Service Definition and Negotiation in the Chameleon Architecture

Carlos Alberto Kamienski and Djamel Sadok

Computer Science Center
Federal University of Pernambuco, Recife, Brazil
Caixa Postal 7851, Cidade Universitária, Recife/PE 50732-970 Brazil
Tel: +55 81 271-8430, Fax: +55-81 271-8438
E-mail: {cak, jamel}@cin.ufpe.br

Abstract **-** This paper presents the approach adopted by the Chameleon Architecture for defining and negotiating advanced end-to-end services in the Internet. It proposes the use of Well-Known Services and a hierarchical model for service negotiation. A simulation study showed that the hierarchical approach has significant gains, related to efficiency and scalability.

Key-words: quality of service, well-known services, service negotiation.

I. INTRODUCTION

Current Internet provides just one service, called best effort, which is not suitable for supporting new advanced applications, such as interactive audio and video. These applications need advanced services with Quality of Service (QoS) performance guarantees. Furthermore, in order to be effective, QoS must be constant along the entire path between source and destination, i.e., end-to-end.

At present, domains maintain several bilateral peering agreements in order to exchange traffic and routing information. These agreements are negotiated statically and the time scale for renegotiation typically is in the term of months. With the introduction of new services, this situation tends to worsen, because of more frequent changes on service utilization patterns and the urge of finding new routes for meeting services' performance requirements. Therefore, *dynamic* service negotiation will be necessary, based upon efficient and scalable negotiation models.

In the Chameleon Architecture, domains interested in being involved in the deployment of advanced services have to agree in some common services, called Well-Known Services (WKS). WKSs are not tied to any particular implementation technology (e.g., IntServ, DiffServ), and even a domain may choose just to use over-provisioning. Chameleon also permits different service negotiation models to be used, while strongly recommending the use of the hierarchical model, for its efficiency and scalability.

The rest of the paper has the following structure. Section II presents the Chameleon Architecture. Sections III and IV emphasize service definition and negotiation respectively. Results of a simulation study are shown in Section V, while in Section VI some conclusions are drawn.

II. THE CHAMELEON ARCHITECTURE

The Chameleon Architecture, which aims to provide advanced end-to-end services in the Internet is divided in three logical planes, in order to provide flexibility to service definition and negotiation, efficient implementation and control of contracted services. The planes are: service plane, operation plane and monitoring plane. This organization provides users with a homogeneous view of the network, although the deployment of end-to-end services may need the cooperation of many networks, possibly using different QoS technologies.

A. Service Plane

The Service Plane plays a fundamental role in the Chameleon Architecture, providing an abstract interface for service negotiation, so that all domains offer a similar external behavior. An abstract interface is provided by the combination of WKSs, standard SLSs and a common service negotiation model.

The functionality of the Service Plane is implemented by the Service Broker (SB), which is responsible for, e.g., traffic prediction, service negotiation, resource provisioning and admission control. The SB may be seen as an extension of the Bandwidth Broker (BB) [4] proposed for DiffServ, but with two basic important differences. Unlike the BB, the SB is not tied to any QoS technology, and it is able to negotiate services based on other QoS parameters (delay, jitter), not only capacity.

B. Operation Plane

Each domain, through an internal set of policies, maps negotiated services to some mechanism (e.g. IntServ or DiffServ) used for resource provisioning and network configuration. Chameleon aims to separate service offerings from technology decisions.

C. Monitoring Plane

This plane is orthogonal to the other two planes. It collects information from the operation plane and feeds it into the service plane. It is responsible for continuously measuring QoS parameters for each service, sending the results to interested parties and possibly taking correcting actions.

III. Service definition

In the Chameleon Architecture, service definition is based upon the concept of Well-Known Service (WKS), which is a service that has a clear and unambiguous definition of the performance guarantees that a provider offers or wants to receive when an SLA (Service Level Agreement) is being negotiated. It must have the same behavior in every domain where it is implemented, to make it possible to deploy an end-to-end service to the end-users. Domains specify the desired WKS in the SLS (Service Level Specification) during negotiations by means of an identifier (WKSID).

There are some reasons for creating Well-Known Services. The most important is the difficulty for a domain to capture the semantic of a service and all implications for its implementation just by observing some parameters in the SLS during the negotiations, as proposed in [3]. WKSs help the development of some activities related to service deployment, like network planning, resource provisioning.

In the Chameleon terminology, a WKS is a transport service. Transport services make up the necessary infrastructure for deploying end-to-end services, which are implemented and negotiated by domains. On the other hand, end-user services are those meaningful to end-users, like voice and video.

Instead of attempting to define all possible transport services as WKSs, a better way is to group them in some classes, which encompass semantically similar services. From these classes new services can be instantiated through careful parameter configuration. Some required information to specify a class is: a) description of the service semantics; b) mathematical proof, depending upon the class; c) configurable performance parameters; d) an identifier (WKSID).

WKSs do not need to be standardized. Rather, a group of domains just has to agree on a WKS in order to deploy a new service.

IV. Service negotiation

The negotiation of an end-to-end service comprises two phases, end-user and transport service negotiation. End-user service negotiation does not produce resource provisioning. On the other hand, transport service negotiation results in resources provisioning in every domain involved in an end-to-end communication.

Resource provisioning may be done through advanced or immediate reservations. Advanced reservations are based on utilization statistics and must guarantee that most service activation requests will be accepted. On demand reservations are necessary in order to adapt available resources to instantaneous needs of users/applications. This paper only considers advanced reservations.

A. Bilateral Negotiation Model

Bilateral negotiation is a traditional model, which has been proposed for Bandwidth Brokers [4]. In this model, a domain negotiates with a neighboring domain, which in turn negotiates with its next domain, “rippling” through until the destination, using a negotiation protocol[6].

B. Hub Negotiation Model

In the Hub model [2], a domain which plays the role of a Service Provider is the responsible for negotiating with every domain that is in the route from source to destination in an end-to-end service. Therefore it is not considered scalable and its geographic scope should be restricted to just a few domains.

C. Hierarchical Negotiation Model

This model introduces the new concept of Service Exchange (SE), which is a central entity that coordinates service definition and negotiation. The SE performs negotiations on behalf of some participant domains, unlike the other models, where each domain needs to have individual agreements with several neighboring domains.

In order to achieve its goal, a SE needs information such as available services (WKSs), topology of its area and characteristics of inter-domain links. Furthermore, each SB periodically sends purchase and sale information (Figure 1a). Then, the SE performs negotiation "rounds", which result in end-to-end service permission (total or partial) or refusal.

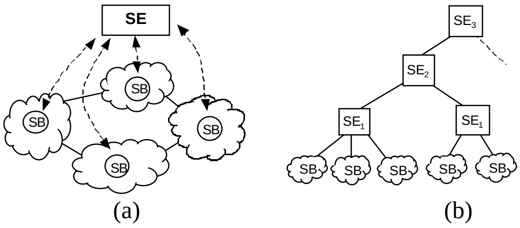

Figure 1 – Hierarchical model; a) Communication between SBs and SE; b) relationship between SEs

There are two types of domains in this model: service buyers and sellers. Some domains may play both roles. Seller domains send information to the SE about their external links and internal capacity. For each service, it comprises the WKSID and some QoS parameters. For buying services, important information include WKSID, scope, traffic specification and QoS parameters.

This model aims to provide scalability and efficiency to service negotiation. A SE always receives service purchase requests for aggregate traffic. Within an SE area, SBs aggregate traffic by service and destination, make predictions and send them to the SE. SEs also have a hierarchical organization. Each SE aggregates the requests where the destination is outside its area and sends them to a higher level SE, and so on (Figure 1b). SEs at a same level are not allowed to communicate peer-to-peer. In order not to overload a SE, if the number of domains gets too large in some area, the SE may be internally divided.

The efficiency from the hierarchical model is due to the knowledge each SE has of its area allowing it to perform negotiations considering several criteria, e.g., inter-domain QoS Routing (QoSR).

V. EVALUATION

This section presents a short comparison between negotiation models based on simulations (using the ns-2 [5] simulator).

A. Simulation Model

Simulations have been done using a modified version of the ATT network, with 45 Mbps links and 155 Mbps of domain internal capacity (Figure 2). Each domain was configured with delay of 10 ms and links according the their physical length to 10, 15, 20, 30 and 40 ms.

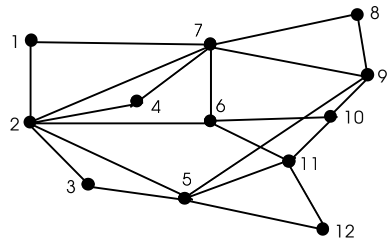

Figure 2 – Simulation topology

Simulated traffic refers to a voice service. The voice call arrival rate is modeled as a Poisson process and call duration is exponentially distributed with a mean of 120 s. Voice sources are modeled as an on-off process with average duration of 1.004 s ("on periods") and 1.587 s ("off periods"). Each source generates CBR traffic at 80 Kbps when "on" and 0 Kbps when "off ".

To support this end-user service, a simple transport WKS was defined with just one parameter: a maximum delay of 150 ms. A measurement and negotiation time interval of 1 minute was used based on previous simulations and on results presented in [1].

Domains constantly generate calls targeted to other domains (uniformly distributed). System load, given by the number of simultaneous active calls, has been varied between 1000 and 20000.

B. Efficiency in Resource Allocation

We define the efficiency of a negotiation model by the relation between the total amount of resources requested by a domain and the respective granted resources. For each system load, we computed the efficiency as the average of resources granted in every negotiation performed during the simulation.

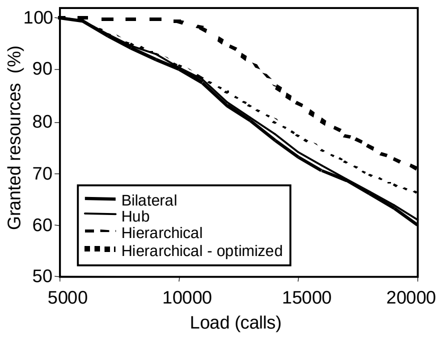

Figure 3 – Granted resources

Figure 3 shows that up to 5000 calls, all negotiation models were able to grant 100 % of requested resources, because system load is low and there is plenty of resources. The bilateral and the hub negotiation models had quite similar results for all simulated loads. Compared to the hierarchical model (not optimized), both had a similar performance for loads up to 12000 calls. From this point, the bilateral and hub models suffer from a sub-optimal performance because of interactions between messages that pre-reserve resources while the availability in downstream domains is checked.

Our comparison showed that the gain of the hierarchical model varies from 1 % to 6 % depending on the system load. Using an optimized algorithm for the hierarchical model (based on a kind of QoS Routing) the gain increases to 8 % to 12 %. Further simulations, varying topologies, link capacities and delays, showed a gain up to 10 % and 20 % with the hierarchical model with and without QoS Routing respectively.

C. Scalability

We evaluate the scalability of the negotiation models based upon the number of peering agreements that have to be maintained and the number of messages that have to be exchanged.

For our evaluated scenario, with 12 domains, the bilateral model have 42 peering agreements while in the hub model each domain has to maintain one agreement with every other domain (i.e., 132). For the hierarchical model each domain has to maintain just one agreement (with the SE), no matter the number of domains. If A is the number of agreements, than for N domains, A = N * N-1 for the hub model, A = k * N, k <= N-1, for the bilateral model, and A = N for the hierarchical model.

The number of exchanged messages is 488 and 976 for the bilateral and hub models respectively. In the hub model the number of messages is always twice the number in the bilateral model, because it needs two phases, one for requesting resources and another for notifying granted resource. For the hierarchical model, the number is 24. For N domains, the number of messages using the hierarchical will always be N * 2.

VI. CONCLUSION

In this paper, we presented the Chameleon Architecture and its approach based upon definition and negotiation of services for the deployment of end-to-end services in the Internet. Three negotiation models were presented: the traditional bilateral and hub models and the proposed hierarchical model, which uses the Service Exchange (SE) for coordinating the negotiation process.

A comparative evaluation showed that, for our simulated scenario, the hierarchical model was more efficient and scalable than the other two models. It could grant up to 6 % more resources without any optimization and up to 12 % with a kind of optimization based on QoS Routing. We showed further that the hierarchical model is more scalable, because it demands less peering agreements and it also generates less signaling messages than the bilateral and hub models.

REFERENCES

[1] Chuah, C. N., et. al., “QoS Provisioning Using a Clearing House Architecture”, IWQoS’2000, June 2000.

[2] EURESCOM Project P1008 Deliverable 2, “Selected Scenarios and requirements for end-to-end IP QoS”, January 2001.

[3] Goderis, D. et. al., “Service Level Specification Semantics and Parameters”, Internet Draft, November 2000.

[4] Nichols, K., Jacobson, V. & Zhang, L., “A Two-bit Differentiated Services Architecture for the Internet”, RFC 2638, June 1999.

[5] Network Simulator (version 2.1b7), http://www.isi.edu/nsnam, October 2000.

[6] Teitelman, B. & Chimento, P., “QBone Bandwidth Broker Architecture”, Internet2 QBone Draft, June 2000.