A Novel Model for Inter-Domain QoS Management for Real-Time Applications

Efrayim Metin[1,2], Mahbubul Alam[1,3] and Neeli Rashmi Prasad[1,4]

[1] Center for PersonKommunikation, Aalborg University, Fr. Bajers Vej 7A5, DK-9220, Denmark
[2] Delft University of Technology, P.O.Box 5031 2600 GA Delft, The Netherlands
[3] Cisco Systems International BV, 1101 CH Amsterdam, The Netherlands
[4]Ben Nederland BV,
E-mail: {emetin, jami, np}@cpk.auc.dk

Abstract

This paper proposes a novel model for inter-domain Quality of Service (QoS) management architecture consisting of a Common Open Policy Services (COPS) based Policy Control Management System, Clearing-House (CH) acting as a policy architecture that regulates the resource allocations to different groups of traffic, Open Settlement Protocol (OSP) used to negotiate between CHs and Differentiated Service (DiffServ) network with Dynamic Advanced Resource Allocation.

Key words

Quality of Services, Clearing House, Common Open Policy Service, Open Settlement Protocol, Differentiated Services, Resource Allocation, Advanced Admission Control.

1. Introduction

Currently proposed IP Quality of Service (QoS) techniques, such as Differentiated Service (DiffServ) [2], Integrated Service (IntServ) [3], IntServ over Diffserv [5] and Multi-Protocol Label Switching (MPLS) protocols are either non-scaleable, too immature or both to enforce and to manage end-to-end QoS in a commercial large-scale networks.

The guarantee of end-to-end QoS requires an efficient resource management mechanism that can reserve/control resources like bandwidth, delay and jitter according to a policy for immediate and future resource utilization. To provide guarantee of delay and jitter, strict admission control and delay calculation is necessary, and also the bandwidth usage has to be strictly policed [1].

The various mechanisms and protocols needed to provide QoS in IP networks, managing and coordinating them across a network can be a difficult task. It is not possible to manually configure every network device with the right queuing and traffic processing mechanisms to provide consistent, priority-based service everywhere necessary in a large-scale network. In addition, QoS applications must continue to work properly even if the network is dynamic and network topological changes frequently. The "traditional" network management applications cannot meet those requirements. To fully automate the decision-making process, using a third-party policy-based networking (such as bandwidth broker) offers a new way of controlling the QoS capabilities in the network.

Several Bandwidth Broker (BB) implementations have been proposed as a scalable mechanism for QoS provisioning over DiffServ architecture [4] but they do not optimize end-to-end path selections, considering only traffic between two domains and no advanced resource reservation.

This paper presents the ideal level of granularity of CH performance in respect of the trade-off between end-to-end QoS on one hand and scalability on the other hand and to understand the behavior of AC algorithm used. The architecture is designed to be modular and hierarchical, allowing future modification and addition to existing framework. The major units of the architecture are the Policy Decision Points (PDP) such as CH [8], the access edge Policy Enforcement Point (PEP), the core PEP such as the Edge Router or it can be Gateway GPRS support Node (GGSN) for a 3G network. The simulation for this inter-domain model describes the usage of COPS [6] for dynamic Resource Allocation in a DiffServ network. COPS- DiffServ Resource Allocation (COPS-DRA) combines the outsourcing and provisioning model [7].

2. QoS Management Model

The guarantee of end-to-end QoS requires an efficient resource management mechanism that can reserve/control resources like bandwidth, delay and jitter according to a policy for immediate and future resource utilization. To provide guarantee of delay and jitter, strict admission control and delay calculation is necessary, and also the bandwidth usage has to be strictly policed. The various mechanisms and protocols needed to provide QoS in IP networks, managing and coordinating them across a network can be a difficult task. It is possible to manually configure every network device with the right queuing and traffic processing mechanisms to provide consistent, priority-based service everywhere necessary in a large-scale network. In addition, QoS applications must continue work properly in the face of dynamic network and organizational changes. The "traditional" network management applications cannot meet those requirements. To fully automate the decision-making process, using a third-party (CH) policy-based networking offers a new way of controlling the QoS capabilities in the network.

This QoS model distinguishes it self from the traditional approaches in several aspects. Firstly it has hierarchical and distributed management architecture, which is strongly attributing to the scalability issue. Secondly, the network resource utilization is optimized, by separating the signaling plane from the data plane. This is achieved by separating the call set-up process, Admission Control (AC) [10,11] and Resource Allocation (RA) from the data traffic. Segregation of signaling/control messages from the data traffic allows to attain easily measured knowledge about the network utilization prior to data flow, as a result of in advance, measurement based admission control is performed on the edge of each Autonomous System (AS). Thirdly, for the intra-domain the local CH interprets traffic specifications received from access and/or core edge router and selects the best packet traversal routes and makes flow based resource reservation within the Logical Domain (LD), aggregated resource reservation between the LDs within the same Basic Domain (BD) and aggregated advanced provisioned resource for the inter-domain while maintaining the billing information for each user or for the higher layer CH as shown in Figure 1.

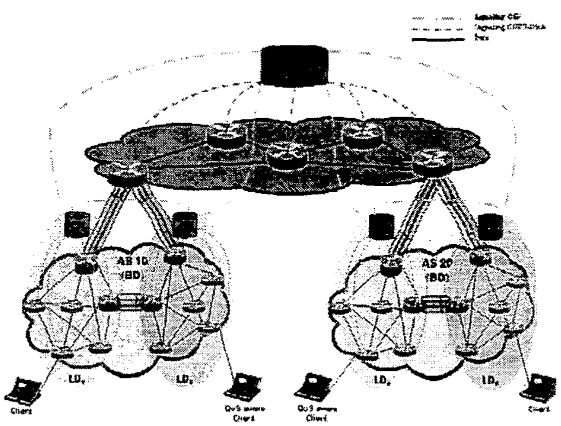

Figure 1: High-level Intra- and Inter- domain QoS management model

The local CH (CH_0) in each LD within a BD uses its internal cache information about intra-Internet Service Provide (intra-ISP) reservation status and the call specification to search for the optimal path within its own logical domain. If the Callee resides in a different ISP, the call specification(s) will be forwarded to the Global CH (CH_1) after being aggregated to check the inter-ISP reservation status on the end-to-end path. The CH_1 aggregates reservation requests for calls that travel from the same ISP. The clustered reservation requests are sent periodically to each of the ISPs involved in the path. The path is optimized based on the desired quality of service (e.g. network latency), reservation availability.

2.1. Traffic Management and Admission Control

Traffic management is a set of policies and mechanisms that allow network to efficiently handle a diverse range of requested services. Traffic management includes the following operations: Admission Control, Scheduling, buffer management and flow control. Here Admission Control is the key component in ensuring QoS objectives in a communication network and has important consequences for the performance of it. The admission control algorithm depends on the QoS criteria (e.g. reservation, priority, etc.). Its main goal is to support QoS service demands of (real-time) applications via resource reservation. It determines whether a new traffic flow can be admitted to the network such that all users will receive their requested performance or not.

In general admission control schemes can be grouped into two categories, parameter based and the measurement based admission control. The parameter-based method performs admission control based on the requested traffic parameters (peak rate, mean rate, effective bandwidth, etc.). The drawback of this method is the need in storage of huge amount of traffic parameters and status of all requested connections in the node, which will give problems in the scalability aspect when deployed in large IP networks.

Whereas, Measurement-Based Admission Control (MBAC) method is a dynamic algorithm making decisions using the measurements about the current (actual) state of the network after every certain time interval (Time window, *Tw* with sample frequency *s*). In this model the network nodes needs 'constantly' to monitor the (quantitative value of the) traffic passing through it, rather than a priori specified traffic characteristics. Using MBAC resources can be allocated according to measured properties of *aggregated* traffic rather than individual flows. An important assumption at this point is the recent past behavior (maximal/peak rate envelope) of the (aggregate) traffic flow will continue to bound future packets arrivals [9].

The main goal of MBAC is to estimate the network status and decide whether to accept or reject the new connection requests (based on its QoS violation probabilities). The Admission control function has to also take care about the guarantees given to QoS requirements of the already established flows, which may not be violated by the new connection requests. In addition, in MBAC the admission decisions are made based on network states, which can be either deterministic or statistical traffic parameters.

MBAC provides higher network utilization and is more suitable to support real-time applications that are tolerant of occasional QoS degradation. Although it cannot provide guaranteed service, it is sufficient to provide better-than-best-effort service.

Furthermore MBAC has the following advantages above the parameter-based approach:

- Is scalable, it does not need to maintain the network state information in all the nodes.
- Increased utilization when the sources submit conservative estimates of the traffic
- Decreased congestion when the sources send packets that are in excess of the declared traffic descriptors.

2.2. Resource Reservation

The CH architecture can support two types of reservation: advanced and direct reservations. An advanced reservation (AR) is time-limited and resources are allocated in advance, based on the statistical estimates of aggregate traffic over a particular link (saved in a Repository). This method is used to reduce the violation in QoS assurance if the traffic arrives before the resources are properly reserved. Advanced reservations do not reflect to the rapid fluctuations of local traffic produced by end-users. Direct reservations (DR) on the other hand, can be made on demand when existing reservations become insufficient to accept the new admission request.

In this CH model resources will be allocated/reserved only if the requested Bandwidth (BW) is available in both the sender's domain and the destination domain as well as on the inter-domain link. A traffic predictor is used to estimate the required capacity based on Gaussian approximation of the aggregate traffic arrival, which can be simply characterized by two parameters: its mean μ, and variance σ^2.

Since only aggregate reservations are established for particular link and not for individual flow, this approach only requires maintenance of aggregate state information in all the routers.

2.3. QoS Metrics and Parameter

The new Internet multimedia applications have significant bandwidth requirements. Others have strict timing requirements, or function one-to-many or many-to-many (multicast). These require network services more than the normal "best-effort" service currently provided by IP. As a consequence, IP networks needs to be "intelligence", they're expected to anticipate. Network applications can be characterized in terms of how predictable the data rate is and how tolerant in delay delivery. Generally, two-way applications are more sensitive to delay than one-way.

The representation of QoS is a bottom up approach, which needs to define a minimum set of QoS parameters. The initial QoS approach is to gather statistical data for the QoS parameters that reflect the call setup quality and the call quality. End-to-end QoS can be broadly categorized into call setup quality (e.g., call setup time, dropping probability) and call quality (e.g., end-to-end delay and speech quality). There are many contributing factors and the initial set of QoS parameters that are of interest are:

- Packet-loss (in both directions)
- Dropping probability
- New call blocking probability. In a buffer-less multiplexer, packet loss occurs whenever the aggregate input arrival rate exceeds the link capacity (C). Loss probability is subdivided in two variants:
  - *Tail probability* of the queue length distribution $P_t(Q>B)$, which refers to the fraction of time an infinite buffer queue's occupancy exceeds B.
  - *Loss probability* P_l, fraction of bits dropped by a queue that has finite buffer space B
- Call delivered ratio
- Bandwidth (guarantees/availability)
- Throughput
- Requested resources
- Aggregated traffic rate

Both packet loss and delay related to a specific call, while the CH computes the call delivered ratio. The call delivered ratio formula is:

$$Call\ delivered ratio = \frac{of\ calls\ completed}{of\ call\ attempts},\ \text{where,}$$

$$of\ call\ completed = \sum (answered, busy,\ unanswered)\ calls$$

3. QoS Management Model for 3G

The Go interface in 3GPP Release 5 architecture will allow service-based local policy and QoS inter-working information to be "pushed" to or requested by the PEP in the GGSN from a Policy Control Function (PCF). The PCF is a logical entity of the Proxy Call Session Control Function (P-CSCF) and its functionality is similar to Local CH in the proposed model. COPS for Policy Provisioning (COPS-PR) is used to communicate service-based local policy information between PCF and GGSN. The GGSN sends requests to and receives decisions from the PCF. COPS-PR can be extended to provide per-flow policy control along with a 3GPP Go Policy Information Base. As we all know that per-flow based policy control is neither scalable nor it utilizes the network efficiently. Therefore, we believe that the proposed QoS management model namely, COPS-DRA which combines both outsourcing and provisioning model, providing a scalable, flexible and efficient usage of network resources and therefore, would be more suitable for 3GPP architecture.

4. Simulation Model

The simulations described in this paper are performed to identify the ideal level of granularity of Clearing House performance in respect of the trade-off between end-to-end QoS on one hand and scalability on the other hand. The model used for this purpose uses a coarse grained model of the Internet as shown in Figure 2. The inter-network is modeled as interconnected autonomous systems (AS). Some of these systems are host networks, which act as traffic sources and sinks, the rest are Internet

Service Providers (ISPs), which act as pure transport networks. The Simulator implements the control logic of the Clearing House architecture and mechanisms.

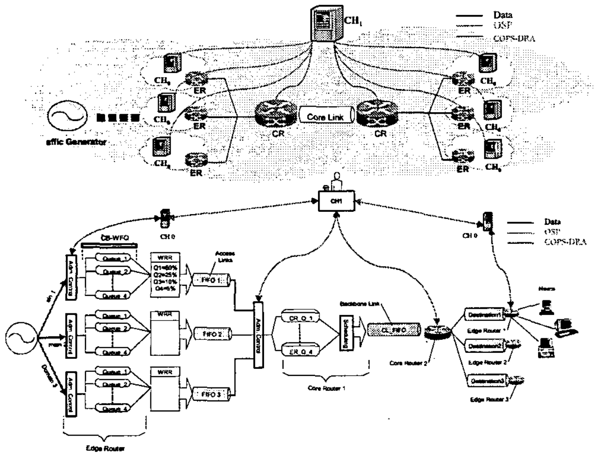

Figure 2: Detailed architecture of the simulation model for Inter-domain QoS management

4.1. Simulation Results

The CH architecture is designed to improve the end-to-end performance by rationing the number of high priority admitted by the edge router and managing network resources on congested links based on continuous network monitoring.

Different scenarios with mixed prioritized and best effort traffic accross the network will be performed to gather statistics on the load control performance of the Clearing House architecture. The call arrival rate in each domain *i* is modeled as an independent Poisson process of intensity λ_i packets per second. In this simulation a global CH (CH_1) keeps track of the reservations along the various links in the topology, see Figure 2. The six local CH (CH_0) nodes perform local admission control while the CH_1 processes call requests between the ISPs and coordinate aggregate reservation on the various inter-domain links. The reservation status is maintained in the database, which is constantly updated (measurement based). The load is defined here as the number of inter-domain reservation requests per second arriving at the CH_1 node. Throughput is measured as the number of calls serviced by the CH-node per second.

Throughputs for different scenarios were measured and plotted in Figure 3. It can observe from Figure 3 that for the proposed model the CH at different levels is stabilizing very soon. Using an *Adaptive & Dynamic Scheduling (ADS),* the CH can successfully take a load of 3200 calls/s while the CH using a First In First Out (FIFO) scheduling can only take a load of 1700 calls/s.

Adaptive & Dynamic Scheduling is an enhanced version of a combination of weighted Fair Queuing (WFQ) and Weighted Round Robin WRR). Here the weights, i.e. the service rate of each class, are not fixed. According to the arrival rate and buffer occupancy the ADS weights are adapted so that the delay and delay spacing between classes can be properly controlled, to assure the mission-critical packets to be delivered on time. ADS scheduling increases the throughput of the system by servicing call requests efficiently. The throughput of the ADS drops once the load increases beyond 3700 calls/s. At this point its throughput is 78 %.

A call is "blocked" when reservation request is dropped by the scheduler either due to insufficient resources, buffer or excessive load. Within this number the packet loss is included as well. Here the packet losses caused by buffer overflows in routers as well as discarded packets are also taken into account. It should be clear that all non-conforming packets must be dropped before any conforming packet is. Conforming traffic must be delivered with very low losses.

From figure 4 it can be seen that the blocking rate obtained using a FIFO scheduler is much less then that of ADS scheduler. As the number of flows increase, the amount of policing and state needed at the edge router also increases. The call blocking-rate is almost negligible until a load of 1200 calls/s. For the Adaptive scheduling the call blocking rate is less than 9% until a load of 2900 calls/s.

After a high-priority flow is admitted, it is very important to verify that this flow only uses its allocated share of network resources. Policing refers to monitoring of aggregated groups of admitted flows and detecting specific groups that violate their total allocated bandwidth. After a certain threshold, the blocking rate increases linearly with the load indicating a saturation point of throughput.

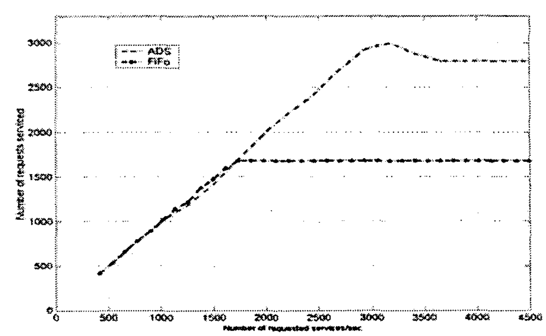

Figure 3: Throughput of a Clearing House as the traffic load is varied

In the context of this paper over-provisioning is used to express the amount of over-allocated bandwidth for the high-priority traffic. The estimation of the required bandwidth (BW) can represented as $BW = \mu + \delta \cdot \sigma$, with

a mean of μ and standard deviation of σ. Here δ is a scalar
factor that controls the amount to which the bandwidth predictor put up variability in the samples. The predictor is based on a Time window (T_w). The predictor can over-estimate the bandwidth requirements that vary at a shorter time, resulting in possible QoS violation. Under-estimation will lead to call rejections. Figure 5 presents the percentage of the total bandwidth versus simulation rounds for different values of T_w. The over-allocated bandwidth at the start of the simulations is due to the initiation parameters and the very low injection of high priority flows. However, the used algorithm reaches equilibrium of ≈ 20% in the Access Link. The plot shows that the transmission link will be more over-provisioned for a smaller T_w for a constant value for the δ.

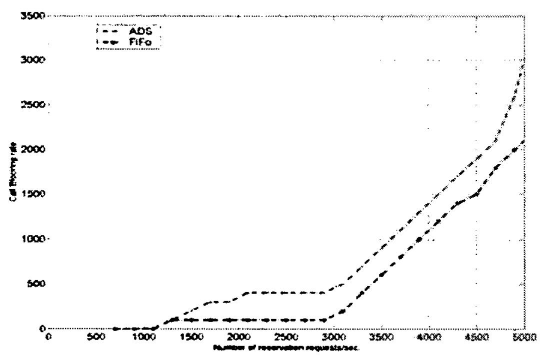

Figure 4: Call Blocking rate as the traffic load is varied.

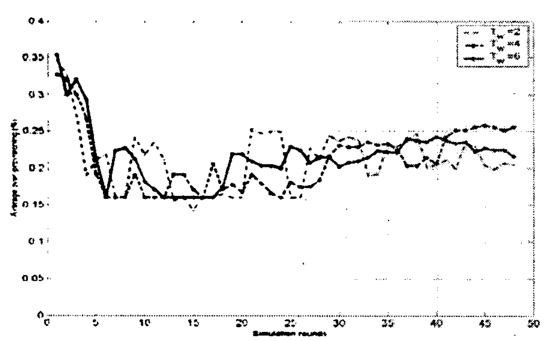

Figure 5: Over-provisioning in the Access Link

5. Conclusions

Enabling QoS introduces complexity in several aspects, starting from applications, different networking layers and architectures to network management issues. This paper presented a novel QoS management model and a new approach to reserve and control resources in IP networks. From the simulation results it can be concluded that CH hierarchical and aggregation model allows to realize a fully distributed and scalable resource control framework. In this model wide area network is treated as a collection of smaller routing domain called a basic domain (BD) which can be a local POP network of an ISP. Several BDs can be aggregated to form logical domain (LD). This introduces a hierarchical tree of LDs and a local CH associated with each LD to regulate the intra-domain aggregate reservation. CH serves the distributed resource control system in which the computational load for control are distributed to various nodes at different level of granularity. In this model, the resource reservations are performed on aggregated bases for the flows that share the same ingress-egress points. These reservations are adapted dynamically based on bandwidth predictors using dynamic algorithm making decisions using the measurements about the current state of the network after every certain time interval.
However, further study through experimental result are needed to show that the QoS model can improve the flexibility and assurance provided by current solutions, while maintaining a high level of network utilization.

References

[1] M. Alam, J.R. Farserotu and R. Prasad, 'Quality of Service among IP-Based Heterogeneous Networks' IEEE Personal Communications, December 2001.

[2] S. Blake, "An architecture for Differentiated Services," RFC2475, Dec 1998.

[3] R. Braden, D. Clarck and S. Schenker, "Integrated Services in the Internet Architecture: an Overview," RFC 1633, June 1994.

[4] Qbone Bandwidth Broker Architecture, http://qbone.internet2.edu/bb

[5] Yavatkar, F. Baker, L. Zhang, M. Speer, R. Braden, B. Davie, J. Wroclawski, E. Felstaine, "A Framework for Integrated Services Operation over Diffserv Networks," RFC 2998, Nov. 2000

[6] D. Durham, J. Boyle, R. Cohen, S. Herzog, R. Rajan and A. Sastry, "The COPS (Common Open Policy Service) Protocol," RFC 2748, Jan 2000

[7] S. Salsano, "COPS usage for DiffServ resource Allocation (COPS-DRA)," work in progress, Oct 2001.

[8] Randy H. Katz, Chen-Nee Chuach, Anthony D. Joseph, "QoS Provisioning Using a Clearing House Architecture," IEEE Magazine 2000.

[9] J. Qiu and E.W. Knightly, "Measurement-Based Admission Control with Aggregate Traffic Envelopes," IEEE/ACM Transactions on Networking, Apr 2001.