Provisioning for Interdomain Quality of Service: the MESCAL Approach

Michael P. Howarth, Paris Flegkas, George Pavlou, Ning Wang, and Panos Trimintzios, University of Surrey

David Griffin and Jonas Griem, University College London

Mohamed Boucadair and Pierrick Morand, France Telecom R&D

Abolghasem (Hamid) Asgari, Thales Research & Technology Ltd.

Panos Georgatsos, Algonet S.A.

ABSTRACT

This article presents an architecture for supporting interdomain QoS across the multi-provider global Internet. While most research to date has focused on supporting QoS within a single administrative domain, mature solutions are not yet available for the provision of QoS across multiple domains administered by different organizations. The architecture described in this article encompasses the full set of functions required in the management (service and resource), control and data planes for the provision of end-to-end QoS-based IP connectivity services. We use the concept of QoS classes and show how these can be cascaded using service level specifications (SLSs) agreed between BGP peer domains to construct a defined end-to-end QoS. We illustrate the architecture by describing a typical operational scenario.

INTRODUCTION

Network quality of service (QoS) is a key consideration for future multiservice networks, as the demands placed on the Internet continue to increase with deployment of multimedia applications and distributed data retrieval systems. Extending the current best effort Internet to support QoS is thus recognized as an important next step [1].

Most research to date has focused on supporting QoS within a single administrative domain. However, delivery of end-to-end QoS to support end-user applications requires that autonomous systems (ASs) administered by different organizations cooperate to deliver the required level of service. The problem we therefore seek to address is how to provide QoS across multiple domains in a way that takes into account the commercial Internet's multi-organizational structure, builds incrementally on existing protocols and approaches, and is scalable.

Compared to the intradomain case, the ability to deliver interdomain QoS requires different IP network providers (INPs) to negotiate service contracts with each other and to engineer their networks to provide the required level of performance. The service contracts are called service level agreements (SLAs), and they include a technical component that is called a service level specification (SLS). The contracts specify the relationship between an INP and either its customers or a peer INP. A key aspect of the SLS is definition of the QoS classes an INP can offer for its customers' traffic. Each INP then has to provision and configure its network resources so that traffic is forwarded in accordance with the agreed QoS levels. Thus traffic engineering (TE) plays an important role in achieving end-to-end QoS, enabling the network to deliver defined performance (measured typically in terms of throughput, delay, and packet loss) while also optimizing the use of network resources.

The ability to support interdomain QoS provides a number of challenges: awareness of QoS capabilities in other domains (for both long-term planning and provisioning, and short-term dynamic response); the ability to engineer domains to deliver QoS; and to achieve all this in a way that is scalable. In this article we propose a framework that brings together all the functions described above in a way that meets these challenges. The framework therefore encompasses business-related processing of service planning and exchange of QoS capabilities between providers; QoS-based inter- and intra-domain TE in the management plane; QoS-enabled routing at the control plane; and traffic enforcement in the physical network at the data plane. Our approach does not require that automated processes always be used to implement the functions described here; many of the management functions could be implemented by manual processes, or by manual processes with automated support.

Some work on interdomain QoS provisioning exists in the literature. Key components of an interdomain QoS architecture have been described in [2]; at the service management level, [3] proposed SLA policies to enable INPs to agree how to distribute QoS across multiple domains. At the control level, QoS extensions to the underlying border gateway protocol (BGP)

 0163-6804/05/$20.00 © 2005 IEEE

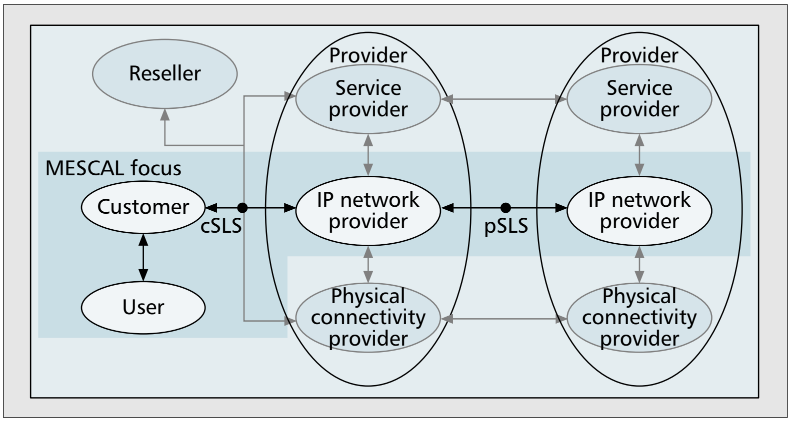

Figure 1. *The business model and MESCAL focus.*

have also been proposed [4, 5]. Among other authors, at the resource management level [6] has described interdomain TE heuristics to perform outbound path selection. The work presented here is, however, we believe the first to provide a full description of the functionality required to fully support interprovider QoS including service and resource management, control, and data plane levels.

The rest of this article is organized as follows. We review our assumed business model and define the principal actors in the Internet. We define key concepts and entities that provide a vocabulary for describing and implementing QoS between INP domains. We then describe our proposed architecture that supports interdomain QoS across the multiprovider commercial Internet, based on work done in the EU project Management of End-to-End Quality of Service Across the Internet at Large (MESCAL). Finally, we describe how the components of the architecture are used in a typical system scenario, and we introduce a set of service scenarios to show how the architecture can be implemented to support different end-user QoS requirements.

The MESCAL Business QoS Model

The business model assumed in MESCAL, as shown in Fig. 1, is for the purpose of illustrating how our work relates to current business practices. This model depicts the stakeholders, capturing their business roles and relationships in the chain of IP QoS-based service delivery. The entities in this model are described below.

A *customer* can subscribe to QoS-based services offered by *providers*. Customers are the target recipients of QoS-based services offered on the basis of respective SLA agreements. They interact with providers for the purpose of buying services to meet their communication needs. A *user* is an entity, either a human being or in general a computer process, that uses the QoS-based services bought by the customer.

Providers are responsible for offering and provisioning QoS-based services. Depending on the type of services offered, three types of providers are distinguished: *service providers*, *INPs*, and *physical connectivity providers*. Service providers offer higher-level QoS-based services encompassing both connectivity and informational aspects (e.g., telephony or content streaming services). They may not necessarily own or administer an IP network, but they need to administer the necessary infrastructure required for provisioning the offered services. Service providers may rely on the connectivity services offered by INPs.

INPs offer QoS-based IP connectivity services (i.e., services that provide reachability between hosts in the IP address space with particular QoS parameters). These INPs must own and administer an IP network infrastructure. For connecting customers to their IP infrastructure, INPs may interact with access providers, or customers could be connected through facilities provided by the INPs. For the purpose of expanding the geographical span of the offered connectivity services, INPs can interact with each other on a one-to-one peering relationship basis.

Physical connectivity providers offer physical connectivity services (i.e., up to the link layer). Access providers offer services for connecting customer premises equipment to an INP or service provider's equipment.

The focus of the MESCAL project is the business relationships between customers and INPs, and between INPs, for the purpose of realizing QoS-based IP connectivity services. The business relationship of primary interest to MESCAL is represented by SLAs and more specifically by their technical aspects, the SLSs. In the rest of this article the term *provider* is used to mean an INP unless otherwise specified.

The MESCAL Internet QoS Model

In this section we describe the principal notions and entities required for the framework, and describe the relationships between them. We extend the intradomain QoS model devised and validated in the TEQUILA project [7, 8] so that the MESCAL model can cover QoS-based services potentially spanning the entire Internet.

QoS-Based Services

The term *QoS-based service* denotes a service that offers QoS-based added value to customers (e.g., matching their usage requirements). Services are offered on the basis of SLAs. The latter are established between customers and providers, and describe the characteristics of the service and their mutual responsibilities for using and providing the offered services. The SLS is an integral part of an SLA, denoting the technical characteristics of the service such as bandwidth, delay, and topological scope. Two types of SLSs are identified in MESCAL, extending previous work on intradomain SLSs [9]:

- Customer SLS (cSLS), established between end customers and INPs
- Peer SLS (pSLS), established between INPs with the purpose of expanding the geographical span of their offered QoS services

QoS Classes

A QoS class (QC) denotes a basic network-wide *QoS transfer capability* of a single provider's domain. A QoS transfer capability is a set of attribute-value pairs, where the attributes express various packet transfer performance parameters such as one-way transit delay, packet loss, and interpacket delay variation (jitter), and their particular values. A provider domain's supported QCs are divided into local QoS classes (l-QCs) and extended QoS classes (e-QCs), to allow us to capture the notion of QoS capabilities across domains:

- *l-QC* denotes a QoS transfer capability provided entirely within the local provider domain itself.
- *e-QC* denotes a QoS transfer capability provided using both the local domain and other (service-peering) domains. An e-QC is provided by combining an l-QC with appropriate l-QCs or e-QCs of other domains. The topological scope of an e-QC therefore usually extends outside the boundaries of the local domain.

From a service *offering* perspective, QoS classes correspond to the performance (transfer quality) guarantees expressed in c/pSLSs. From a service *provisioning* perspective, QoS classes segregate the network QoS space into a number of distinct classes, and hence set the traffic-related objectives of TE functions. The concept of l-QC could be compared to the differentiated services (DiffServ) per domain behaviors (PDBs).

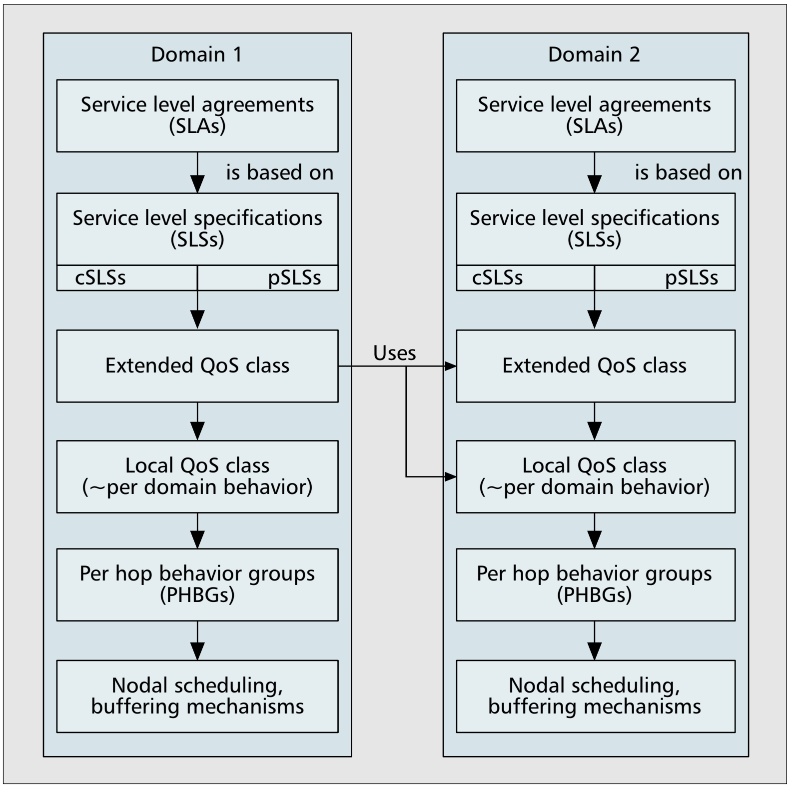

Figure 2. *The MESCAL Internet QoS service model.*

The Internet QoS Model

The MESCAL model for Internet QoS-based services is shown in Fig. 2. It is layered, built around the notion of QoS classes introduced earlier. For a single provider domain, QoS classes abstract the network element QoS-enabling capabilities into sets of network-wide packet transfer capabilities. This provides the necessary abstraction level for building QoS-based services and linking service-peering provider domains so that they can expand the geographical scope of their QoS-based services independently of the underlying network-level capabilities and technologies employed in the different provider domains.

Interdomain QoS Peering Models

The term *peering* is used throughout this article to denote two providers interacting for the purpose of expanding the topological scope of their offered QoS-based services with financial settlements;[1] peering here implies the existence of some form of customer-provider relationship [10]. There exist many models for interconnection and service-layer interactions between providers for offering QoS services across domains. These models rely mainly on experience in the telecommunications industry in provisioning international telephony services. We build on these models to establish a set of pSLSs and thus to construct end-to-end QoS-based services across the Internet. Conceptually, there are a number of peering models such as hub, source-based, cascaded, and hybrid. In the source-based model, an INP negotiates pSLSs directly with downstream providers to construct end-to-end QoS service. With this model, service peers are not necessarily physically adjacent.

Providers would prefer to offer services that reflect the current loosely coupled Internet structure and for whom the use of a *cascaded* model would be more appropriate: this also has advantages of improved scalability and allowing incremental deployment. Therefore, the MESCAL solution adopts a hop-by-hop cascaded model for interactions between providers at both the service and network layers. In the cascaded model, each INP makes pSLS contracts with the immediately adjacent interconnected INPs. Thus, the QoS peering agreements are between adjacent neighbors, and not between providers more than one hop away. This type of peering agreement is used to provide QoS connectivity from a customer to reachable destinations that may be several domains away. Figure 3 gives an overview of the operations in this model. l-QC3, l-QC2, and l-QC1 are supported by AS3, AS2, and AS1, respectively. AS2 negotiates a contract (pSLS2) with AS3, enabling AS2's customers to reach destinations in AS3 with an e-QC2. Although not shown in the simple example of Fig. 3, in general there may be many options for combining a domain's local QoS capabilities with those of adjacent providers. We use the term *QC mapping* to mean the process of identifying this set of options. We then use the term *QC binding* to describe the process of selecting which of the possible QC mappings are put into effect (e.g., in the case of Fig. 3, associating AS2's internal l-QC2 with the external l-QC3 offered by AS3). QC binding might result in a number of QC bindings for a given e-QC, for example, using different peers.

This binding process can be cascaded to further domains. Thus, AS1 can negotiate with AS2 to enable AS1 customers to also reach destinations in AS3, although at no point do AS1 and AS3 negotiate directly.

Each INP needs to know the e-QCs supported

[1] This definition of peering is more generic than the one used today, which assumes peering between providers does not include any financial settlements (i.e. there is no customer-provider relationship).

A provider domain's supported QCs are divided into local QoS classes (l-QCs) and extended QoS classes (e-QCs), to allow us to capture the notion of QoS capabilities across domains.

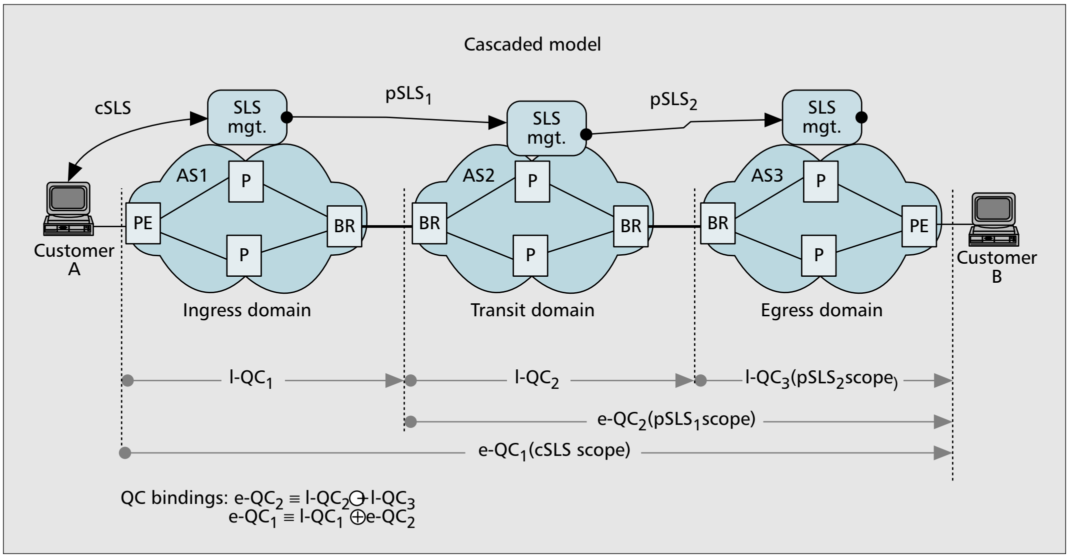

Figure 3. *A cascaded QoS peering model.*

by its neighboring domains for binding with its own l-QCs in order to construct its own e-QCs, which it then advertises to its customers and peers. Since pSLSs are established for aggregated traffic demands, each INP typically only has to manage a limited number of pSLSs. Thus, the number of pSLSs that needs to be established by an INP is only on the order of the number of the adjacent domains, making the cascaded model scalable. A limitation of the cascaded approach is that it gives the pSLS service initiator less control of the whole IP service path.

Functional Architecture

This section introduces, from the perspective of a single provider, the functions required for the provision of interdomain QoS services. The architecture decomposes the functions required to provide interdomain QoS: this enables the development of interdomain QoS solutions by breaking the overall problem down into manageable entities while maintaining a holistic view of the problem. Figure 4 shows the components of the functional architecture grouped into their major functional areas.

The functional architecture is divided into three planes: management, control, and data.

The management plane includes offline functionality, typically located in management servers outside the network elements. Relevant functions are responsible for:

- Interacting with customers and service peers to negotiate contracts
- Implementing the business decisions of the INP through planning, dimensioning, and configuring the network

The control plane covers intra- and interdomain routing, handling the admission of traffic flows, and dynamic resource management including load distribution and capacity management functions. Typically, control plane functions are embedded within network elements, although they are not involved in packet-by-packet decisions.

Finally, the data plane is responsible for per-packet treatment, and is configured by the control plane.

The management plane functions run at the epochs of the so-called resource provisioning cycles (RPCs). In MESCAL we define two RPCs: the intradomain RPC, for offline intradomain TE, and the interdomain RPC for offline interdomain TE. At RPCs network resources are optimized to meet predicted demand, including sufficient spare capacity to avoid network reconfiguration at each SLS subscription or renegotiation, while avoiding the inefficiencies of massively overprovisioned resources.

We now discuss the principal functional groupings within the architecture.[2]

Service Planning and QoS Capabilities Exchange

QoS-based service planning encompasses the business-related activities responsible for defining the services a provider offers. These are specified according to the business objectives, and include l-QCs within the provider's own network, and e-QCs that result from combining its local QoS-based services with those offered by adjacent peers. This in turn requires that a domain be aware of the QoS class capabilities of other domains.

Prior to any pSLS agreement with a peer, a provider has to *discover* from potential service peer providers the peer's QoS capabilities to various destination prefixes and their associated costs. This is achieved using the *QoS capabilities discovery* function. Once l-QCs and e-QCs have been defined and engineered within a domain (by intra- and/or interdomain TE), the *QoS capabilities advertisement* function is responsible for promoting the offered services so that customers and service peer providers are aware of the offerings. It is envisaged that a variety of advertising means could be used, ranging from digital marketplaces or other automated peer-to-peer processes to conventional offline techniques.

Network Planning and Provisioning

Network planning includes the offline processes responsible for determining the type, quantity, and geographical location of the physical

[2] *For most providers, an important aspect of service differentiation is a means of charging for different service levels. Metering, rating, billing, and other commercial aspects of QoS delivery are outside the scope of this article, but are discussed in [11].*

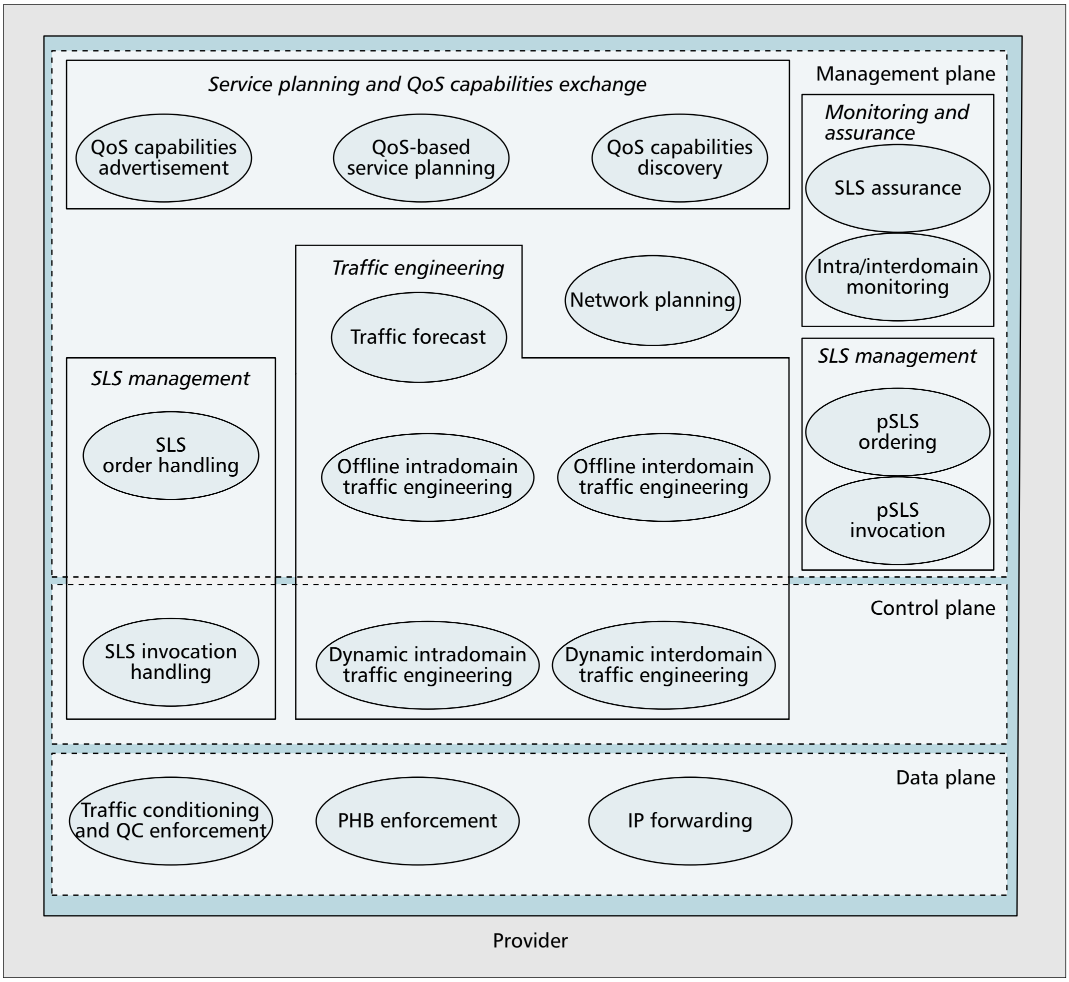

Figure 4. *Functional architecture for interdomain QoS delivery.*

The number of pSLSs that needs to be established by an INP is only of the order of the number of the adjacent domains, making the cascaded model scalable. A limitation of the cascaded approach is that it gives the pSLS service initiator less control of the whole IP service path.

resources (e.g., points of presence, IP routers, and communications links) required by an INP. It also encompasses network provisioning to ensure that the physical resources are deployed as planned, with the appropriate physical configuration. This is distinct from TE, which is responsible for managing the distribution of traffic, optimizing the use of *existing* physical resources, and ensuring QoS in a cost-effective manner. While many management activities can be achieved in an automated manner through network configuration, the implementation of planning decisions usually involves manual installation or configuration of physical equipment.

Offline Traffic Engineering

Traffic forecast is responsible for aggregating and forecasting traffic demand. During an RPC, the set of subscribed cSLSs and pSLSs is retrieved from *SLS order handling* and an aggregation process derives a traffic matrix between ingress and egress points of the domain. This is then used to calculate and provision the intra- and interdomain resources needed to accommodate the traffic from both established SLSs and those anticipated to be ordered during the provisioning cycle.

Traffic engineering is divided into inter- and intradomain functions. Although we consider them as separate blocks, it is important to recognize that an optimal TE solution for end-to-end QoS requires the two to work together closely. For example, an interdomain TE solution that assigns certain traffic flows to certain interdomain links but results in some intradomain links being overloaded is not a good solution.

Offline interdomain TE performs the QC mapping and QC binding operations described earlier to construct potential e-QCs that meet the service requirements defined by *QoS-based service planning*. It then works with *offline intradomain TE* to select a subset of these e-QCs while making optimal use of intra and interdomain network resources (this is QC binding). It also identifies a set of optimum pSLSs that need to be established with downstream providers.

Offline intradomain TE computes the intradomain network configuration (routing constraints and capacity requirements per QC) that satisfies the predicted traffic demand.

Dynamic Traffic Engineering

Dynamic interdomain TE runs within an interdomain RPC and is responsible for interdomain routing. An example of how this would be implemented is a QoS-enhanced version of BGP [5]. Dynamic interdomain TE also dynamically performs load balancing between the multiple paths defined by *offline interdomain TE*. It uses real-time monitoring information, changing appropriately the ratio of the traffic mapped to the interdomain paths.

We have analyzed three potential end-to-end service options, each of which could be supported by a particular customization of algorithms within the architecture proposed earlier. We call each such configuration a solution option.

Dynamic intradomain TE is the dynamic management layer defined in [7]. It includes intradomain routing, load balancing, and dynamic bandwidth assignment for managing in real time the resources allocated by *offline intradomain TE,* in order to react to statistical traffic fluctuations and special arising conditions. It controls the network resources, and is responsible for controlling the routing processes dynamically and ensuring that the bandwidth is appropriately distributed among the traffic classes or DiffServ per hop behaviors (PHBs).

SLS Management

This includes two distinct phases: ordering (i.e., establishing contracts between peers) and invocation (i.e., committing resources before traffic can be admitted).

For ordering, *SLS order handling* implements the server side of the SLS negotiation process. Its purpose is to perform subscription-level admission control. It receives from the offline intradomain TE block the resource availability matrix (RAM), which indicates the available capacity of the engineered network to accept new SLS orders from local customers (cSLS) and peer domains (pSLS). SLS order handling maps incoming SLS requests onto the e-QCs it can offer and investigates whether there is sufficient intra- and interdomain capacity based on the RAM for that e-QC. *pSLS ordering* is the client side of the pSLS negotiation process: it receives requests from offline interdomain TE for new pSLSs, and negotiates them with service peers (i.e., by communicating with SLS order handling in the peer domains).

Requests for invocation of pSLSs are handled by *pSLS invocation*. Admission control is needed to ensure that the network is not overwhelmed with traffic; this allows the network to adopt a policy of overbooking resources at the subscription level. *SLS invocation handling*, the server side of the invocation process, contains the admission control algorithm, and receives requests from customers or peer providers for cSLS/pSLS invocations. It checks whether the invocation conforms to the subscribed SLS and if there is sufficient capacity in the local AS (and also on the interdomain pSLS in the case of SLSs not terminated locally).

Data Plane Functions

Traffic conditioning and QC enforcement is responsible for packet classification, policing, traffic shaping, and marking according to the conditions laid out in previously agreed SLSs and the invocation of those SLSs. At ingress routers *traffic conditioning* classifies incoming packets to their e-QC and marks them with the appropriate DiffServ code point (DSCP) for the required l-QC. At the egress router the *QC enforcement* function may need to remark outgoing packets with the correct DSCP as agreed in the pSLS with the service peer. Thus, QC enforcement implements the data plane binding from l-QC to e-QC.

PHB enforcement represents the queuing and scheduling mechanisms required to realize the different PHBs with the appropriate configuration.

Monitoring and SLA Assurance

Monitoring is responsible for node and network monitoring, collecting data at the request of other functional blocks and notifying them when thresholds are crossed on both elementary data and derived statistics. *SLS assurance* compares the monitored performance and traffic statistics to the contracted QoS levels agreed in the SLSs to confirm that the network or service peer networks are delivering the agreed service levels.

System Scenarios

Illustrating the Architecture

We now illustrate the functional architecture by describing a working scenario when a new interdomain QoS-based service is required. The numbers in parentheses refer to the numbered interactions in Fig. 5, which extends Fig. 4 by showing the high-level interactions between functional blocks. Figure 5 also depicts some of the functional blocks in upstream customer and downstream service peers to show the interaction of the functional blocks with those of the neighbors. The arrows depict the direction of the main flow of information, generally implying a configuration or the invocation of a method in the direction of the arrow.

QoS-based service planning (1) identifies a new interdomain service that could be offered to customers, say, for viewing high-quality streamed video from a set of servers located in remote INP domains. Business planning will specify the technical parameters of the e-QC (bandwidth, delay, etc.) that could be formed from combinations of its existing l-QCs and e-QCs already offered by its peers (known from the QoS capabilities discovery block) to the remote destinations. Part of this function will also determine the expected demand from customers and the cost constraints, including the price it is prepared to pay peers for the e-QCs. The e-QC QoS parameters, required destinations, and cost constraints are passed to offline interdomain TE (2) to trigger a new interdomain RPC. The anticipated demand is passed to traffic forecast to generate a new traffic matrix for this RPC (2).

Offline interdomain TE algorithms are invoked (3) to discover suitable bindings of l-QCs and e-QCs. Appropriate peer INPs and their available e-QCs are identified via the QoS capabilities discovery block (4). After selecting feasible l-QC/e-QC bindings, it will decide on the most suitable bindings (and the bandwidths of the required pSLSs) that meet all of the traffic demands specified in the traffic matrix (5) that includes the new service.

While interdomain TE optimizes interdomain resources (QC bindings and peer pSLSs), it is necessary to ensure that:

- There are sufficient intradomain resources (l-QC capacity) between the anticipated customers and selected egress routers
- The intradomain configuration to meet the selected interdomain bindings is not suboptimal

An iterative algorithm therefore runs between offline inter- and intradomain TE (6), with intradomain TE receiving the intradomain traffic matrix from traffic forecast (7).

As a result of this iteration, offline interdo-

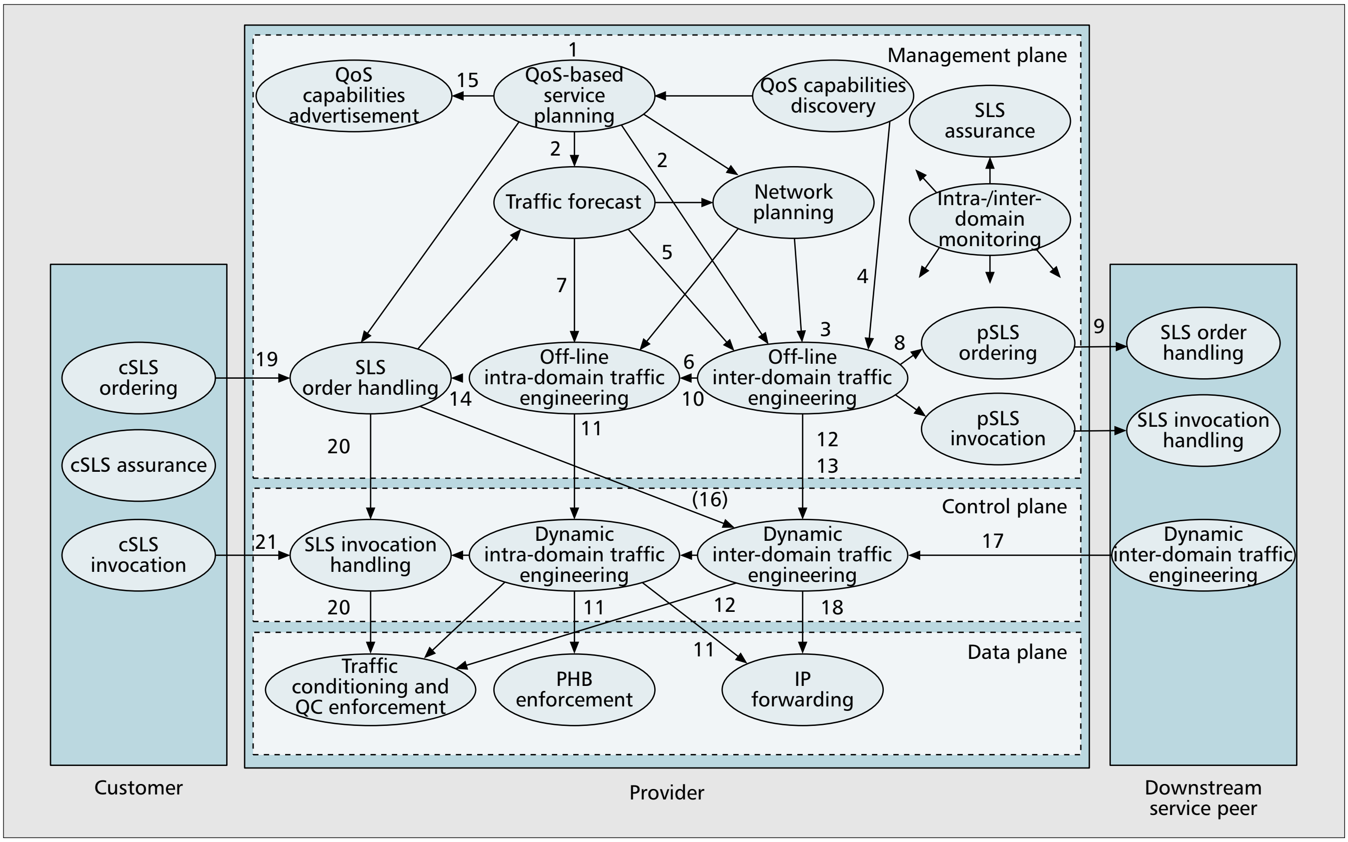

Figure 5. *System scenario.*

main TE selects candidate solutions to be negotiated with peer domains by pSLS ordering (8). pSLS ordering negotiates these pSLSs with the relevant peer INPs (9).

Once the pSLSs have been agreed, offline interdomain TE triggers offline intradomain TE (10) in order to configure the selected intradomain solution. Intradomain TE will configure network parameters (e.g., routing plans and PHB capacities) and deploy these in the routers via dynamic intradomain TE (11). One approach for dynamic interdomain TE is to use a QoS-enhanced BGP (q-BGP). If such an approach is used, offline interdomain TE also configures egress routers with the correct DSCP mappings for the selected l-QC/e-QC bindings (12), and configures the q-BGP processes in the dynamic interdomain TE blocks (13) with appropriate policies for processing the q-BGP messages that will arrive from the downstream peer ASs where new pSLSs have been established.

Offline interdomain and intradomain TE will also forward to SLS order handling the inter- and intradomain RAM for the chosen configuration (14). These will allow SLS order handling to determine whether there is capacity for future c/pSLS subscriptions from customers or upstream peer INPs. The new e-QC capabilities are advertised to upstream INPs and potential customers via the QoS capabilities advertisement function (15).

In the downstream peer INPs, once a new pSLS has been agreed, SLS order handling will configure, for example, the q-BGP processes (16) to forward q-BGP announcements to the new customer ASs for the destinations and e-QCs in the new pSLS. q-BGP announcements will subsequently be received from the downstream ASs (17). The dynamic TE processes will select appropriate interdomain routes, enforcing the policies previously configured by offline interdomain TE (18). From this point on the INP is able to forward packets to remote destinations with the required QoS, assuming of course that the INP's customers (end customers as well as upstream INPs) first establish and invoke SLSs to use these capabilities.

A customer wishing to subscribe to the new interdomain service will initiate a cSLS or pSLS negotiation with SLS order handling (19). The latter will consult the RAM and the repository of existing SLSs to determine whether there is sufficient capacity for the request. Once the SLS has been agreed, the traffic conditioners in the ingress routers will be configured for the new SLS (20). In the case of an end customer, when a policy of SLS overbooking is deployed in the INP, each micro-flow that is part of the overall pSLS subscription will signal its requirements (21) via the SLS invocation handling component in the ingress routers, where admission control algorithms will determine whether there is sufficient capacity to avoid QoS deterioration. The extent to which admission control is required depends on how hard or soft a QoS guarantee is required.

APPLYING THE ARCHITECTURE: SOLUTION OPTIONS

The architecture described above provides a framework for all the components required to implement interprovider QoS, allowing coordination between neighboring domains, to provide end-to-

Service options	Usage	Characteristics			
		Topological scope	E2E QoS performance	E2E bandwidth	Routing mechanism
Loose	Improved Internet service for large population size	Any reachable destination	Qualitative	No guarantee	IP or MPLS
Statistical	Statistically bounded QoS for specified destinations[1]	Specific destinations	Qualitative or quantitative	Statistical guarantee	IP or MPLS
Hard	Hard guarantees based on paths/tunnels for corporate customers	Specific destinations	Quantitative	Guaranteed	MPLS

[1] A range of customers can be identified that require QoS performance guarantees between the two extreme cases, for example ,hard upper bounds on delay to a large but limited set of destinations with statistically guaranteed throughput.

Table 1. *MESCAL service options.*

end QoS through a cascaded model. The model is generic, allowing a variety of different performance guarantees to be provided. For example, residential customers may need to subscribe to QoS-based IP services in order to get to any reachable destination at any time simply with better-than-best-effort service levels. On the other hand, corporate customers may require hard upper bounds on QoS parameters and a constant bandwidth for supporting particular mission-critical services such as IP VPNs to a limited set of destinations. In order to satisfy a wide range of QoS requirements, and therefore potential customers, we have analyzed three potential end-to-end service options (Table 1), each of which could be supported by a particular customization of algorithms within the architecture proposed earlier. We call each such configuration a *solution option*.

The three solution options that correspond to the three service options of Table 1 are as follows:

- The **loose guarantees** solution option extends our QC definitions, using what we call a *meta QoS class* (m-QC) [12]: this is an abstract class, based on qualitative metrics. m-QCs are useful for defining Internet-wide QoS parameters that are understood by all QoS service providers. It is envisaged that providers throughout the Internet would implement a small number of well-known m-QCs. Interdomain QoS services are then created by constructing paths across those domains that support a particular m-QC. m-QC examples are a voice traffic m-QC with low delay, or a low-packet-loss m-QC. Thus, traffic can be sent across multiple domains using dedicated m-QCs with specific QoS performance constraints. Because m-QCs are globally agreed, the QC binding function simplifies to the tasks of mapping a domain's l-QCs to the closest corresponding m-QC. The end result can be considered a series of parallel Internets, each supporting a different m-QC. The m-QC service is provided to any reachable destination.
- The **statistical guarantees** solution differs from the loose guarantees solution by providing end-to-end guarantees associated with specific destination prefixes and defined by the strict cascaded approach of Fig. 3. The QoS characteristics and bandwidth provided to any destination prefix are thus more tightly specified than in loose guarantees.
- The **hard guarantees** solution option provides end-to-end guarantees by reserving resources through the construction of explicit interdomain multiprotocol label switching (MPLS) QoS-based label switched path (LSP) tunnels. The paths would be engineered by coordination of a number of path computation servers (PCSs), one located in each domain [13].

Summary and Future Work

In this article we have addressed the issue of how to engineer the Internet to support QoS across multiple domains. We have defined a QoS vocabulary, defining l-QCs that describe QoS transfer capabilities within a provider domain and e-QCs as QoS transfer capabilities constructed by a combination of these l-QCs and offered across multiple domains. We have distinguished between cSLSs and pSLSs. We have applied our vocabulary to a cascaded interdomain peering model. We have presented an architecture that defines the functional blocks an INP needs to deploy in order to support interprovider QoS, and described an operational scenario that illustrates how the components of the architecture interwork. Our approach shows how adjacent INPs negotiate pSLSs with each other and engineer their network based on predicted traffic. A QoS-enhanced BGP can also be used to support dynamic interdomain TE. We have introduced three service options, each of which is supported by our architecture, as an illustration of the type of guarantees users and applications may require.

We have completed the detailed design stage in which the proposed functional blocks of the architecture have been specified in terms of interfaces to other blocks, behavior, and algorithms. We are currently at the stage of implementation, validation, and experimentation, from which some preliminary results of our proposed interdomain TE algorithms have already been published [14]. We will continue experimenting and validating the system through both testbed environments using Linux-based routers and simulators in order to be able to deal with large-scale networks, stress conditions, failures, and so on. Detailed evaluation results on various aspects of the proposed framework will appear in future research papers.

We are also investigating how BGP [15] could be extended to convey QoS-related information between peer ASs and have proposed a QoS-inferred BGP (q-BGP) protocol that extends work on the QoS_NLRI attribute described in [5]. This allows domains to exchange at the routing level parameters such as QoS service capabilities, and QoS performance and traffic characteristics. q-BGP is applicable to any interdomain QoS delivery solution requiring exchange of QoS information and especially to all three of our solution options.

We are additionally investigating the role of admission control mechanisms in the architecture. We are also in the process of specifying PCS functions, including a communication protocol and a mechanism for remote PCS discovery. The framework for using PCS elements to provide the Hard Guarantees solution option is described in [13].

Acknowledgment

This work was undertaken in the Information Society Technologies (IST) MESCAL project, which is partially funded by the Commission of the European Union. We would also like to thank the rest of our MESCAL colleagues who have implicitly contributed to the ideas presented here.

References

[1] G. Huston, "Next Steps for the IP QoS Architecture," IETF Informational RFC 2990, Nov. 2000.
[2] P. Fuzesi *et al.*, "Provisioning of QoS Enabled Interdomain Services," *Comp. Commun.*, vol. 26, June 2003, pp. 1070–82.
[3] P. Pongpaibool and H. S. Kim, "Providing End-to-End Service Level Agreements Across Multiple ISP Networks," *Comp. Net.*, vol. 46, 2004, pp. 3–18.
[4] L. Xiao *et al.*, "QoS Extension to BGP," *IEEE Proc. ICNP 2002*, Paris, France, 2002.
[5] G. Cristallo and C. Jacquenet, "The BGP QoS_NLRI Attribute," IETF Internet draft, draft-jacquenet-bgp-qos-00.txt, Feb. 2004, work in progress.
[6] T. C. Bressoud *et al.*, "Optimal Configuration for BGP Route Selection," *Proc. IEEE INFOCOM 2003*.
[7] P. Trimintzios *et al.*, "A Management and Control Architecture for Providing IP Differentiated Services in MPLS-Based Networks," *IEEE Commun. Mag.*, vol. 39, no. 5, May 2001, pp. 80–88.
[8] E. Mykoniati *et al.*, "Admission Control for Providing QoS in IP DiffServ Networks: The TEQUILA Approach," *IEEE Commun. Mag.*, vol. 41, no. 1, Jan. 2003, pp. 38–44.
[9] D. Goderis *et al.*, "Service Level Specifications Semantics, Parameters, and Negotiation Requirements," IETF Internet draft, draft-tequila-sls-02.txt, Feb. 2002; http://www.ist-tequila.org
[10] G. Huston, "Interconnection, Peering and Settlements," *Internet Protocol J.*, vol. 2, no. 1, Mar. 1999, pp. 2–16.
[11] P. Georgatsos *et al.*, "Provider-Level Service Agreements for Interdomain QoS delivery," *Proc. 4th Int'l. Wksp. Advanced Internet Charging and QoS Tech.*, Sept. 2004.
[12] P. Levis *et al.*, "The Meta-QoS-Class Concept: a Step Towards Global QoS Interdomain Services," *Proc. IEEE SoftCOM*, Oct. 2004.
[13] M. Boucadair and P. Morand, "A Solution for Providing Inter-AS MPLS-Based QoS Tunnels," IETF Internet draft, draft-boucadair-pce-interas-00.txt, Oct. 2004, work in progress.
[14] K-H. Ho *et al.*, "On Egress Router Selection for Interdomain Traffic with Bandwidth Guarantees," *Proc. IEEE High Perf. Switching and Routing Conf.*, Phoenix, AZ, Apr. 2004.
[15] Y. Rekhter and T. Li, "A Border Gateway Protocol 4 (BGP-4)," IETF RFC 1771, Mar. 1995.

Biographies

MICHAEL HOWARTH (m.howarth@surrey.ac.uk) is a lecturer in networking at the Centre for Communication Systems Research (CCSR), University of Surrey, United Kingdom. He holds a Bachelor's degree in engineering science and a D.Phil. in electrical engineering, both from Oxford University, and an M.Sc. in telecommunications from the University of Surrey.

PARIS FLEGKAS (p.flegkas@surrey.ac.uk) is a research fellow at CCSR, University of Surrey. He holds a Diploma in electrical and computer engineering from Aristotle University, Thessaloniki, Greece, and an M.Sc. in telematics from the University of Surrey, and is currently working toward his Ph.D.

GEORGE PAVLOU (g.pavlou@surrey.ac.uk) is professor of communication and information systems at CCSR, University of Surrey, where he leads the activities of the Networks Research Group. He holds a Diploma in engineering from the National Technical University of Athens, Greece, and M.Sc. and Ph.D. degrees from University College London (UCL).

NING WANG (n.wang@surrey.ac.uk) is a research fellow at CCSR, University of Surrey. He holds a B.Eng. in computer science from Changchun University of Science and Technology, P.R. China, an M.Eng. in electrical and computer engineering from Nanyang Technological University, Singapore, and a Ph.D. in electronic engineering from the University of Surrey.

PANOS TRIMINTZIOS (p.trimintzios@surrey.ac.uk) was a research fellow at CCSR, University of Surrey, and now works as a researcher at the Foundation for Research and Technology — Hellas (FORTH), Institute of Computer Science, Crete, Greece. He received a B.Sc. in computer science and an M.Sc. in computer networks both from the University of Crete, Greece, and his Ph.D. from the University of Surrey.

DAVID GRIFFIN (dgriffin@ee.ucl.ac.uk) is a senior research fellow in the Department of Electronic and Electrical Engineering, UCL. He has a B.Sc. in electrical engineering from Loughborough University, United Kingdom, and is currently completing a part-time Ph.D. in electrical engineering from the University of London. Before joining UCL he was a systems design engineer at GEC-Plessey Telecommunications, United Kingdom, and then a researcher at FORTH, Institute of Computer Science, Crete, Greece.

JONAS GRIEM (jgriem@ee.ucl.ac.uk) is a postgraduate student in the Department of Electronic and Electrical Engineering, UCL. He has an M.Eng. in electronic engineering with computer science from UCL, and is currently completing his Ph.D. in which he is studying IP-based traffic engineering for intradomain networks.

MOHAMED BOUCADAIR (mohamed.boucadair@francetelecom.com) was an R&D engineer with France Telecom R&D in charge of dynamic provisioning, QoS, multicast, and intra/interdomain traffic engineering, and now works on VoIP services. He graduated from the Ecole Nationale Supérieure d'Ingénieur de Caen (Institut des Sciences de la Matière et des Rayonnements), a French school of engineers.

PIERRICK MORAND (pierrick.morand@francetelecom.com) is the project coordinator of the EU IST-MESCAL project. He also leads a research group within France Telecom working to develop and enhance VoIP services for corporate business. He graduated from the Ecole Nationale Supérieure d'Ingénieur de Caen (Institut des Sciences de la Matière et des Rayonnements), a French school of engineers, and received his degree from the University of Caen, France.

ABOLGHASEM ASGARI (hamid.asgari@thalesgroup.com) is an assistant chief engineer with Thales Research and Technology, United Kingdom, specializing in data communication systems/networks. He holds a B.Sc. from Beheshti University, Tehran, Iran, and an M.Sc. from the University of Auckland, New Zealand, both in computer science, and a Ph.D. in the design and analysis of ATM networks from the University of Wales, Swansea, United Kingdom.

PANOS GEORGATSOS (pgeorgat@egreta.gr) leads the Telecommunications R&D Group within Algonet S.A., Greece. He received a B.Sc. degree in mathematics from the National University of Athens, Greece, and a Ph.D. in computer science with specialization in data network routing and performance analysis through queuing theory techniques from Bradford University, United Kingdom.

We are investigating the role of admission control mechanisms in the architecture. We are also in the process of specifying PCS functions, including a communication protocol and a mechanism for remote PCS discovery.