Cost-sharing and Billing in the National Research Networks: the MIRA Approach

Carlos Veciana, Josep Solé-Pareta, Jordi Domingo-Pascual

Universitat Politècnica de Catalunya - UPC,
Advanced Broadband Communications Lab. - CCABA[1]
Jordi Girona, 1-3, Mòdul D6 (Campus Nord),
08034 Barcelona, Catalunya, Spain
{carlosv,pareta,jordid}@ac.upc.es

Abstract:

In many countries, National Research Networks (NRN) are funded by public funds. Nowadays, there are many Internet providers and telecommunication operators that can provide the same service than NRN. There should be strong reasons to maintain public funds to the NRN to avoid unfair competence between NRN and commercial Internet services. Since, it is nearly impossible to detect academic traffic among other purpose traffic (commodity traffic), some presumptions upon the traffic should be done. On the other hand, to ban the access to commercial networks and services is not the objective of network administrators. The guideline of current NRN administrators is to promote the academic usage and to allow academic community to use other services too, but under some restrictions. These restrictions consist of charging for the non-academic traffic based on a certain accounting scheme. A possible accounting scheme could be based on classifying the traffic under several classes, and charging per each class based on volume. The revenues of this charging method would be used to maintain a full Internet-connected NRN where academic usage is funded and other usage is paid.

The MIRA project has proposed and tested a cost-sharing method based on the usage of the network resources and it has been applied on the Spanish National Research Network (RedIRIS). A traffic capture platform has been developed in order to capture and analyse the content of the traffic. Moreover, a classification of the potential academic usage of the traffic based on the origin and destination of the IP traffic has been implemented too. In this paper we describe a cost-sharing method, based on two cross-related traffic classifications, which helps to measure the academic and research usage of the network resources. These two classifications can be combined in order to calculate the bill for each entity connected to the NRN. This billing method justifies the public funds for the NRN while the academic and research traffic represents most of the traffic. Also, NRN administrators can use the revenues derived from our billing method for cost sharing between all academic and research community.

[1] Centre de Comunicacions Avançades de Banda Ampla, http://www.ccaba.upc.es

1 Introduction

The Internet has been widely deployed in the commercial arena during last years. National Research Networks are still connected to the Internet but now, they have to coexist with commercial networks.

On the one hand, the objectives of the Academic and Research Networks [1] are different from those of commercial networks. National Research Networks should cope with two basic objectives. The first objective is to provide a testbed for the network research activity; new protocols, new applications and new network technologies can be tested before new proposals are presented. The second objective is to provide scientific and academic communities with a collaborative environment.

On the other hand, the objective of commercial networks is to take economic advantage of services provided to the users (residential, business and even academic users). Since National Research Networks are funded by public institutions, the services provided and the activities that take place in those networks should not overlap with those offered in commercial networks in order to maintain fair competition.

It is very important to characterise the usage the users make of the network resources, in order to verify that Academic and Research activities are taking place in Academic and Research Networks, and economic and leisure activities take place in commercial networks.

Charging and billing is a current topic for most network providers and Internet Service Providers. We propose a charging and billing method based on the usage of the network resources. Two traffic characterisations, one based on the contents of the data interchanged and other based on the origin and destination of the user's connections, are combined with the amount of data interchanged. Then, different charging schemes based on these two classifications can be applied to bill each entity connected to the NRN. Our traffic characterisation tries to detect the academic traffic among the whole traffic in order to charge more the non academic traffic and therefore stimulate the academic usage.

The first objective of the MIRA project was to develop a hardware/software tool that gathers knowledge about the usage that the people make of the Academic Network resources. The second objective is to propose a cost sharing between the entities connected to the National Research Network, based on the usage of the resources they make. The cost sharing proposes a bill per Regional Node and provides audit reports with detailed information about the most important servers and services accessed by the users, and possible irregular usage of network resources.

This paper will show the results of the MIRA project. We describe two traffic characterisation procedures. The usage based classification and the origin and destination based classification. The combination of these two classifications allows us to define a charging and billing scheme useful for cost sharing in the Spanish National Research Network (RedIRIS).

2 Billing and Charging Models Applicable to NRN

The initial design decision about the Internet was called for flat rate pricing in the first days of ARPANET. The intention was to encourage the growth in usage. Nowadays, other charging schemes have been experimented in some networks in order to share costs based on the usage. These charging schemes are usually applied to the traffic of the most expensive links [2].

Janet Model

The JANET, a United Kingdom NRN, charged its members universities by volume, but only for those traffic to the USA link [3]. The charging was applied during high activity hours of the day, trying to reduce the traffic and to redistribute the traffic to the low activity hours. It has had no effect on the traffic growth nor the traffic distribution. But a new charging scheme, charging only transoceanic high cost links, was implemented helping to fund at least the most expensive link.

Switch Model

SWITCH network provides connectivity to universities and research institutions in Switzerland. It recovers its costs from member institutions, but these costs are not distributed inside institutions among its departments. Each institution pays a fixed rate depending just on the connection- one third of the costs-. Moreover, the other two thirds of the costs are charged based on volume. Then, volume drives the most part of the charging. In the SWITCH case, the effects of volume based charging have been noticed in the behaviour of the users. For example, a reduction of WWW traffic has been detected by the use of caching proxy servers. But the growth continues, and sometimes the costs derived from the deploying of such caching servers and the complexity brought to the network does not balance the reduction of costs in the links.

UUNet Model

The UUNet Model [4] does not come from a NRN, but it introduces an interesting characterisation. The UUNet provides a charging scheme based on a characteristic of the traffic, the measure of its bursts. Traffic bursts are measured periodically. Then a statistic measure is applied in order to charge the whole traffic. Notice that simple measures like the amount of bytes are easy to account in edge devices. But other measures about the characteristics of the traffic, like “burstines” or effective bandwidth, are not feasible to be measured for the whole traffic, only in samples.

Although the desire for most users is a simple flat rate for its Internet connections, there are strong arguments for complex pricing. Most times, these arguments are the desire to have a premium service. Other times, the arguments are a most rational usage of network services. A combined charging scheme seems to be feasible to be deployed and accepted [5].

2 A Cost-sharing Model for the National Research Networks

The MIRA approach in the area of charging and billing is based on statistical high-speed traffic capture and data analysis heuristics.

The data analysis heuristics allow us to characterise the information that flows through high-speed links with two new attributes: the origin and destination (O&D) attribute and the usage attribute. These new attributes characterise the traffic in a qualitative way (like the UUNet “burstable” model), in front of other quantitative measures based on bytes or packets transmitted.

The O&D attribute will be one of three values: local traffic (from or to the same NRN, LOC), European research traffic (from or to the Pan-European NRN Ten-155, EUR), national traffic (from or to the Spanish a neutral point Ibernet, SPA), or default traffic (to USA links).

The usage attribute can be one of these four values: academic, leisure, commercial, and unknown. The value of the attribute is assigned after being applied a pattern matching process to the payload of all the packets belonging to the same connection.

We call cost-sharing besides charging because our intention is not to give an accurate bill to the NRN users (academic and research entities) but to provide information based on heuristic classification of the usage. The statistical capture will provide knowledge about the main trends in usage of the network, and the final cost-sharing bill will be based on this main trends.

The MIRA platform is divided into subsystems. The Traffic Capture Subsystem (TCS), and the Traffic Analysis Subsystem (TAS). The TCS is a low-cost and passive traffic capture hardware for IP over ATM links, that performs the packet gathering process. Traffic samples are periodically analysed by the TAS, which groups packets into flows and adds new attributes to these flows. Then, reports about usage are produced in different degree. This process is widely explained in [6].

Once each flow has been characterised with the usage attribute and the O&D attribute, each byte of that flow is charged based on these attributes. The O&D attribute values depend upon the network topology of each NRN. Our example is based on sets of Autonomous Systems, and corresponds to the external connections in the Spanish NRN RedIRIS. Table 1 shows all the possible attribute combinations.

Table 1 Attribute combination

Origin & destination/usage	Academic (ACA)	Leisure (LEI)	Commercial (COM)	Unknown (UNK)
RedIRIS (LOC)	LOC-ACA	LOC-LEI	LOC-COM	LOC-UNK
Ten-155 (EUR)	EUR-ACA	EUR-LEI	EUR-COM	EUR-UNK
Ibernet (SPA)	SPA-ACA	SPA-LEI	SPA-COM	SPA-UNK
USA (USA	USA-ACA	USA-LEI	USA-COM	USA-UNK

There are four classes of traffic for each attribute, resulting sixteen classes of traffic when combined. Taking into account the sense of the traffic, the incoming or outgoing attribute of the flows, results in 32 classes of traffic to be charged.

O&D classes could be seen as distance pricing, but content pricing could be applied with some assumptions. Different values for these attributes can be set based on the costs of the links used or based on the potential academic usage of each link destination.

Usage classification types are based on pattern matching and IP address analysis. The bytes accounted in each class (volume charging) are charged with different fees. The pattern matching tries to find words related to each class of traffic (academic, leisure, commercial). Some symptoms can be detected for each IP connection. Then a verdict based on these symptoms is applied to all the bytes belonging to that connection.

The direction of the flow of the data, from/to the Spanish NRN from/to the external groups of autonomous systems (Ibernet, Ten-155 or USA), is useful to charge more the service of commercial or leisure traffic outside RedIRIS than the access to external commercial and leisure data by RedIRIS users.

The following tables show an example of the tax applied to each class of traffic. Notice that by default, traffic suspicious to be academic is charged under the volume of bytes. Traffic not certain to be academic is charged based on bytes transmitted and finally, traffic more suspicious to be leisure or commercial is taxed heavily.

Table 2 Charging levels

Cost €/GigaByte

0.1	0.2	0.5	1	3	4	5

Table 3 Incoming traffic charging

Incoming Traffic Origin /usage	Academic (ACA)	Leisure (LEI)	Commercial (COM)	Unknown (UNK)
RedIRIS (LOC)	LOC-ACA	LOC-LEI	LOC-COM	LOC-UNK
Ten-155 (EUR)	EUR-ACA	EUR-LEI	EUR-COM	EUR-UNK
Ibernet (SPA)	SPA-ACA	SPA-LEI	SPA-COM	SPA-UNK
USA (USA	USA-ACA	USA-LEI	USA-COM	USA-UNK

Table 4 Outgoing traffic charging

Outgoing traffic Destination/usage	Academic (ACA)	Leisure (LEI)	Commercial (COM)	Unknown (UNK)
RedIRIS (LOC)	LOC-ACA	LOC-LEI	LOC-COM	LOC-UNK
Ten-155 (EUR)	EUR-ACA	EUR-LEI	EUR-COM	EUR-UNK
Ibernet (SPA)	SPA-ACA	SPA-LEI	SPA-COM	SPA-UNK
USA (USA	USA-ACA	USA-LEI	USA-COM	USA-UNK

Finally, as a help to the network administrators to interpret the bill, an audit report about the most important origin and destination locations is given in different detail degree. For the traffic that goes to or comes from Ibernet, volumes per Network are given. For the traffic going to or coming from Ten155 Pan-European network, volumes per AS are given. In this case information about networks is too scattered to be interpreted clearly. The same occurs with the traffic going to or coming from the USA link. The viability of giving per network reports in these two groups of traffic is under study.

4 Field Trial

The MIRA traffic analysis platform has been deployed during last tree years in the Spanish NRN, RedIRIS. Billing information has been added since January 2001. Traffic capture platform is running 24 hours a day, reporting data based on days, weeks, and months. This information can be queried via web interface.

Traffic Capture and Analysis Platform

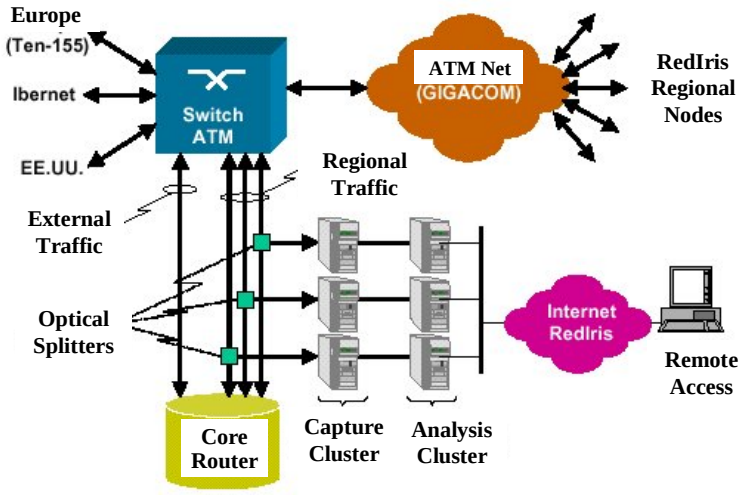

Fig. 1 Mira Platform

Fig. 1 shows the traffic splitter point and hardware related to the MIRA platform. The capture cluster is the traffic capture subsystem (on capture PC for each ATM link, from regional networks to the core router). It is based on a modification of the OC3MON software [7], that runs on low cost PCs equipped with Fore ATM network adapters. The analysis platform collects traffic samples and adds new attributes to the flows. Finally, the characterised flows are classified and charged periodically. The main advantage of this architecture is that it is transparent for the network. The network equipment doesn't need to collect any statistics about the traffic nor SNMP information has to be transmitted trough any network link. The availability of the IP payload allow us to add new attributes, that complement to those derived from the IP headers like other traffic classification platforms/applications [8][9] do. The percentage of the traffic captured depends on the load of the links and the number of capture PC boxes. In [10] you can find a preliminary study about the effect of having samples to estimate some traffic characteristics compared with having the full traffic.

Billing Process

Once all flows are characterised with the new attributes, a charge per byte is applied based on the charging matrix. This charging can be applied to different time aggregations. It is possible to apply a different charging matrix to different time of the day or different day of the week. The billing process is periodically invoked, charging all flows characterised since the last invocation, and reporting a new bill, as shown in Fig. 2.

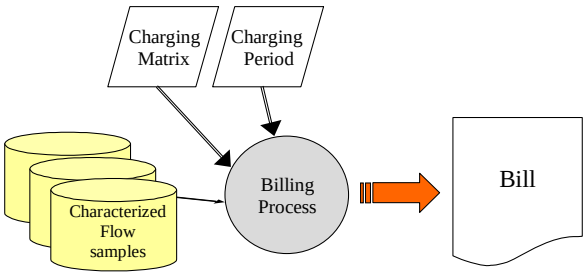

Fig. 2 Billing process

The charging matrix applied is the same explained in section 3. This is the current configuration, and other matrixes that encourage other usage of the network are under study.

The billing process is applied to all the traffic belonging to a set of users. This set of users may be a regional node (all networks connected to the central node of the NRN trough the same ATM link), or an entity (all networks belonging to the same entity, independently of the link of the one that originate the traffic). The examples shown in this paper are based on regional node sets of users.

The following figures show the effect of volume charging in front of usage charging. The results corresponds to one day traffic measurement, April the 6th, one of the regional nodes in Spanish NRN.

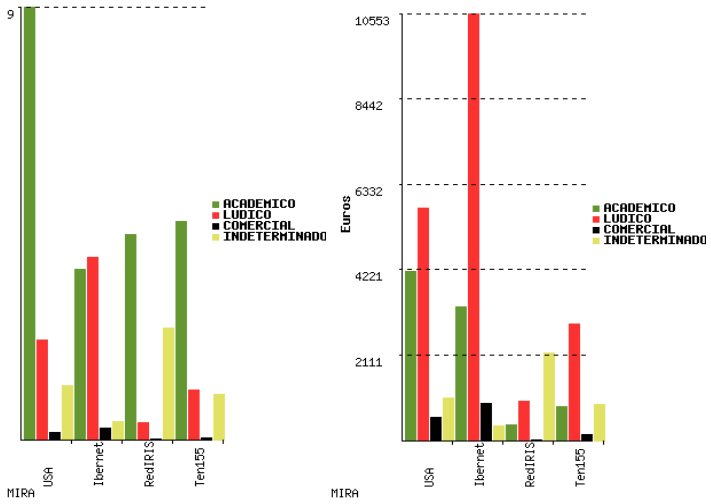

Fig. 3 Traffic classification by usage and origin (GigaBytes), and the bill corresponding to the application of our charging Matrix (Euros).

Fig. 3 shows the traffic classification with two charts. There is the usage and origin classification on the left, in GigaBytes. On the right, you can see the same volumes charged by our charging matrix. One can notice that commercial and leisure traffic has more weight in the final bill, compared to academic traffic. The same happens to USA and Ibernet traffic.

The desired effect of this bill is to reduce traffic to high costs links, deploying local cache/proxy servers and to educate academic and research users in the usage of academic resources.

In order to provide Network administrators with more arguments to carry out this work, some audit reports are provided. These reports have two utilities. On the one hand, they can be useful to detect services, like online newspapers or academic web servers, which can be cached on local servers to reduce the volume. And on the other hand they are useful to justify costs for the RedIRIS entities and commercial Network providers in our country.

Bills per regional node, per entity (set of networks) and per other charging periods (week, month, and working days) can be queried via the web interface developed for the MIRA project. The web interface also allows MIRA administrators to browse the raw data (characterised flow files, networks and AS lists) in order to find more concrete information about network usage.

6 Universities and Internet Applications

Although the first version of the cost sharing scheme was proposed to charge each regional node of the NRN, a new traffic characterisation has been introduced to calculate a bill for each University network. Then, the result of the traffic analysis and billing process for the current version is one bill for each University. With these new classification it is able to apply a different charging matrix for each University, derived from specific characteristics of each University.

The traffic classification based on pattern matching for usage classification may be seen as low confident for some network administrators. For this reason, a simple classification based on application ports, and a later aggregation under groups of Internet applications, is performed now, to substitute the usage classification. Nevertheless, the usage classification is still reported as an audit report.

5 Summary

We present an accounting, charging and billing platform for cost-sharing in the NRN.

It has been designed for high-speed links, where complex accounting is against high performance and simple management.

Our platform gives mechanisms that allow the administrators to define different policies to implement the cost-sharing between the entities that are connected to the NRN.

The billing is mainly based on the destination of the information, but it is also

Our current work is based on statistical validation of our heuristic measures, with some preliminary results to be presented and published in [7]. Also, we are working on a study about the effects of the billing application with long term measures, and the application in Ipv6 networks.

Acknowledgements:

This work has been funded by the CICYT (Comisión Interdepartamental de Ciencia y Tecnología) under contract 2FD97-2234-C03-02 (MIRA). The authors thank the rest of the partners of these projects (Universidad Carlos III de Madrid, Universidad Politécnica de Madrid and Telefónica Investigación y Desarrollo) for their comments and suggestions.

References

[1] Robert J. Aiken., "New Frontiers for Research networks in the 21st Century", The Internet Protocol Journal, Volume 3, Number 4, December 2000. Cisco Systems

[2] A.M. Odlyzko, "The history of communications and its implications for the Internet". http://www.research.att.com/~amo/doc/history.communications0.ps. June 2000

[3] http://bill.ja.net

[4] http://www.uunet.com

[5] Lori M. Allen, "Presentar Factura", translated to Spanish by Mariano Korman. Global Communications, n.43, January 2001. Spanish version of Data Communications, Network Computing and tele.com.

[6] P. Lizcano et al., "MEHARI: System for Analysing the Use of the Internet Services", Computer Networks 31 (1999), pp. 2293-2307.

[7] http://www.caida.org

[8] J.Aspirdof et al. "OC3MON: Flexible, Affordable, High Performance Statistics Collection", INET'97, Lamasya, june 1997.

[9]Cisco Netflow Technology, http://www.cisco.com/warp/public/732/Tech/netflow/netflow_techdoc.shtml

[10] Carlos Veciana-Nogués, Albert Cabellos-Aparicio, Jordi Domingo-Pascual, Josep Solé-Pareta, "Verifying IP Meters from Sampled Measurements", IFIP TC6 Conference on Testing Internet Technologies and Services, TestCom 2002, Berlin, March 18-22. (Proceedings) ISBN 0-7923-7695-1