, which represents traffic between Spanish institutions in different regional Networks. The second group is the Géant group, which represents traffic to other European NRNs. The third group is the Espanix group, which represents commercial networks in Spain. The last group is the USA group. This group is the default route for the traffic, and represents most of the traffic going to the United States or going to other networks not connected via Geant or Espanix. In spite of the type or usage of the applications, these destinations have, a priori, more information about the academic nature of the traffic.

Application: The application class attribute is derived from the analysis of the source and destination ports in TCP/UDP headers. Other important transport protocols with no port information, such as ICMP, are also classified. Applications are classified into groups based on the potential academic usage of each application. In this case seventeen application groups are defined.

6.3 Charging and Billing

There are 5,440 different ways of combining the different types of bytes to be charged in the regional node for Catalonia (40 institutions, 2 directions, 4 locations and 17 application groups). Moreover, if we want to apply different charging matrices for different time periods (work week, weekend and day/night), the number of combinations increases. The application class attribute is derived from the analysis of the transport protocol and the source and destination ports in TCP/UDP headers, according to the well-known and registered port numbers, and the non-standard ports used by the most known applications. Other important transport protocols with no port information, such as ICMP, are also classified. Applications are classified into groups based on the potential academic usage of each application. In this case seventeen application groups are defined.

Table 8: Charging Matrix

Location/Application	**E-mail**		**Games**		**P2P**		**WWW**		**Unknown**	
	In	***Out***	***In***	***Out***	***In***	***Out***	***In***	***Out***	***In***	***Out***
RedIRIS	0.1	0.1	3	4	4	5	0.2	0.2	1	1
GÉANT	0.2	0.2	3	4	4	5	0.3	0.3	1	1
ESPANIX	1	1	3	4	4	5	1.5	1.5	1	1
USA (default)	0.5	0.5	3	4	4	5	0.6	0.6	1	1

Table 8 shows an example of a possible traffic charging matrix. Four applications are shown in order to reduce space. Unknown application traffic is charged based only on volume; no variations between locations are applied. Other well-known applications have different charges. For example, the price of e-mail increases from 0.10 per gigabyte for within the Spanish NRN to 1.00 to Spanish commercial networks. Nevertheless, e-mail service may be considered to be academic, and the price rises depending on the link cost, with one exception. Although a local link to a Spanish commercial network may be cheaper than a link to USA, the use of the mail service in this environment may be considered to be less academic. Other applications clearly out of the scope of academic usage are charged heavily. The charges are higher

for outgoing traffic than for incoming traffic, in order to charge more for the service of a non-academic application than for access to that type of service.

This matrix is applied to volume classifications each charging period. Fig. 3 determined the incoming traffic classification volume for one entity over the course of one day. Each application class has four Location columns. Fig. 4 shows the incoming traffic costs once the charging matrix is applied. For the outgoing traffic the results are similar.

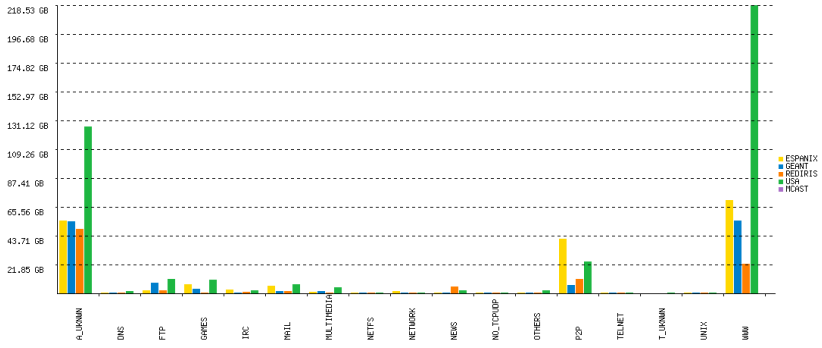

Fig. 3. Incoming Traffic Classification Volume

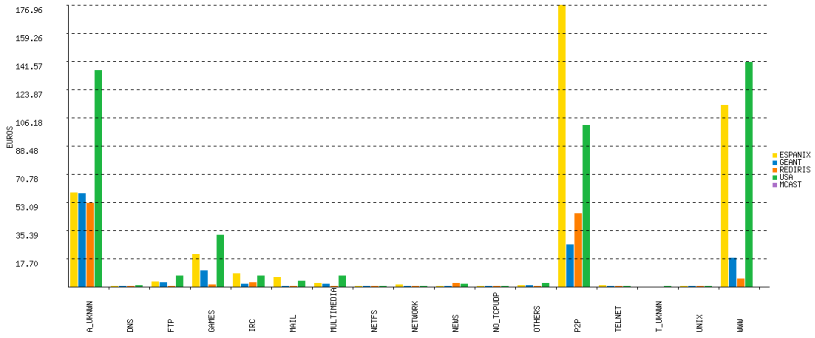

Fig. 4. Incoming Traffic Classification Costs

6.4 Conclusions

Full high-speed traffic analysis in real time with great detail is possible with low-cost resources. The tested traffic accounting and classification platform for cost-sharing in the Spanish NRN was developed exclusively using standard PC equipment. In addition, the traffic-capture tasks do not degrade the network performance, because they are passive. Therefore, the platform can be installed in current networks, and it gives enough information to update some current charging and billing systems for NRNs, since most of them are based only on traffic volume accounting.

Besides the information needed by the billing process, much valuable traffic information is collected during the traffic accounting and classification processes. This information resides in flow registers and is not used for billing, but for gathering

knowledge on institution traffic profiles and usage of network resources. Also, it permits to detect irregular usage or attacks on the studied network and to apply traffic engineering in order to optimize traffic costs.

The transmission speed of network technologies is growing. The backbone network speed increases and the routing/switching process will be faster than real-time analysis. Therefore, full traffic accounting for high-speed links will become unfeasible with low cost resources. To keep costs and resources for analysis low, a sampling method should be applied in the future. A preliminary study on this topic can be found in [44].

7 Overall Conclusions and Outlook

This chapter on pricing and Quality-of-Service in the Internet addressed views of providers and users. While traditional user-provider relationships are defined by consumer and seller roles, more advanced pricing models take into account that neither the provider nor the user dominate the Internet. However, the technology in place, the mechanisms applied, the charging schemes offered, and the QoS delivered to the user define provider-centric tasks. As shown in this chapter, those key components linked into a suitable, applicable, and efficient technology platform, configured by economic goals, and addressing user needs are feasible, though a little complex today to operate.

Therefore, a dedicated and important task set can be outlined, identifying those areas of research, provider support, and user behavior investigations, which demand further work and studies. On the technology side and on the economic part, advances have to be addressed: Appropriate protocols and interfaces have to be defined and standardized to enable the use of cost metrics, which are linked into the communication system to serve user perceived QoS, security, AAA extensions, traffic and accounting analysis, and control algorithms. Research into economically viable charging models and their technically efficient implementations define an important goal of this research, where those control algorithms combine technically measured parameters with business model driven economic targets. A balanced combination of those parameters will allow for a successful user discrimination in the market of Internet Service Provider offering content and various other Internet services.

The authors like to thank J. Diederich for managing parts of the editorial process.

8 References

1. B. Aboba et al.: *Criteria for Evaluating AAA Protocols for Network Access;* RFC 2989, http://www.ietf.org/rfc/rfc2989.txt, November 2000.
2. B. Aboba, D. Mitton (Chairs): *IETF AAA WG;* http://www.ietf.org/html.charters/aaa-charter.html, March 2003.
3. J. Altmann: *How to Charge for Network Services - Flat-Rate or Usage-Based?* Special Issue on Networks and Economics, Computer Networks Journal, August 2001.
4. J. Bormans, K. Hill (eds): *Study on MPEG-21 (Digital Audiovisual Framework) Part 1 v2.0;* Document ISO/IEC JTC1/SC29/WG11/N4040, Singapore, March 2001

5. J. Bormans, Keith Hill (Ed.s): *MPEG-21 Overview;* Document ISO/IEC JTC1/SC29/WG11/N4511, Pattaya, December 2001.
6. B. Briscoe, V. Darlagiannis, O. Heckman, H. Oliver, V. Siris, D. Songhurst, B. Stiller: *A Market Managed Multi-service Internet (M3I);* Computer Communications, Vol 26, No. 4, March 2003, pp 404-414.
7. N. Brownlee: *New Zealand Experiences with Network Traffic Charging;* ConneXions, Vol. 8, No. 12, 1994.
8. CADENUS project: *Creation and Deployment of End-User Services in Premium IP Networks;* http://www.cadenus.org/, March 2003.
9. Cisco Systems: *NetFlow Services Solutions Guide;* http://www.cisco.com/univercd/cc/td/doc/cisintwk/intsolns/netflsol/nfwhite.pdf, 2001.
10. D. Clark: *A Model for Cost Allocation and Pricing in the Internet;* in L. W. McKnight, J. P. Bailey (Eds): Internet Economics; MIT Press, Cambridge, Massachusetts, U.S.A., 1997.
11. J. Cushnie, D. Hutchison, M. Popa: *Internet Charging using the Unique-Kiosk Approach;* COST263 Technical and Management Meeting, Namur, Belgium, http://www.comp.lancs.ac.uk/computing/users/cushniej/jc_papers.html, December 2001.
12. B. Day, K. Hoadley (Consultation Paper): *Network Funding;* http://www.ukerna.ac.uk/archive/response_jisc_399.html, 1999.
13. M. Day, D. Gilletti, P. Rzewski: *Content Internetworking (CDI) Scenarios;* Internet Draft (work in progress), <draft-ietf-cdi-scenarios-00.txt.pdf>, February 2002.
14. M. Day, B. Cain, G. Tomlinson, P. Rzewski: *A Model for Content Internetworking (CDI);* Internet Draft (work in progress), <draft-ietf-cdi-model-01.txt>, February 2002.
15. T. Dolan: *Internet Pricing. Is the end of the World Wide Wait in View?* Communications & Strategies, Vol. 37, 2000, pp 15-46.
16. ETSI - European Organization for Standardization: *NA8 GONOW;* http://www.etsi.org/
17. M. Falkner, M. Devetsikiotis, I. Lambadaris: *An Overview of Pricing Concepts for Broadband IP Networks;* IEEE Communications Surveys, Second Quarter 2000.
18. A. Ganesh, K. Laevens, R. Steinberg: *Congestion Pricing and User Adaptation;* IEEE Infocom, Anchorage, Alaska, U.S.A., April 2001.
19. R. Gibbens, F. Kelly: *Resource Pricing and the Evolution of Congestion Control;* Automatica, Vol. 35, 1999, pp 1969-1985.
20. J. K. Houle, G. M. Weave: *Trends in Denial of Service Attack Technology;* CERT Coordination Center, 2001.
21. IETF: *Content Distribution Internetworking (CDI) Working Group;* http://www.ietf.org/html.charters/cdi-charter.html, March 2003.
22. C. Irvine, T. Levin: *Toward a Taxonomy and Costing Method for Security Services;* 15th Annual Computer Security Applications Conference, Phoenix, December 1999.
23. C. Irvine, T. Levin: *Toward Quality of Security Service in a Resource Management System Benefit Function;* Heterogeneous Computing Workshop, May 2000.
24. ISO: *Common Criteria for Information Technology Security Evaluation - Security Assurance Requirements;* ISO / IEC 15408, Geneva 1999.
25. ISO, Information Technology, Open Systems Interconnection: *Security Frameworks in Open Systems;* IS 10181 - 1 through 7, Geneva 1995.
26. ISO/MPEG Requirements Group: *Current Vision on Event Reporting in MPEG 2;* Document ISO/IEC JTC1/SC29/WG11/N5338, Awaji, December 2002.
27. ITU: *International Telecommunication Union;* http://www.itu.int/, March 2003.

28. S. Jagannathan, K. Almeroth: *Price Issues in Delivering E-Content on Demand;* ACM SIGCOMM Exchanges, May 2002.

29. R. Koenen: *From MPEG-1 to MPEG-21: Creating an Interoperable Multimedia Infrastructure;* Document ISO/IEC JTC1/SC29/WG11/N4518, Pattaya, December 2001.

30. J. MacKie-Mason, H. Varian: *Pricing the Internet;* in B. Kahin, J. Keller (eds): Public Access to the Internet, Prentice Hall, 1995.

31. S. K. Miller: *Facing the Challenge of Wireless Security;* IEEE Computer, July 2001, pp 16-18.

32. MobyDick project: *Mobility and Differentiated Services in a Future IP Network;* http://www.ist-mobydick.org, February 2003.

33. D. Moore, K. Keys, R. Koga, E. Lagache, K. Claffy: *The CoralReef Software Suite as a Tool for System and Network Administrators;* http://www.caida.org/outreach/papers/2001/CoralApps/CoralApps.pdf, 2001.

34. A. M. Odlyzko: *The History of Communications and its Implications for the Internet;* http://www.research.att.com/~amo/history.communications0.ps, 2000.

35. C. Rensing, H. Hasan, M. Karsten, B. Stiller: *AAA: A Survey and Policy-based Architecture and Framework;* IEEE Network Magazine, Vol. 16, No. 6, November/December 2002, pp 22-27.

36. B. Schneier: *Applied Cryptography;* John Wiley, New York, U.S.A., 1996.

37. S. Shenker, D. Clark, D. Estrin, S. Herzog: *Pricing in Computer Networks: Reshaping the Research Agenda;* Telecommunications Policy, Vol. 20, No. 3, 1996, pp 183-201.

38. [7] SMPTE/EBU Task Force for Harmonized Standards for the Exchange of Program Material as Bitstreams: *Final Report: Analyses and Results;* http://www.smpte.org/, 1998.

39. B. Stiller, J. Gerke, P. Reichl, P. Flury: *A Generic and Modular Internet Charging System for the Cumulus Pricing Scheme;* Journal of Network and Systems Management, Vol. 3, No. 9, September 2001, pp 293-325.

40. B. Stiller, J. Gerke, P. Reichl, P. Flury: *Management of Differentiated Service Usage by the Cumulus Pricing Scheme and a Generic Internet Charging System;* 7th IEEE/IFIP Integrated Network Management Symposium (IM 2001), Seattle, Washington, U.S.A., May 2001, pp 93-106.

41. B. Stiller, P. Reichl, S. Leinen: *Pricing and Cost Recovery for Internet Services: Practical Review, Classification, and Application of Relevant Models;* Netnomics - Economic Research and Electronic Networking, Vol 3, No. 2, 2001, pp 149-171.

42. SUSIE Project: *Charging for Premium IP Services;* http://www.teltec.dcu.ie/susie/, 2003.

43. C. Veciana Nogués, J. Domingo Pascual, J. Solé Pareta: *Cost-sharing and Billing in the National Research Networks: the MIRA Approach;* Terena Networking Conference. Limerick, Ireland, June 3-6, 2002.

44. C. Veciana Nogués, A. Cabellos Aparicio, J. Domingo Pascual J. Solé-Pareta: *Verifying IP Meters from Sampled Measurements;* In: Schieferdecker, I.; König, H.; Wolisz (Eds.): IFIP 14th International Conference on Testing Communicating Systems (TestCom 2002). Berlin: Kluwer Academic Publishers, 2002, pp 39-54.

9 Affiliation Data

Burkhard Stiller
Information Systems Laboratory IIS
University of Federal Armed Forces Munich
Werner-Heisenberg-Weg 39
D-85577 Neubiberg
Germany
Phone: +49 89 6004 2826
Fax: +49 89 6004 3898
e-mail: stiller@informatik.unibw-muenchen.de
http://www.informatik.unibw-muenchen.de
and
Computer Engineering and Networks Laboratory TIK
ETH Zürich
Gloriastrasse 35
CH-8092 Zürich
Switzerland
Phone: +41 1 632 7016
Fax: +41 1 632 1035
e-mail: stiller@tik.ee.ethz.ch
http://www.tik.ee.ethz.ch

Pere Barlet-Ros
Carlos Veciana
Josep Solé-Pareta
Jordi Domingo-Pascual
Department Arquitectura de Computadors
Advanced Broadband Communications Laboratory (CCABA)
Universitat Politècnica de Catalunya (UPC)
Jordi Girona, 1-3, Mòdul D6 (Campus Nord)
08034 Barcelona, Catalunya, Spain
e-mail: {pbarlet,carlosv,pareta,jordid}@ac.upc.es
http://www.ccaba.upc.es

John Cushnie Andreas Mauthe Rui Lopes David Hutchison
Computing Department
Lancaster University, UK
SECAMS Building

Lancaster
LA1 4YR
UK
e-mail: {j.cushnie, a.mauthe, r.lopes, d.hutchison}@lancaster.ac.uk
http://www.comp.lancs.ac.uk/computing/

Mihai Popa
R&D Department
SC PROCETEL SA
Calea Rahovei 266-268
Bucharest
Romania
Phone: +40 21 331 7821
Fax: +40 21 423 2609
e-mail: mihpopa@fx.ro

Jim Roberts
France Telecom R&D
DAC/OAT
38 rue du General Leclerc
92794 Issy-Moulineaux
France
e-mail: james.roberts@francetelecom.com
http://perso.rd.francetelecom.fr/roberts/

Denis Trcek
Department of Digital Communications and Networks E6
“Jozef Stefan” Institute
Jamova 39
1101 Ljubljana
Slovenia
Phone: +386 1 477 33 79
Fax: +386 1 426 21 02
e-mail: denis.trcek@ijs.si
http://epos.ijs.si/