Strategies for Provisioning End-to-End QoS-based Services in a Multi-Domain Scenario

Carlos Alberto Kamienski and Djamel Sadok

Centro de Informática, Universidade Federal de Pernambuco
Caixa Postal 7851, Cidade Universitária, Recife/PE, 50732-970, Brazil
E-mail: {cak, jamel}@cin.ufpe.br

Providing end-to-end performance guarantees for QoS-based services, such as interactive voice and video, is a challenging task in the current Internet. In most cases, it involves the cooperation of multiple administrative domains, for the correct resource provisioning along the end-to-end path. Although some QoS standards have emerged in the last years, together with a trend for covering Internet services with performance guarantees through a SLA, end-to-end QoS provisioning remains presently an open research problem. The main reason, advocated by this paper, is the current domain interconnection structure, which does not provide enough financial incentives for QoS deployment. In this paper, we tackle the interconnection problem by giving special attention to interdomain dynamic service negotiation, which is an important step that must be considered prior to the physical resource provisioning. Three strategies for ISP interaction (called negotiation models) are presented and evaluated through simulation and analysis.

1. INTRODUCTION

Deploying QoS-based services in the Internet is a research concern that has been around for some years. Yet, currently there is no general solution for end-to-end QoS, i.e., when source and destination applications are located in different network domains. Some router-level solutions (such as IntServ and DiffServ) have been standardized and implemented in commercial products, enabling routers to offer tailored treatment for traffic from different users and applications. Furthermore, ISPs have been offering Internet services with some performance guarantees, using SLAs (Service Level Agreements). Although most SLAs only cover a single backbone network, recently some ISPs have been offering multi-provider SLAs, that is, SLAs that extend their guarantees to other carefully chosen domains. These attempts have not been completely successful, because they do not deal with the heart of the problem, which is, in our opinion, the interconnection structure of domains in the current Internet.

We argue in this work that deploying QoS-based services is not a matter that can be resolved only by implementing mechanisms for resource reservation or service differentiation at the router level. In the last few years, there has been a trend to deal with the QoS problem at a higher abstraction layer, in the form of an overlay network. Some commercial solutions that have been using the overlay technique for offering QoS-based services over the Internet are Equinix [4] and Internap [8]. A number of research projects have also identified this problem and consequently have focused on building SLA-based overlay networks for QoS provisioning, such as, CADENUS [1], AQUILA [3], EURESCOM P1008 [6] and Internet2/QBone [14], TEQUILA [15].

None of them, however, has the explicit concern of separating service and operation related issues in distinct planes and in proposing and evaluating service negotiation models as

we have done in our Chameleon Architecture [9]. Chameleon's approach focuses on services rather than on QoS technologies, by placing all service-related aspects in a common service plane. The Service Broker (SB) is the entity in charge of the activities of this plane. SON (Service Overlay Network) [2] is a proposal that shares some similarities with Chameleon, but it sees the overlay network as a separate provider that buys capacity from underlying transport domains. The OverQoS architecture [13] follows the same idea. In Chameleon, the combination of the service planes from multiple domains builds a shared overlay network.

We state that the strategies for provisioning end-to-end QoS-based services in the Internet must observe four primary aspects, which are important pieces for solving the QoS puzzle and have been extensively used in the design of the Chameleon Architecture:

1. Technology independence: domains are not expected at all to use the same QoS underlying technology for providing the agreed performance guarantees. Currently, IntServ, DiffServ, MPLS and even over-provisioning have been proposed as QoS technologies.
2. Service definition: Domains need to decide which services are worth to be deployed and how they will agree on the service definition. Service definition in Chameleon is done by a Well-Defined Service (WDS) [9], which allows every domain to have the same understanding of the service requirements, so that a unique end-to-end behaviour may be achieved. A WDS is the technical part of the SLA, comprised of one or more SLSs (Service Level Specification), which contain the performance parameters concerning QoS levels. WDSs are transport services and different user services can be mapped to the same WDS, thereby reducing the complexity of QoS management and control.
3. Interconnection of domains: New business relationships and settlement structures are needed for the interconnection of domains in the Internet, in order to provide domains with the right financial incentives for QoS deployment. The combination of the settlement structures of the two most common forms of interconnection, transit and peering, currently creates a situation where only half of the end-to-end path of a packet is covered by financial compensations [7]. Such a scheme does not stimulate ISPs to provision resources for enabling end-to-end services in a multi-domain scenario.
4. Interdomain service negotiation: As traffic volumes vary over different timescales for different services, domains will probably want to renegotiate the service levels of the SLA with each other in a more dynamic way.

In this paper, we focus on the organizational aspects of the interconnection of domains, giving special attention to strategies (or models) for interdomain service negotiation and their comparative evaluation in order to favor end-to-end QoS provisioning. Three service negotiation models are considered: the Cascade, Hub and Hierarchical ones. They are evaluated according to the selected criteria of efficiency and scalability.

The rest of the paper is structured as follows. Section 2 presents more details of interdomain service negotiation and the Cascade, Hub and Hierarchical models are described. Configurations used in the simulation study are described in section 3. Sections 4 and 5 present evaluation results. Finally, section 6 draws some conclusions and presents suggestions for future work.

2. INTERDOMAIN SERVICE NEGOTIATION

Interdomain service negotiation refers to the process whereby domains communicate with each other in order to deploy services in the Internet. It involves verifying the feasibility of deploying a given service through some sequence of domains in terms of its required performance guarantees. Resource provisioning for satisfying end-to-end QoS requirements is a step with follows service negotiation. In the current Internet, this negotiation is performed in

a static manner, only for the best effort service. Domains maintain several bilateral agreements, and the timescale for renegotiation typically is in the terms of months. With the introduction of new QoS-based services, this situation tends to become more critical, due to the more frequent changes of service utilization patterns and the need to finding new routes for meeting services' performance requirements. Therefore, *dynamic* service negotiation will be necessary, and should be based upon efficient, scalable and financially viable negotiation models.

Domains may play the role of service buyers and/or sellers in the negotiation process. Most concepts related to service negotiation refer to resource (throughput): estimation, request, offering, allocation and grant. Resource estimation: buyer domains determine the amount of resources needed in the end-to-end path, through traffic measurement and prediction, and perform the resource request. Typically, resource requests will be triggered at equally spaced time intervals (e.g., at the time scale of hours). Resource offering: seller domains prepare a service offering matrix, containing the implemented services and their corresponding QoS levels. Resource allocation: process of looking for resources for current requests. Resource grant: a resource grant (total or partial) implies that resources will be provisioned all the way along the end-to-end path.

The service negotiation model determines the particular mode whereby a group of domains interacts in order to achieve dynamic service negotiations.

2.1. Cascade Negotiation Model

In the Cascade negotiation model, each domain is responsible for the connectivity and communication with its immediately adjacent domains. An end-to-end service is built by concatenating the specific bilateral agreements between each pair of domains. In this model, a buyer domain willing to deploy a service that ends in another one, negotiates with its neighbour, which in turn negotiates with its next one, "rippling" messages through until the destination domain and then backwards (Figure 1).
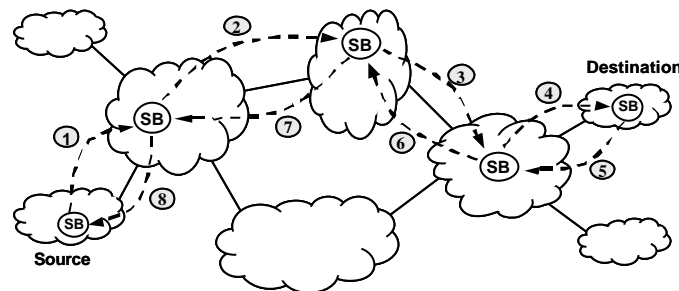

Figure 1 – Cascade Negotiation Model

The cascade model (and also the hub one) has been proposed in EURESCOM projects [6] as an organizational model for the interconnection of operators. It was also proposed for Bandwidth Brokers (BB) [11] and now it has been adapted to service negotiation.

2.2. Hub Negotiation Model

In the Hub negotiation model (Figure 2), a domain playing the role of a User Service Provider (USP) is responsible for negotiating with every other domain along the path from source to destination. In other words, the USP represents the buyer domain.

The hub model requires two stages, although Figure 2 only represents one of them. The first stage is for pre-negotiation, whereby the USP sends messages to all domains involved in

the end-to-end service deployment one at a time, checking for service availability and pre-allocating resources. The second stage is for confirmation of negotiation results.
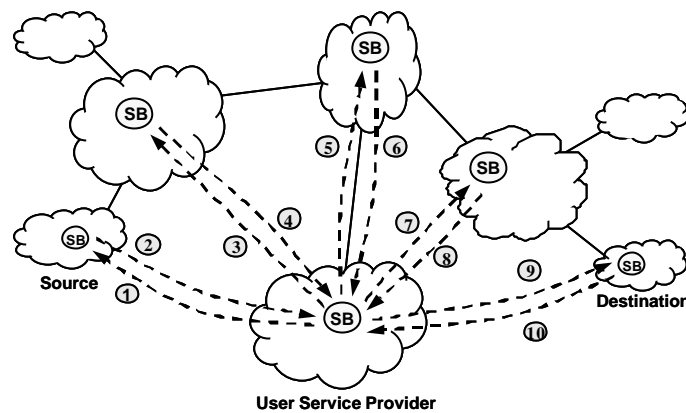

Figure 2 – Hub Negotiation Model

2.3. Hierarchical Negotiation Model

The Hierarchical model [9] (Figure 3) introduces the new concept of Service Exchange (SE), which is a central entity that coordinates service negotiation among a group of participant domains. A SE needs to keep some information for performing negotiations, such as offered services, topology, inter-domain link capacity. Furthermore, each domain periodically sends service purchase and sale information, and the SE performs negotiation "rounds", which result in service permission or refusal. The function of the SE may be played by a commercial entity that offers services for domain interaction or by one of the participants (or an outsourced entity) in a more collaborative environment.
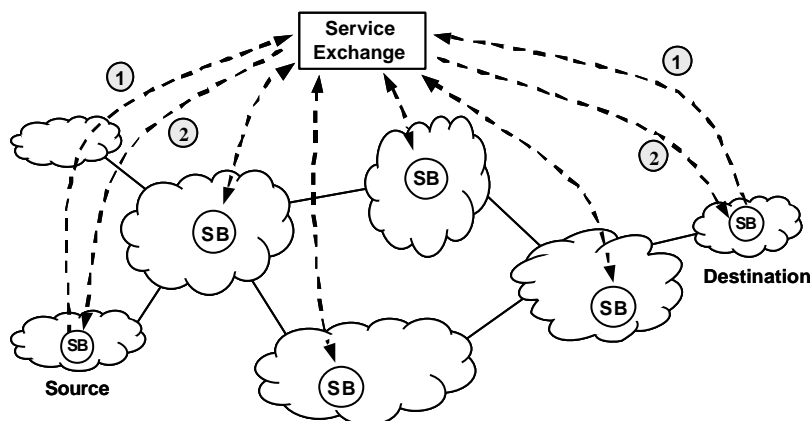

Figure 3 – Hierarchical Negotiation Model

Using the hierarchical model, routes can be configured in such a way that resource utilization is optimized and resource requests are satisfied as much as it is possible. Currently, inter-domain routing in the Internet is performed by the BGP protocol, based upon a shortest-path algorithm. The hierarchical model allows the utilization of algorithms based on more efficient heuristics. However, this gain in efficiency comes at the cost of changing the inter-domain routing within the SE area.

In the hierarchical model, a SE always receives service requests for aggregate traffic. Within each domain, the SB aggregates traffic by service and destination, makes predictions and sends them to the SE. SEs also have a hierarchical organization. Each SE aggregates the requests with a destination that is outside its area and sends them to a higher-level SE, and so on. The lower level SE is called SE_1, the next one is the SE_2 and the SE level n is called SE_n.

3. EVALUATION

Network providers interested in deploying end-to-end QoS-based services may be benefited with a comparative evaluation of negotiation models. We used the Network Simulator (*ns*) [10] extended to implement the required functionalities. The simulation study refers to a very simple WDS, defined by two parameters: throughput and delay. We used a voice user service for generating traffic, where the call arrival rate in each domain D_i is modeled as a Poisson process of intensity l_i calls per second and call duration is exponentially distributed with a mean of $1/m = 120$ seconds. In further simulations, we used a Pareto distribution for call duration and a VBR video service. The results did not significantly differ from this service, showing that the performance of negotiation models is considerably insensitive to these traffic models. The traffic load arriving at each domain is defined as $r_i = l_i/m$. Voice sources are modeled as an On-Off Markov process, which alternates between *on* and *off* periods and is also exponentially distributed with average duration of 1.004 s and 1.587 s, respectively. Each source generates CBR traffic at 80 Kbps when *on* and 0 Kbps when *off*.

A Local Gaussian Predictor was used for resource estimation. It is described as *local* because only the traffic samples collected at the last measurement interval, called T_{meas}, are used in the computations. Traffic samples are collected by border routers at regularly spaced time intervals of 1 second. The mean $\bar{x}$, and the standard variation, s, are computed and sent to the SB for further processing. The estimation of the throughput is calculated by $\hat{T} = \bar{x} + a\,s$, where a is a multiplier that allows some variability in the samples. $\hat{T}$ is expected to be exceeded with probability $1 - G(a)$, where G is the *cdf* of the standard normal distribution. We used 2.5 as the value for a, which yields a low probability of exceeding the predicted throughput of 0.006.

The choice of a network topology for simulation is an important step and it should represent as close as possible a real network. For the purpose of this study, four topologies were chosen (Figure 4): Abilene, GÉANT and RNP2, which are the backbones of the Internet2, European and Brazilian research networks, respectively. The fourth topology is a simple Manhattan network. Abilene has only 12 domains connected through links with identical capacity of 2.5 Gbps. Manhattan is a simple and regular topology, with 16 domain connected through 45 Mbps links. In GÉANT, 26 domains are interconnected through links varying from 34 Mbps to 10 Gbps. RNP2 also has relatively lower capacity links, from 1 Mbps to 25 Mbps. As far as the delay is concerned, interdomain links and domains were set according to a rough approximation of their physical distances. System load was distributed over domains according to the sum of the capacities of the links connected to them. Destinations for voice calls are chosen according to domain's in/out capacity.

For each evaluated scenario, simulations were carried out, with each negotiation occurring at periods of 60 seconds. We are not advocating the use of periods as short as 1 minute, since renegotiation intervals are not being evaluated in this paper. In [9], the performance of the predictor with longer periods is analyzed. However, there is a tradeoff between prediction accuracy and prediction interval [12]. For each scenario, 100 replications were executed, collecting values of the metric of interest at the end of each replication. The results presented next refer to the average of the metric of interest of all replications.
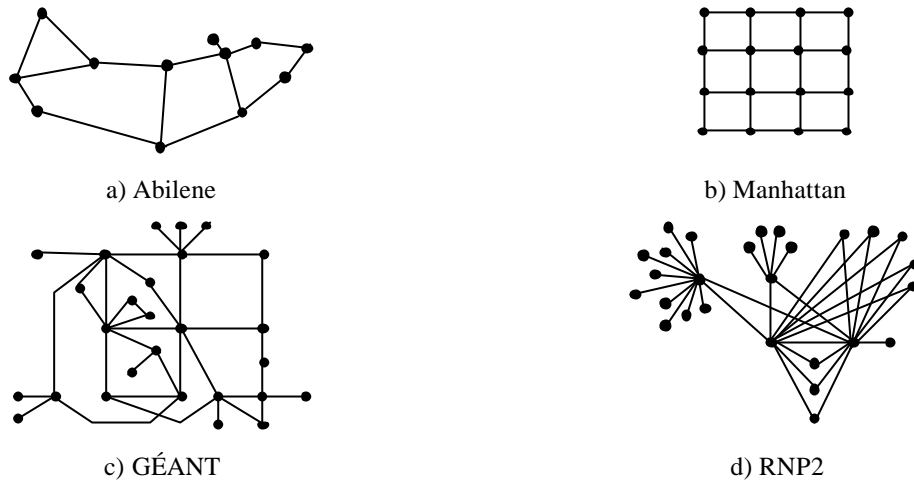
a) Abilene   b) Manhattan

c) GÉANT   d) RNP2

Figure 4 – Simulated Topologies

4. EFFICIENCY

Efficiency in service negotiation refers to the extent to which a model is able to allocate available resources. ISPs willing to deploy QoS-based services may be interested in evaluating the efficiency of such strategies, since despite the current wide supply of raw optic fibers, capacity for exchanging data in the Internet remains always a scarce and expensive resource.

4.1. Resource Grant Rate

In this section, we aim to provide ISPs some indication of how to compare the possible interconnection models, using resource grant as an index for efficiency. Let R_i and G_i be respectively the total amount of resources requested by and granted to domain i ($i = 1,2,\ldots,n$) to all destination domains. The resource grant rate is defined as $GR = 100 \times 1/n \sum_{i=1}^{n} G_i / R_i$.

In addition to using the basic shortest-path algorithm, resource allocation of the hierarchical model was further evaluated based on two different optimization schemes. When two or more paths satisfy a request, both schemes select the first path. When the request cannot be fulfilled by any path, they use different heuristics for finding a feasible path that has: the highest throughput or the lowest delay.

Figure 5 shows converging results for the four simulated topologies, when the voice user service is considered. The cascade and hub models generated values of GR very close to each other, because or their distributed negotiation style. Since there is no central entity for coordinating the resource allocation, and messages have to be sent to the various domains that take part in the end-to-end path for a given service, one negotiation request may easily interfere with the other ones. This happens because until a domain receives a confirmation message, it has to keep resources pre-allocated, thus creating a situation where existing resources remain unallocated even though other requests were not totally fulfilled.

For the hierarchical model, the negotiation is centralized. Consequently, it can allocate more available resources than the other two models. In general, the higher the load, the better the results of the hierarchical model. The highest observed GR gain was 20 for Manhattan, which represents an increase of 49.4%. These results show that, for the simulated scenarios, there is a clear incentive for using the hierarchical model.
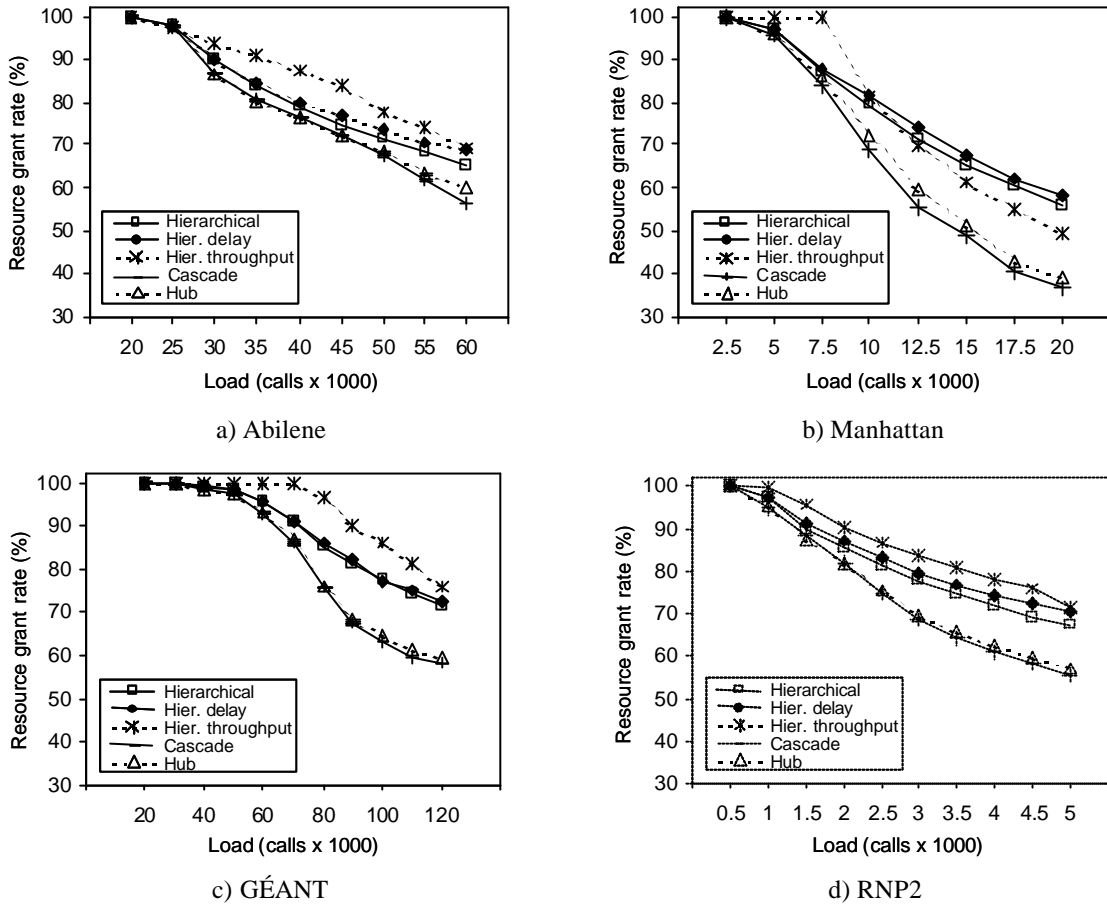
a) Abilene

b) Manhattan

c) GÉANT

d) RNP2

Figure 5 – Resource grant rate (GR)

The results achieved by the improved versions of the hierarchical model showed that, in most cases, the highest throughput heuristic is the more efficient, being able to grant more resources than the other two algorithms. One remarkable exception is for Manhattan, from 10000 calls on, where the heuristic fails and the curve falls below the other two ones. The lowest delay heuristic was not able to obtain any significant improvement. Our conclusion is that the benefits of deploying a more sophisticated algorithm are always observed for low to medium load. As the load increases reaching its maximum, the improvements are less significant (or even negative). For Manhattan, the GR gain of the highest throughput heuristic achieved its best improvement, of 12.4, which represents an increase of 14.2%.

4.2. Resource Provisioning Rate

The object of this section is to compare the negotiation models through an index that shows the degree of provisioning of the interdomain links. This index reflects the percentage of the resources offered by seller domains that were sold as a result of the resource allocation, and then effectively provisioned. The more resources are provisioned, the better is the use of existing resources that a given negotiation model can make. In turn, this is very useful when performing interdomain traffic engineering. Let O_i be the quantity of resources offered by domain i including all interdomain links and S_i the resource effectively sold during the negotiation. Then, the resource provisioning rate is defined as $PR = 100 \times 1/n \sum_{i=1}^{n} S_i / O_i$.

Figure 6 depicts the results of the PR. It can be observed that under the highest throughput optimization for the hierarchical model PR is always higher than under the other models. In

the best case, for Manhattan, *PR* of the highest throughput heuristic is 48 points (145%) and 60 points (181%) higher than the basic hierarchical and the cascade model, respectively. Another remarkable result is that even the basic hierarchical model is able to yield much higher provisioning rates than the cascade and hub models.
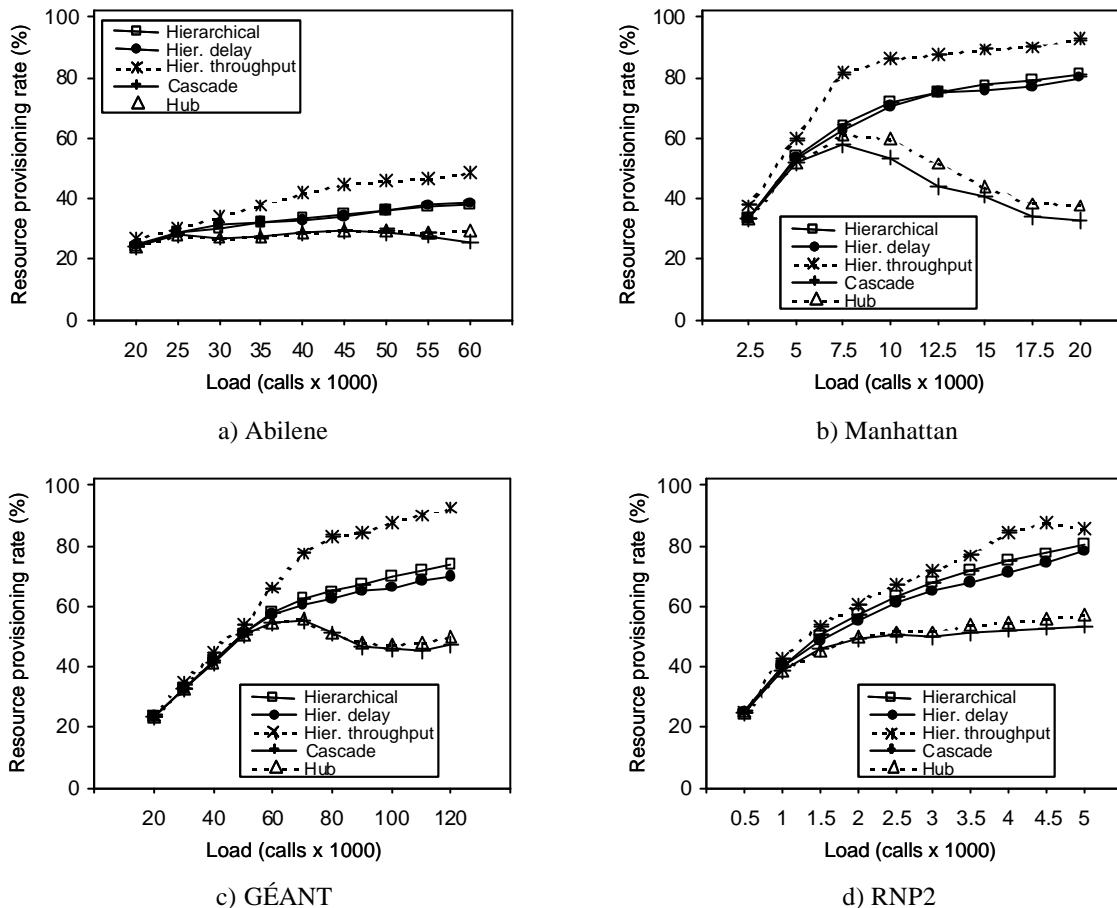

a) Abilene

b) Manhattan

c) GÉANT

d) RNP2

Figure 6 – Resource provisioning rate (*PR*)

The relevance of the results of Figure 6 becomes more evident when they are compared to the results of resource grant in Figure 5. Two important observations can be made. Firstly, there is an obvious relationship between resource grant and provisioning: the higher the grant of a negotiation model, the higher the provisioning it produces. Secondly, as the load increases, the grant rate decreases and the provisioning rate increases. Both conclusions fail however under certain conditions. Figure 5b shows that the highest throughput model is more efficient with light load but falls below the basic hierarchical and the lowest delay models from 10,000 calls on. However, in Figure 6b the curve for the provisioning of this model remains higher than that of the other two models. In this case, the highest throughput model found longer paths, so that although it generated a higher provisioning rate, the grant rate was lower. The reason is that more resources are provisioned to some requests, that otherwise could be allocated to others. The second conclusion fails for the cascade and the hub models. Because of the cross-interference among different requests, the allocation process is not able to grant them so many resources as the topology supports.

5. SCALABILITY

Another important criteria for ISPs to be able to carry out service negotiations and resource provisioning involving small and large groups of domains is scalability.

5.1. Allocation Requests

The number of allocation requests generated by a given negotiation model has a huge impact on its scalability, since the more requests are sent to be allocated, the more complex and processing intensive becomes the negotiation. As the number of allocation requests grows with the number of domains, it may also interfere with the efficiency of the negotiation.

For the hierarchical model, the main issue is the processing burden imposed by the resource allocation process, since it is carried by centralized entities. Let A be the number of requests generated by a given negotiation model. Then, $A = \sum_{i=1}^{l} \sum_{j=1}^{S_i} \left(M_{ij} \times \left(M_{ij} - 1 \right) \right) / 2$, where l is the number of levels of SEs, S_i is the number of SEs level l, and M_{ij} is the number of participants of SE_{ij}. Participants of a SE are domains for a first level SE and SEs of the immediate lower level for higher level SEs. It is straightforward to observe that when there is only one level of SE, that is $l = S_i = 1$, then $A = \left(n \times \left(n - 1 \right) \right) / 2 = C_2^n$, where n is the number of domains. Hence, there is one request from each domain to every other domain to be processed. For the cascade and the hub models, $A = \left(n \times \left(n - 1 \right) \right) / 2 = C_2^n$, that is also the worst case for the hierarchical model. The complexity for C_2^n is $O\left(n^2\right)$.

5.2. Message Exchanges

In order to send requests for service negotiation and receive responses regarding service grants, domains have to exchange protocol messages. Each time a message is received by an entity and processed, it is counted as a message exchange. For the hierarchical model, each domain sends one request to its SE and receives a response. SEs in turn also send requests and receive responses from their higher level SE. Let M be the number of messages and n the number of domains. Then $M = 2n + 2\sum_{i=1}^{l-1} S_i$, where l is the number of levels of SEs and S_i is the number of SEs at level i. $M = 2n$ for only a single level of SE. The worst case is when each SE has only two participants, such that $S_i = \lfloor n/2^i \rfloor$. Since 3 levels of SEs are expected to be enough for the whole Internet, the complexity is $O(n)$.

For the cascade and hub models, unless each domain has a peering link with every other domain, messages have to be processed by intermediate brokers. For the cascade model, $M = 2\sum_{i=1}^{n}\sum_{j=1}^{n} H_{ij}$, where H_{ij} is the number of domains between i and j, if $i \neq j$, and $H_{ij} = 0$ otherwise. The worst case is when the domains are interconnected in a linear way. Thus, the upper bound is $M = 2\sum_{i=1}^{n} i(i-1)$, because whenever a new domain is connected at one end, there is an increase in the number of messages of $2C_2^n$. In this case, the complexity is $O(n^3)$. For the hub model, the number of messages is twice the number for the cascade one. Considering that its upper bound is 1,333,332,000 messages for 1000 domains, scalability may be a concern, even though the real number of messages is smaller than the upper bound.

6. CONCLUSION

In this paper, we look into a very important aspect of the interconnection of domains: dynamic interdomain service negotiation. Negotiation models are strategies for domain interaction needed before the effective resource provisioning takes place. We recommend ISPs to use dynamic negotiation and to cooperate with each other as much as possible in order to use well-defined negotiation models that can help them to find the right financial incentives for QoS deployment.

Some of the lessons learned with our work are that the cascade and hub models achieved very similar results in the simulation study. Furthermore, the results showed that the hierarchical model significantly outperforms the other ones. As far as efficiency is concerned, the higher increase in the grant rate generated by the hierarchical model compared to the cascade model was from 15% to 50%. The increase in the provisioning rate was from 20% to 145%. Concerning scalability, it was found that the hierarchical model requires a significant lower number of resource allocations and message exchanges than the other two approaches. The basic reason for this behaviour is its very nature, based on centralized negotiation and hierarchical organization structure. We recommend the use of the cascade model for small groups of domains, due to its simplicity. The hierarchical model presents more benefits for large groups, due to both its complexity and better results achieved in our evaluations. A more detailed interpretation of the forthcoming results can be found in [9].

As future work, we intend to investigate economic models and migrating scenarios that can help us in removing the barriers imposed by the current interconnection situation to the deployment of end-to-end QoS-based services.

REFERENCES

[1] Cortese, G. et al., "CADENUS: Creation and Deployment of End-User Services in Premium IP Networks", IEEE Communications Magazine, January 2003.

[2] Duan, Z., Zhang Z. & Hou, T. T., "Service overlay Networks: SLAs, QoS and Bandwidth Provisioning", 10th International Conference on Network Protocols, November 2002.

[3] Engel, T. et al., "AQUILA: Adaptive Resource Control for QoS Using an IP-Based Layered Architecture", IEEE Communications Magazine, January 2003.

[4] Equinix, Inc., http://www.equinix.com.

[5] Goderis, D. et al., "Service Level Specification Semantics and Parameters", Internet Draft, <draft-tequila-sls-02.txt>, January 2002.

[6] Hatch, C. (editor) et al., "Selected Scenarios and requirements for end-to-end IP QoS management", EURESCOM P1008 Deliverable D2.1, January 2001.

[7] Huston, G., "Interconnection, Peering, and Settlements", INET'99, June 1999.

[8] Internap Network Services Corporation, http://www.internap.com.

[9] Kamienski, C.A., "An Architecture for Providing End-to-End QoS-based Advanced Services in the Internet", PhD Dissertation, Federal University of Pernambuco, Brazil, February 2003.

[10] Network Simulator (version 2.1b8a), http://www.isi.edu/nsnam/ns/.

[11] Nichols, K., Jacobson, V. & Zhang, L., "A Two-bit Differentiated Services Architecture for the Internet", RFC 2638, July 1999.

[12] Sang, A., & Li, A., "A Predictability Analysis of Network Traffic", INFOCOM'2000, March 2000.

[13] Subramanian, L., Stoica, I., Balakrishnan, H. & Katz, R. H., "OverQoS: Offering QoS using Overlays", First Workshop on Hop Topics in Networks (HotNets-I), October 2002.

[14] Teitelbaum, B., "Internet2 QBone: Building a Testbed for IP Differentiated Services", IEEE Network Magazine, September 1999.

[15] Trimintzios, P. et al., "A Management and Control Architecture for Providing IP Differentiated Services in MPLS-Based Networks", IEEE Communications Magazine, May 2001.