



Mediation for Convergent Billing Systems

White Paper

Contents

1	Introduction	3
2	Scenario Description	4
3	Solution Description	6
3.1	Define Information Requirements	6
3.2	Identify Points of measurement.....	6
3.3	Identify Available Metrics.....	7
3.4	Map Requirements to Metrics	8
3.4.1	Global Identifiers	8
3.4.2	Timing Issues.....	8
3.5	Implementation.....	9
3.5.1	Framework	9
3.5.2	Capacity Planning and Performance.....	10
3.5.3	Duplication, Loss and Corruption	10
3.5.4	Aggregation.....	10
3.5.5	Rounding.....	10
4	Conclusion	11
5	Abbreviations	12
6	Contact Details.....	13

1 Introduction

The convergence of the Internet, telecommunications, mobile data and content provision have created many opportunities for new products and services and many options for new business models and value chains which can only be validated by market success. Survival in this arena is closely tied to the ability to rapidly introduce new products and ways of charging for them. These developments have presented new challenges to billing systems in particular. These issues have been addressed by modern convergent, event-based billing systems such as Geneva and Portal.

It is true that such billing systems may be relatively easily reconfigured to support new products, services and rating and discounting schemes. It is, however, unrealistic to expect this to involve no more than invoking the appropriate GUI to specify event bindings (the mapping of event attributes associated with new services or devices to new rating or discounting parameters). In fact this is just the beginning of the story. Many other OSS/BSS components are involved and this paper looks at mediation - one of the Cinderella parts of OSS/BSS.

On the face of it the main tasks of the mediation systems are simple:

- collect detailed call and usage information from the network elements;
- format and check these details for syntactical correctness;
- summarise at the level of granularity needed for billing; and
- deliver these details in a common format to the billing system.

However, a crucial and often neglected aspect is that mediation crosses functional areas. It entails all the problems and issues inherent in mapping between the business/service and the network/technical layers of OSS/BSS. Mediation bridges the gap between billing systems that know as little about routers and switches as the latter know about customers, products or subscriptions. It must deal with heterogeneous hardware, multiple network management systems and a variety of OSS/BSS components from different vendors. As such, a successful design and implementation must consider order handling, customer care, provisioning and service inventory in addition to the requirements and constraints of billing systems.

2 Scenario Description

In this section we describe a hypothetical scenario, drawn from our experiences of implementing mediation systems, in order to give some context to the discussion of mediation principles and examples of potential pitfalls described in subsequent sections.

Imagine a service provider that operates a broadband network. One of their key value-add services is IP-VPN allowing any-to-any communication between customer sites with guaranteed maximum access bandwidth and QoS for on-net (within-domain) access. Access to the public Internet is an inclusive, but separately charged, feature which is supported at lower QoS.

The salient tariffing requirements are:

- Customers subscribe to a VPN, including Internet Access, as a single product so that a single administration and management charge may be levied and discounts relating to numbers of sites and traffic volumes within the VPN may be applied;
- There is a monthly charge depending on the access bandwidth available at each site;
- On-net (internal VPN) and off-net (external Internet) traffic has different QoS and hence is subject to different volume-based tariffs;
- Individual dial-up access is also supported;
- Itemised, per-site, billing is required on the basis of traffic volumes in Gbytes per month;
- IP-VPN traffic is charged solely on the basis of incoming traffic to avoid customers feeling they are being charged twice for each byte; and
- Internet traffic is charged separately in both directions to cater both for browsing and for customers who are hosting their own e-commerce sites.

The network is implemented so that for all accesses up to E1, customer sites are connected to an IP Service Platform within the Service Provider network. Accesses at E3 and above have to be made direct to a peering router. Similarly interconnection between the service provider network and the public Internet is also made via the peering routers. This is depicted in Figure 1.

The IP Service Platform supports differential QoS by allowing the configuration of pairs of “bit-buckets” for each customer site virtual connection and provides accounting statistics identified by different class of service identifiers. In our scenario the interface to the IP Service Platform is via a CORBA interface. The peering routers have a separate virtual interface for each customer connection and for each connection to a different carrier. Accounting statistics in our scenario are retrieved separately for each interface via SNMP-V2. (In real life not all routers adhere strictly to SNMP V2, though most “endeavour” to. It is important to thoroughly investigate and document existing or potential problems with software release versions and protocol support for these before embarking on implementation.)

It should be noted that we have chosen a relatively simple scenario where there is only one charging domain. In the world of mobile data several parties, e.g. operator, portal and content provider, may collaborate to provide the service, each incurring different costs. These must be interconnected, mediated and settled!

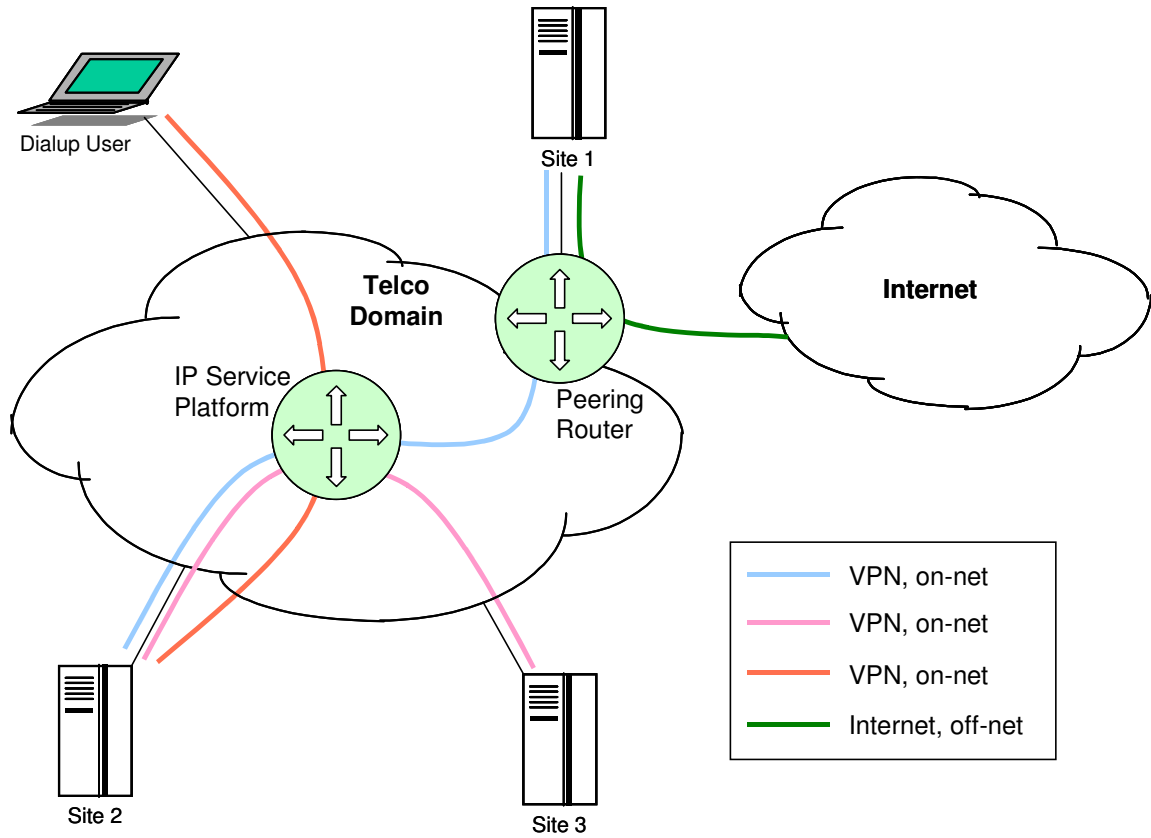


Figure 1. Scenario network and services.

3 Solution Description

3.1 Define Information Requirements

The task here is essentially to build an abstract information model which captures all the requirements for billing in terms of rating parameters, charging attribution, units of aggregation and granularity for invoicing, and to ensure that the relationships are fully specified so that the information can all be pulled together. The starting point is obviously the service definition. However, such definitions are often ambiguous if not completely silent on vital issues. This is often a reflection of the customer's desire to be quick-footed and flexible in responding to changing market conditions and competing products. Although it may seem an obvious pitfall, it is not unusual for products/services to be defined that cannot be mediated!

In the scenario above, the key parameters are as follows.

- Customer identification and bill itemisation
 - Customer ID;
 - VPN Product ID;
 - VPN ID;
 - Site ID;
- Product/service usage
 - Access bandwidth at each site;
 - Charges for each access bandwidth offered;
 - Charge per Mbyte of IP-VPN (QoS1) traffic received;
 - Charge per Mbyte of Internet (QoS2) traffic received;
 - Charge per Mbyte of Internet (QoS2) traffic sent;
 - Mbytes of IP-VPN (QoS1) traffic received at each site;
 - Mbytes of Internet (QoS2) traffic received at each site;
 - Mbytes of Internet (QoS2) traffic sent from each site;
 - What constitutes rateable traffic, e.g. just the IP packets excluding the ATM wrapper;
 - Rounding rules, e.g. are you really going to charge for a full Mbyte if only 1 byte was received in a month?;
- Discount and Rebate
 - Discount structure;
 - Rebate structure;
 - An SLA defining the conditions under which QoS1 and QoS2 are deemed to have been met or, conversely, to invoke rebate.

3.2 Identify Points of measurement

Metrics suitable for inclusion in billing records, or UDRs¹, may be collected from many different devices across the network, but collection should be from as few as possible in order to reduce implementation and maintenance effort, and, critically, keep errors to a minimum. Errors in collection of UDRs will lead to under billing or over billing and, clearly, every effort must be made to avoid this.

¹ We will use the term UDR, or Usage Data Record, throughout this white paper to refer to a dataset recording metrics for one or more units of service or product usage. Of course, a number of acronyms have been coined specifically for one service or another. CDR, or Call Data Record, is the original and was coined for voice records. However, in these days of convergent systems we would prefer to use one acronym to fit all.

It is also important to collect information from the highest possible level, even if it that means that information is not being collected from the device that masters it. For example, in our hypothetical scenario information from a RADIUS server is used to provide the identity of dial-in VPN users. This information could be extracted from the RADIUS server or from the IP Service Platform that interacts with the RADIUS service to obtain that information anyway. So even though the information is mastered by the RADIUS server, the mediation system collects it from the IP Service Platform.

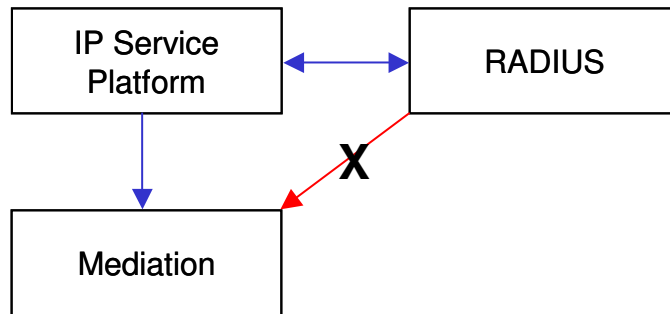


Figure 2. Identification of dial-in users.

So, for our scenario the information requirements to satisfy are:

- Traffic of each QoS type to each customer site connection must be measured;
- Traffic must never be double counted; and
- The fewer measurement points there are the better.

Two points of focus within the network can provide this information:

- the IP Service Platform (one instance of which could serve 32,000 customer connections); and
- the Peering Routers.

3.3 Identify Available Metrics

The metrics available at the IP Service Platform CORBA interface include:

- "User Name" (which can be set to group a series of connections into a single VPN)
- "Circuit ID" (automatically generated as each connection to a customer site is provisioned);
- Domain name;
- Byte counters in each direction for each bit-bucket (related to the two levels of QoS available at each customer connection);
- Start of accounting period;
- End of accounting period; and
- Counters in each direction for each bit-bucket indicating the number of bytes dropped due to exceeding thresholds, congestion, etc.

The metrics available at the Peering Routers SNMP interface include:

- InterfacelD (which relates to a single customer connection)
- Domain name;
- Byte counters in each direction;

- Counters in each direction indicating the number of bytes dropped due to exceeding thresholds, congestion, etc.

3.4 Map Requirements to Metrics

3.4.1 Global Identifiers

Somehow the metrics available from the devices that are being mediated have to be linked to the customers and their services as understood by the billing system. The same link needs to be made to the customer care system, if this is distinct from the billing system. This is especially important for products and services that need to be “hot rated”, i.e. rated in real time, like prepaid services for mobile data.

Traffic volumes must be attributable to appropriate levels in the hierarchy of product, product instance, customer site and customer. In other words a consistent data model is required. In our relatively simple scenario, we need to distinguish VPN from Internet, VPN instance (a customer may operate more than one VPN), volume of traffic per product instance at site level and volume of traffic per product instance at customer level. In principle, therefore, we need Global Identifiers, i.e. unique tags, used by all system elements involved in billing, usage data collection, and customer care, for the customer, the customer site and the product instance. These identifiers must be configured into the system elements when a service is provisioned. If any of these systems do not have the capability to store these identifiers, or cannot store the identifier because it has an incompatible syntax, then mapping tables must be created, and the mapping tables must also be updated when a service is provisioned.

In common with most devices using SNMP MIBs, the peering router in our hypothetical example is focussed on technical network ID's rather than commercial identifiers. A mapping table is therefore required from customer, product and product instance to Interface ID.

Similarly, the IP Service Platform is also a network layer entity whose responsibility is ensuring the uniqueness and separacy of circuits and information at the technical level. Due to the large number of customer connections involved with this key platform, automated provisioning is required and is best achieved in this case by picking up the Circuit IDs generated by the platform and using these as Subscriber ID within the billing system, thus avoiding the maintenance overhead and extra risk implied when using a mapping table. An often overlooked issue is the life time of such identifiers in the different systems. For example, in our scenario the billing system may keep the “Subscriber ID” in perpetuity, whilst the network device may re-use its “Circuit ID” once the original user has been removed.

3.4.2 Timing Issues

Even when all global identifiers and mapping tables have been defined and provisioned, it may still not be possible to directly deduce rateable traffic flows. In our scenario, for example, we need to distinguish between high QoS VPN traffic from low priority surfing and this means combining information from the IP Service Platform and the peering router. Herein lie two problems related to time. In fact, whenever one deals with combination of data from multiple devices the issues of synchronisation and periodicity of collection interval arise.

In our scenario the collection intervals have to be different. The interval is short on the peering router because of the size of counter in comparison with the volume of traffic, whereas it is long on the IP Service Platform because of the amount of post-processing of data records that takes place here. In real life there will also be practical constraints brought by the particular hardware and software implementation of a device, e.g. maximum or minimum collection periods, file sizes, etc.

In order to correctly rate the different traffic types, we must subtract one flow from the other i.e.

internet usage = usage on peering router – usage on IP Service Platform.

For this sum to make sense the usages must relate to the same periods and hence synchronisation is essential.

Such issues of synchronisation and periodicity become particularly important when time of day rating applies.

3.5 Implementation

3.5.1 Framework

Many billing and SLA monitoring systems include a standard framework for mediation, i.e. an architecture, tools and a set of mechanisms which support the collection, filtering, transformation and reliable transport of event data back to the core system. As an alternative, an off-the-shelf system such as XACCT or even the Tivoli Management Framework may be used. Whilst these may seem expensive, it is certainly not worth implementing such a framework from scratch.

In terms of reliability and maintenance a consistent and proven infrastructure is a must-have. (In a similar way standard interfaces should be used wherever possible to collect information.) Another reason for this is that whilst individual switches and even, on occasion, the billing system can be allowed to go down without significant loss of revenue, **if the mediation system is down then event data and its associated revenue is lost forever**. Mediation is thus one context where “five nines” (99.999%) availability really matters. Hence, one typically expects fault tolerant, high availability systems to be implemented.

Lost, duplicate and invalid records can mean lost revenue and lost customers. Hence the framework should support the ACID characteristics of a transaction processing system, where a set of operations which affect data are:

- Atomic - all of them happen or none of them happen;
- Consistent - the data goes from a consistent state to a consistent state;
- Isolated - partial results of the set of operations are not visible outside; and
- Durable - once complete, the changes are not reversed by failure.

In principle the mediation system must retain a copy of all information until the billing system has taken responsibility for it.

Finally, events must be collected from “focus points” across the network to ensure that they are captured once and once only. Hence the transport element of mediation typically involves traversing a WAN. Furthermore, it effectively provides a conduit between the (public) transport network and the NOC (hopefully via a separate management network). Naturally, security comes in to play. For example, access to MIBs is often controlled via a community string password and careful thought must be used on how to manage this. The important thing to remember is that your security policy, as well as providing for encryption and other security mechanisms, must allow the transport of events and UDRs through the firewalls via acceptable protocols.

3.5.2 Capacity Planning and Performance

There is a tension between the general principle to minimise aggregation until the last moment and ensuring that the various counters involved in the mediation and billing systems do not become overloaded, or that all of the processing becomes invested in the billing system itself and thus performance suffers. Another aspect of this is the period over which metrics are gathered. Here technical restrictions, perhaps related to the size of log files, size of counters or the time taken to process log files, may limit the available options. For example, the software may have been implemented on systems that only support 32-bit counters. These will wrap at about 4 thousand million. If they are measuring bytes on a 155 Mbps interface wrapping can occur in about 4 minutes. This means that the mediation system frequently has to deal with wrap. It also gives the mediation system a very small window of opportunity for retry in the event of failure. There are actually three events to distinguish:

- normal wrapping;
- wrapping more than once due to missed polling; and
- counter reset due to hardware or software failure.

3.5.3 Duplication, Loss and Corruption

A mediation system should be design to detect and respond robustly to three types of error within the usage data that it handles. These errors are:

- Duplication of data – to ensure that no record is billed twice;
- Corrupt data – to ensure that only valid records are passed on to the billing system; and
- Missing data – to maximise revenue.

The most appropriate response to the detection of such errors is often manual intervention. The process associated with this must be bullet proof.

3.5.4 Aggregation

Part of the pre-processing function of mediation is to aggregate the usage information into sensible sized units for the billing system. This minimises traffic across the management network and reduces processing load on the billing engine. The limit to aggregation should be governed by lowest level of itemised billing that may be required. Since the goal is to support flexibility and rapidity in the introduction of new services and innovative charging schemes, this should not be compromised by design decisions for the mediation infrastructure, particularly with respect to aggregation.

3.5.5 Rounding

With volume-based charging one inevitably has to round up or down to transform measured units, e.g. bytes, to charged units, e.g. Mbytes. As a matter of principle, to minimise rounding errors one should always do this as late and infrequently as possible. For example, if a product is charged per gigabyte per month but the rounding is applied on a daily basis then a customer could be billed for 30 gigabyte when they have actually used only 30 bytes. This is an extreme example of a common problem because keeping usage statistics in their raw form means that the size of numbers becomes too large to handle.

4 Conclusion

Mediation is the first link in the revenue collection chain. Errors in mediation will lead to erroneous billing and will affect revenue and customer satisfaction. This white paper has shown that whilst the function of a mediation system can be easily described, designing and building a robust yet flexible system is far from easy. Framework products provide an important foundation for the solution, but these are just tools that help to ensure a competent implementation. They provide mechanisms to address complex, distributed, heterogeneous systems. However, the major challenge remains an intellectual one – that of problem analysis and design. A properly thought out design will pay for itself many-fold during the implementation, testing and operational phases.

5 Abbreviations

3G	Third Generation (mobile telephone technology) – UMTS is a member of the 3G family
ACID	Atomic, Consistent, Isolated, Durable – properties of a transaction
ATM	Asynchronous Transfer Mode
CDR	Call Data Record
CORBA	Common Object Request Broker Architecture
E1	2 Mbps access standard
E3	34 Mbps access standard
GUI	Graphical User Interface
IP	Internet Protocol
IPDR	IP Data Record
MIB	Management Information Base
NOC	Network Operations Centre
OSS/BSS	Operational Support Systems / Business Support Systems
QoS	Quality of Service
RADIUS	Remote Authentication Dial In User Service
SLA	Service Level Agreement
SNMP	Simple Network Management Protocol
UDR	Usage Data Record
UMTS	Universal Mobile Telecommunications System
VPN	Virtual Private Network
WAN	Wide Area Network

6 Contact Details

For further information or advice contact *MorganDoyle* Limited. The latest contact details can be found on our website at <http://www.morgandoyle.co.uk/>