SCALABLE RESOURCE MANAGEMENT ARCHITECTURE FOR VoIP

Christian Prehofer[1], Harald Müller[1], Josef Glasmann[2]
[1]Siemens AG, 81359 Munich, Germany
e-mail: Christian.Prehofer@mch.siemens.de, H.Mueller@icn.siemens.de,

[2]Lehrstuhl für Kommunikationsnetze, Technische Universität München,
Arcisstr.21, 80333 München, Germany
e-mail: glasmann@ei.tum.de

ABSTRACT

This paper presents a new approach for scalable Quality of Service (QoS) based on a resource manager with the focus on Voice over IP (VoIP) in corporate networks. We first review the requirements of VoIP and consider typical network scenarios as well as existing approaches. In our approach several network domains have local resource managers (RM), which are in charge of the local network resources. This architecture has the advantage that only one new network element is needed, which also simplifies QoS signaling. We furthermore discuss ways for automatic configuration of the resource manager and enhance VoIP signaling (H.323) by QoS signaling with the resource manager. We show that the architecture fulfills the main requirements for VoIP, in particular minimal call setup delay and low management and installation effort.

1. INTRODUCTION

This paper first discusses the requirements and current approaches to QoS for VoIP. Then we present a new approach for scalable QoS based on an resource manager architecture.
With VoIP, an IP-based network typically carries data traffic as well as multimedia traffic. While this integration can lead to savings in network infrastructure, the current best-effort IP networks require considerable over-dimensioning to enable real-time applications. This is possible in many local area networks (LANs), but it is not a general solution for heterogeneous networks. Since the user expects QoS as in current switched phone service, this must however be extended to the wide range of networks.

In our approach, we focus on an end-to-end solution for VoIP in corporate networks. We assume that a call may traverse several network domains where each domain has a local resource manager, controlling the local network resources. Furthermore, we consider VoIP signaling with a gatekeeper-routed model, where the call signaling proceeds over one or more gatekeepers. With our QoS extension, the gatekeepers additionally request the needed network resources from the resource managers. An efficient solution for this QoS signaling is one of the main contribution of this paper.

We show that our architecture fulfills the following main requirements (of QoS) for VoIP:

- complementary to H.323 architecture
- minimal call setup delay of QoS signaling. In most scalable approaches, extra signaling is needed for QoS. The impact on call setup time shall be as low as possible.
- suitable for heterogeneous QoS networks. In some networks, like LANs, bandwidth is abundant and simple solutions like layer two priority mechanisms (IEEE 802.1Q) are sufficient. In other networks, like backbones or corporate networks call admission has to prevent overload situations in the network.
- minimal management and configuration overhead.
- migration strategy with incremental introduction with minimal changes to existing network infrastructure.

2. VoIP WITH H.323

VoIP systems are an emerging technology, which will both succeed and complement conventional PBX technology in corporate networks. Figure 1 shows a schematic VoIP scenario. It includes both PC-based clients as well as VoIP telephones on two LANs, which are connected via a WAN. We assume here the H.323 [H.323] VoIP architecture consisting of two H.323-Zones managed by one gatekeeper each. The main tasks of a gatekeeper are of administrative nature (registration, address resolution, call detail recording….). In addition, the gatekeeper may perform bandwidth management. In our architecture we separate this task from the other tasks and establish a separate resource management system behind the gatekeeper (GK). Within its domain, the RM is responsible for

resource management and those parts of admission control related to bandwidth management.
Inter-working with the public switched network (PSTN) is achieved via a gateway.

We focus on the H.323 protocol suite since H.323 is mature and widely available. However, most of the discussion is not specific to H.323. For instance, the basic signaling with H.323 fast connect is comparable to the SIP call setup.

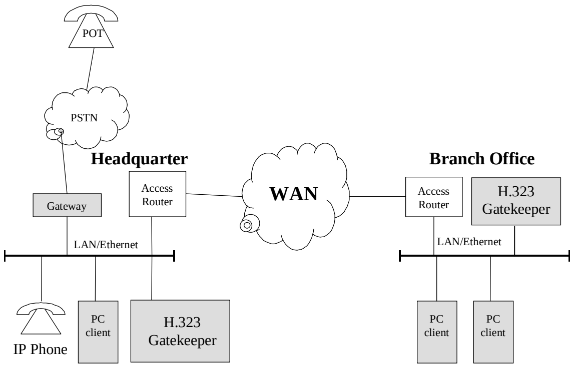

Figure 1: VoIP Scenario

3. QoS REQUIREMENTS FOR VoIP

Customers of VoIP solutions compare the new VoIP system with their old ISDN/PBX system. Thereby QoS aspects like voice quality and call setup time are one of the most important technical issues. QoS for VoIP comprises mainly two separate issues:

- Call setup time
- Voice quality

The call setup time heavily depends on call processing and additional backend services.
Regarding voice quality, the main network aspects are delay, jitter and packet loss in the IP network. Notice that high jitter can result in loss in the playback buffer at the receiving side. In addition, we must consider the end-system performance.

For both of the above issues, we need to consider network QoS parameters such as delay and loss. Since realtime critical multimedia and data traffic are currently both treated in the same way as best effort traffic, the voice and video quality heavily depends on the data traffic load. The general problems with data traffic are

- Data traffic is heavily bursty and unpredictable ("self similar")
- Greedy TCP traffic takes as much as it gets until packet loss occurs
- TCP self-control does not apply for short-lived flows, which are typical for Web applications. (Flows terminate often before TCP congestion control applies)

We discuss in the following the main quantitative parameters for VoIP QoS. For further details, we refer to the ETSI Tiphon recommendations [Tiphon-Qos].

3.1. Call Setup time

An important characteristic of end-to-end QoS besides the call quality is the "call setup time". i.e. the time elapsed from the end of the user interface command by the caller (keypad dialing, email alias typing, etc) to the receipt of a meaningful tone (alerting) by the caller.

The call setup time depends on the used signaling protocol variant, the call processing overhead and the network QoS. By using H.323 Fast Connect Procedures, the media channels are established end to end with the first H.225.0 end to end backwards message received at the calling endpoint.

For the call setup time a main problem is packet loss, since this leads to retransmissions and delay (in the worst case, these are triggered after waiting for a TCP timeout). Note that QoS signaling may influence call setup time.

Call setup time for calls within the LAN should be less than 0.5 seconds. For calls via gateways to public networks the overall call setup time depends on the public network call setup time, which can be significantly higher.

3.2. End-to-end Delay

Several studies about delay have been conducted and reported in the scientific literature; they lead to the following conclusions [Tiphon-Qos]:

- Small delays (10 ms to 15 ms) are not annoying even in absence of echo suppression, since the human ear dose not perceive the effects as echo;
- Delays up to 150 ms require echo control but do not compromise the effective interaction between the users;
- If the delays are in the range 200 ms to 400 ms, the effectiveness of the interaction is lower but can be still acceptable, the perceived voice quality starts to degrade;
- If the delay is higher than 400 ms, interactive voice communication is quite difficult and conversation rules are required (as for "Walkie Talkie" communications).

Packet switched data networks also have another problem: delay is usually variable. While telephone services require fixed delay transmissions, data networks cannot provide this because of their "best effort" policies; different packets may have different delays because of traffic conditions: this variation is usually known as network jitter.
Figure 2 illustrates the end-to-end delay budget for voice payload in a VoIP system. On the sending side, this includes encoding and packetization as well as packet/protocol processing. On the receiver side, an additional delay is introduced by the playback buffer, which is needed to compensate jitter.

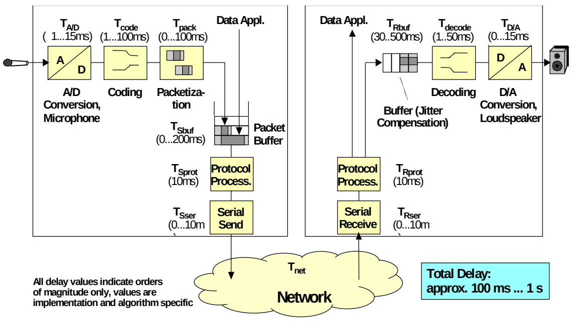

Figure 2: End-to-end Delay

Note that delay requirements on the network are quite high, since the end systems already consume a large part of the delay budget. With respect to the tolerable overall system delay we recommend a network delay lower than 50ms.

4. TYPICAL QoS PROBLEMS AND SOLUTIONS

In the following we show different network scenarios and discuss the appropriate approach for QoS. This also serves to illustrate the need for an architecture, which can accommodate different local solutions.

4.1. LANs

In switched layer two networks bandwidth is normally cheap and abundant. Hence over-dimensioning is feasible for QoS within LANs. Typical LANs with up to Gigabit Ethernet speed can provide ample bandwidth. However, it is often difficult to estimate the bandwidth requirements of high-end applications. Huge delays in the terminals require low networking delays. Even in the case of overprovisioning the jitter problem still exists. Even provisioning of high link capacities does not protect from short time bursts, which may lead to temporary overload situations. This overload can occur in a switch, when traffic is aggregated or bursty data traffic is multiplexed with realtime traffic.

This scenario discusses potential problems in typical LANs with high-end applications, demanding huge amounts of bandwidth. Examples for such applications are

- Application servers
- File servers
- Web servers, incl. Multimedia

In this scenario, problems can arise from the following:

- Congested links i.e. packets must be dropped by the switch if the queues are exhausted.
- Congested switches; most switches have a maximum throughput which is below the possible input rate.

In the end systems performance is needed to handle the input queue (at the driver or at the socket level). Typical remedies are task priorities, if the operating system permits.

A simple approach, which is feasible in most LAN scenarios, is using a separate traffic class for realtime critical multimedia traffic with DiffServ and/or 802.1Q and to dimension this for the worst-case multimedia traffic load. Hence a resource manager may simply admit any calls or use some simple thresholds.

4.2. Enterprise Networks

Typically, larger enterprise networks are heterogeneous in nature. They may consist of LANs, which may be connected with each other via leased lines or backbone networks. In our scenario two LANs are connected via a low bandwidth link, see e.g. Figure 1. The bandwidth in the LANs itself is in many cases sufficient for multimedia and data traffic. The link between the two LANs is typically realized with PSTN lines. Compared to links within the LAN, their cost per bandwidth is very expensive. It is therefore desirable to minimize the bandwidth of these links. This leads to a congestion spot which causes delay and packet loss.
A simple approach for this scenario is to extend the above LAN solution by an explicit bandwidth control for those bottleneck links. Hence a resource manager must be able to detect whether calls proceed over this link and may only admit calls up to the limit of the bottlenecks.**Backbone Networks**

In this scenario a lot of participants (with different user behavior) are connected to a backbone. Overprovisioning is very expensive and in most cases impossible. In the contrary to LANs the mean network utilization in backbones is much higher than in LANs.
Another important point for network operators here is the increasing number of participants. In former

well-dimensioned networks overload situations accumulate. These bottlenecks can not be upgraded at once.
The consequence is that traffic becomes less predictable and fluctuant traffic load situations occur over time. Hot spots normally occur on different links in the backbone dynamically. In most cases, reservation based resource management is preferred over measurement based approaches. In spite of Gbit/s links jitter may become very high. Relatively high loss probability can not be precluded globally.

The main problem of backbone networks is that per-flow bandwidth management in routers is overly expensive. Hence only aggregated bandwidth, e.g. for aggregated videoconferencing calls should be done. This is possible with a resource manager, which manages traffic aggregates in the backbone.

5. CURRENT APPROACHES TO QoS

In view of the above requirements and typical problems, we discuss the usage of existing and upcoming QoS mechanisms. We discuss below the two main approaches:

- introduction of separate traffic classes for voice and data
- explicit bandwidth allocations on a per-flow basis with RSVP and IntServ.

Another, more general approach is currently discussed in the Internet2 [I2-BB] (without the focus on VoIP). A similar, general architecture concept is considered in the ETSI Tiphon standardization [Tiphon-Qos].

5.1. QoS by Prioritization

The state-of-the-art QoS approach with prioritization can be summarized as follow:

- Protect multimedia traffic and H.323 signaling from data traffic bursts in LANs.
- Limit multimedia traffic or provision for worst case multimedia load

Traffic prioritization can be achieved on Layer 3 [DiffServ] and/or on Layer 2 [802.1D] for Ethernet. In both cases, this comprises the following steps:

- Classification of the packets to be prioritized into several QoS classes.
- Marking of these packets by a few bits in the IP or MAC layer.
- Priority forwarding of packets based on the marking, typically by separate queues for the QoS classes.

Clearly, the first two steps have to be done only once, preferably in the end-system. Priority forwarding should be done by as many systems (switches, routers) as possible. Note that we only use the notion of packet here and do not distinguish packets (layer 3) and frames (layer 2).

The QoS prioritization standards recommend to use at least two queues, as illustrated in Figure 3.The traffic of separate classes is put into separate queues in the switches and routers. Packets, stored in a queue with low priority are only served by the scheduler if the high priority queue is empty. Figure 3 shows one queue with multimedia traffic and a second one for data traffic.

On layer 2, usually only simple priority queuing is possible, while layer 3 devices permit to allocate bandwidth to queues. For the voice/video queue, sufficient bandwidth has to be allocated by appropriate configuration.

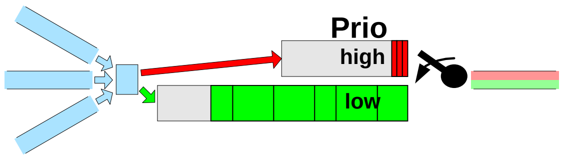

Figure 3: Prioritization via Separate Voice/ Video and Data Queues

Note that with this approach no per-flow queuing is performed, but for LANs aggregate priority treatment for all multimedia traffic is usually sufficient.

With simple priorities, the end points can mark multimedia traffic according to the VoIP traffic class. In the network, there is no call admission control, which takes QoS into account. Hence, network and traffic class dimensioning must consider the worst-case voice and multimedia traffic load. Otherwise, overload situations can lead to QoS degradation for all voice and multimedia calls. This approach is feasible and cost effective in LANs, including the scenarios with high data traffic load. It is not suitable for extended scenarios with bottlenecks in WAN links.

5.2. RSVP Signaling and Integrated Services

A well-known QoS signaling protocol is the IETF resource reservation protocol RSVP [RSVP]. For RSVP, we mostly focus on the issue of QoS signaling with RSVP and not on per-flow bandwidth guarantees possible with IntServ. Although this protocol standard exists for several years, it is not widely used and has scalability problems. These are due to the overhead of per-flow traffic policing and queuing needed at each hop.

RSVP is a hop-by-hop signaling protocol, which follows the end-to-end data path. For RSVP, the call setup time is another critical aspect for its usage in VoIP systems. The following figure illustrates a possible H.323 fast connect procedure which is complemented by RSVP signaling. The dotted messages (RSVP path, resv and conf) are possible as soon as the details of the payload flows (port number, codec) are known. It is quite evident that the session setup for the two RSVP flows impacts the call setup time since the proceeding of the H.323 signaling is pending until the reservation succeeds. Every router along the path must process the RSVP messages.

Since the RSVP approach adds an extra end-to-end signaling procedure, this results in following main disadvantages:

- RSVP signaling adds a significant extra delay for call setup time (currently approx .3 to 1 sec.). This is mainly due to RSVP processing in the routers and an extra round trip time for signaling, which cannot be fully interleaved with H.323 signaling.
- End-to-end network services are only possible if every router along the data-path supports RSVP.
- RSVP is needed in all end systems, which impedes incremental installation.
- The known scalability problems with routers with respect to the number of RSVP sessions can be a limitation and can lead to performance problems with routers [RSVP-appl].
- Rerouting mechanisms are too slow for VoIP requirements. Since rerouting and re-establishing of reservations on new routes is triggered via timeouts, this is too slow for the usage with VoIP.
- Synchronization of H.323 with RSVP is needed, as discussed below.

Figure 4 shows a possible Call setup with RSVP and H.323 fast connect. Note that the RSVP messages must be processed at each router, which is not shown here. The figure shows a possible interleaving of the RSVP messages with H.323 (E.g. Call proceeding can be interleaved with the first RSVP setup). The earliest points to initiate the RSVP reservations are as follows:

- The receiver-initiated RSVP signaling for stream to B is possible as soon as the H.323 open logical channel (codec: B -> A, port numbers A_rec, B_send) info is known, i.e. after the fast start setup message is received at B and B has been granted admission by the gatekeeper.

Sender-initiated RSVP signaling for stream to B is possible as soon as OLC (codec: A -> B, port numbers B_rec, A_send) info known, e.g. after the Alerting message is received by A.
Alerting should only take place as soon as both RSVP sessions have succeeded. Otherwise, the behaviour of the system is very annoying for the user.

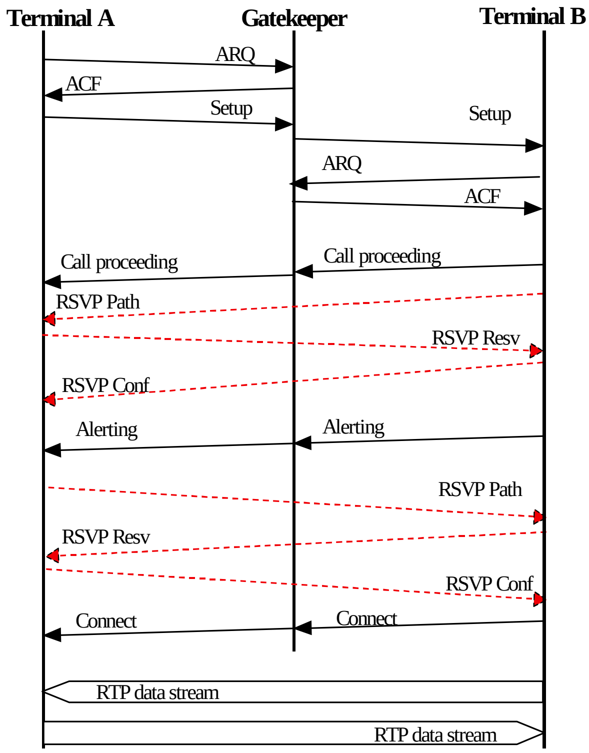

Figure 4: Possible Call Setup with RSVP and H.323 Fast Connect

The problems of combining RSVP signaling with H.323 can be summarized as follows:

- The first problem is that IP addresses, port numbers and used codec (bandwidth) must be known for initiating an RSVP reservation.
- Both end systems must be aware that RSVP signaling is used. This is achieved by setting the „qos“ bit in the H.323 open logical channel (OLC) info (see H.323 implementer’s guide)
- Synchronization of RSVP signaling (for both directions) and H.323 alerting is needed and actually discussed in the Study Group 16 of the ITU-T. The alerting and the alerting message can only take place if the RSVP reservations have succeeded. This requires introducing another end-to-end signaling message before the alerting message.

6. A SCALABLE RESOURCE MANAGEMENT ARCHITECTURE

6.1. Overview

The main idea of the resource management architecture is that each network domain (RM domain) has its local resource manager (RM). A RM manages the network resources (buffer, bandwidth) in switches/routers within its domain. The main task of a resource manager is the processing of connection admission control for their local network and to guarantee sufficient network resources for the granted reservations. These tasks are separated from the GK since they require knowledge of the network topology and resources in the RM's network domain. In contrast to this, the area of control of the GK (H.323 zone) by definition has no direct relation to the physical or logical structure of the underlying network.

Per flow signaling takes only place between the gatekeepers and one or a small number of resource managers along the data path. Therefore, there is no per-flow signaling among routers, which could lead to scalability problems. In addition, the RM may also support reservation in advance in order to avoid call blocking for e.g. important conferences. From the signaling point of view there is no significant difference between normal reservation and reservation in advance.

In larger networks with several RM domains, all resource managers on the data path of the H.323 realtime traffic must be asked for call admission. Any method can be used for local resource management. A RM domain may consist of a single router, a layer 2 access network, an OSPF routing area, or even an autonomous system.

Dividing a large network into RM domains might be necessary to obtain a scalable RM architecture, since smaller domains require less monitoring and administration effort. RMs need to store less topology, routing and call/reservation state information. This together with lower arrival rates of reservation requests reduces the performance requirements of RMs and allows smaller databases, faster table lookups and faster reaction times.

However, a larger number of RM domains increases the average RM domain hop count for H.323-calls, which means that more RMs are involved in a distributed end-to-end reservation setup process. This slows down the speed of the reservation setup procedure and may lead to non-acceptable call setup delays. When designing RM domain topologies this trade off between RM server complexity and call setup times has to be taken into account.

With respect to the H.323 architecture, a Gatekeeper (GK) is uniquely assigned to a RM, while a RM may administer the resources of several GK-zones. Therefore, a network administrator can design GK zones and RM domains independent of each other.

In our VoIP scenario, the gatekeeper signals reservation requests to a resource manager, which keeps track of the network resources in its domain and admits/rejects the call. To be able to do this, the GK has to be extended by a new interface to the resource management system. The reservation and release of resources are triggered by the H.225.0 RAS (Registration, Admission and Status) signaling. This is illustrated in Figure 5.

The resource management relies on a dedicated traffic class (DiffServ and possibly 802.1D), which is used exclusively for H.323 multimedia traffic. This class is assigned a maximum amount of bandwidth and buffer space on each network link. The RM controls the access to these resources exclusively by performing a call admission control mechanism.

In the following, we briefly describe the interworking between the GK and the RM system. Before a multimedia call starts, the terminal (H.323 client) sends an admission request ARQ (H.225.0) to its GK. In reaction to this, a reservation request to the RM system is made by the GK. This may exceed bandwidth limitations on individual links along the data path through the network. The resource management system blocks the reservation request if at least one violation occurs. In this case, the GK is informed about the lack of network resources by the RM system and responds to the admission request of the terminal with an admission reject ARJ (H.225.0 RAS). As a consequence, the terminal is not allowed to start call signaling. In our scenario the GK is able to monitor call state information, since call signaling goes through the gatekeeper (GK routed call signaling). So correct user behavior can be supervised by the GK.

In addition to that, a GK has the opportunity to act as a terminal proxy and release resources if e.g. a fatal error occurs on the H.323 client machine.

The resource manager approach has the following main advantages:

- Scalability with respect to large networks
- Flexibility in design of RM domains
- Heterogeneous QoS networks are supported.
- Simple introduction of QoS services as only one new network component is needed per domain. No changes in the end systems are needed.
- Reservation in advance

In the following, we show how the resource manager concept can fulfill the other requirements stated in Section 1:

- We show two options for retrieving the network topology and resources, which is essential for automatic configuration of the

RM. It also important for simple management that the RM can detect changes in the network.

- We introduce a RM discovery procedure
- We present a novel protocol for inter-RM signaling for H.323 call setup, which leads to minimal call setup delay.

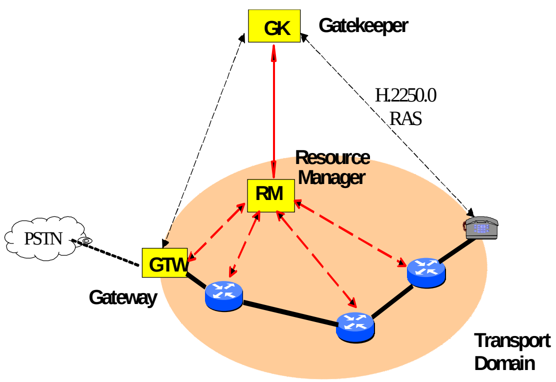

Figure 5: Resource Management Architecture

Note that topology discovery is independent of the actual mechanism used for QoS; the path of the multimedia data flow through the network is essential in almost any approach.

6.2. Topology Discovery and Resource Management

In larger networks automatic configuration is indispensable for the RM approach. For automatic re-configuration, interaction with the network devices is needed. The RM has to retrieve information about the network, i.e. topology, routes, link capacity and link configuration (max. bandwidth reserved for multimedia traffic), and must be notified in case of network changes.

The tasks include the following:

- Automatic discovery of the network topology and traffic routes.
- Adaptation in case of changes in topology or routes.
- Discovery of the link bandwidth and configuration for all (relevant) links.

For these tasks of the RM, several solutions are possible, depending on the target networks.

- The layer-3 topology and the traffic routes can be obtained by some routing protocols, e.g. OSPF. If such a protocol is available, this is a natural choice for topology information.
- The RM can obtain the device configuration via SNMP protocol by querying MIBs at each individual network device. However, the RM has to contact every network element and has to handle proprietary MIBs. This is only recommended if other options are not applicable.
- Policy servers can provide information about network devices and network changes. The policy system adds an extra layer of abstraction and will have to handle the wide variety of enterprise network devices. This is a desirable option, but standards are still lacking for the interface to the RM.
- Dedicated software tools to infer topology and routes. Some common management tools like HP's OpenView and IBM's Tivoli employ SNMP-based algorithms. Others use only basic IP-primitives (e.g. ping and traceroute). All of them discover topology statically, but do not handle network changes.

Regarding these options, the RM needs an open architecture to adapt to various network scenarios. Technically, the preferred solution for topology discovery is interaction with the routing protocol, as discussed above. For the link bandwidth discovery, SNMP can be used to access the appropriate (possibly proprietary) MIBs on the network devices. The most convenient solution would be to retrieve this information from a policy system, if available.

Recall that Layer-2 topology (including a possible 802.1 spanning tree construction) is not considered here, since it is assumed that bandwidth is sufficient in layer 2 networks.

It is assumed, that a network administrator sets link configuration (max. bandwidth reserved for multimedia traffic) statically. In overload situations (high blocking rates) it is conceivable that a resource manager has the ability to change the actual configuration of dedicated links e.g. via policy servers and policy enforcement points using the COPS protocol [COPS].

6.3. RM-Discovery

When initializing the RM system, three discovery procedures have to be performed. First, a gatekeeper has to find the RM, which is responsible for resource management in its zone. Each gatekeeper registers with only one RM while one RM may be in charge of several GKs. A possible discovery procedure may be similar to the GK discovery procedure standardized by the ITU (H.323).
Second, a RM has to receive knowledge of the scope of its domain e.g. number and location of the resources to administer. Thereby it may be assisted by a policy server, which provides the respective rules. These rules may consist of a simple set of routers or an IP network address and the corresponding network mask.

Third, in case of multiple RM domains every RM has to discover its counterparts of all neighboring domains. This is necessary to contact the appropriate RM whenever an inter-domain call setup has to be performed. Therefore, all RMs in a network communicate with each other via IP broadcast or multicast address and well-known port numbers. They are exchanging their IP addresses and information about the scope of their domains e.g. the IP addresses of border routers between different domains.

6.4. Inter Resource Manager Signaling

During a setup process all resource managers on the data path of the multimedia traffic must be asked for call admission. The first RM of a multi-domain reservation is informed about the destination address of the call and has to investigate whether the remote party is in its domain or not. If not, it has to map the destination IP address onto the IP-address of the subsequent RM along the reservation path.

After the discovery of network topology, the RM has knowledge of all routers within an area e.g. its RM domain, which may be its OSPF routing area and the dedicated routing information. With it a RM also knows the border routers to neighboring RM domains.

In the case of multiple domains, resource reservation has to proceed over several RMs. After the RM-discovery a RM knows which routers are the border routers to neighboring domains as well as which are the respective neighboring RMs. With the destination address of a reservation request, a RM is able to determine the routers along the data path within its domain. The border router along this path can be associated with the IP address of a neighboring RM.

A sample scenario of an H.323 call extending over three RM domains is sketched in Figure 6. The resource reservation layer (RMs) and the gatekeeper layer perform different signaling tasks, which have to be coordinated.

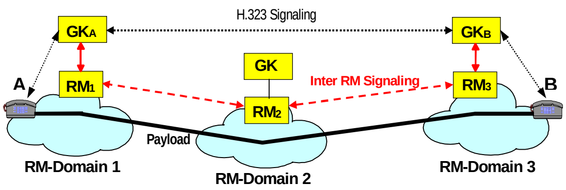

Figure 6: Multiple Resource Manager Scenario

Figure 7 shows a suggested H.323 fast connect signaling [H.323] between the Terminal-A, GK-A, GK-B, Terminal-B and the appropriate reservation signaling procedure between the GK-A respective GK-B and the affected three RMs. H.323 messages are drawn in bold lines, while our new messages are drawn in dotted lines. First, the Terminal-A makes an ARQ to its GK-A. As described above, GK-A is registered with RM-1 and initiates a reservation setup via RM-1. This reservation request message contains the destination address as well as QoS- and traffic- parameters e.g. complex token bucket parameters, which are specified in [TB_TS]. The QoS-parameters may be an upper end-to-end delay bound and information whether hard or soft QoS-guarantee is required by the application. The first one can be used for QoS-dependant call routing and is required for multi-domain reservations, the second one for choosing the appropriate admission control algorithm.

The codec used is not fixed until the H.323 capability exchange has taken place. Thus, the exact traffic parameters are not known, yet. Therefore, the reservation of resources has to be established in two steps. First, a preliminary bi-directional reservation in form of a worst-case reservation has to be made. A worst-case reservation means that the client calculates the aggregated bandwidth for the codecs with the highest bandwidth requirements. To make this preliminary reservation as efficient as possible, the RAS signaling has to be enhanced with more detailed information about the media streams. Later, when the exact call settings are settled, excessive resources are to be released in a second step (Res_Upd_Req/Conf). This is e.g. the case when the remote terminal has received the H.323 call setup message and makes an admission request to its GK.

After receiving a reservation setup request, RM-1 recognizes that the remote party does not lie in its domain. RM-1 determines the route through its domain and performs admission control on all or only on dedicated links. If resources are available domain wide, RM-1 grants admission for the call to Terminal-A by sending a reservation confirmation to the GK, which results in an admission confirm ACF to Terminal-A. Terminal-A can start H.323 call signaling now.

In parallel, RM-1 determines the border router, looks for the address of the RM-2 residing in the neighboring domain and forwards the reservation request to it.

The procedure is repeated by the intermediate RM-2. When receiving the reservation request, RM-3 finds out that it is the terminating RM. It performs admission control only if the reservation in the intermediate RM-Domains were successful. RM-3 stores the reservation together with the H.323 call context and waits for GK-B to refer to the existing reservation.

When the Setup message arrives at Terminal-B, Terminal-B can select one of the codecs offered by Terminal-A. Now, Terminal-B has the knowledge of the complete call information e.g. IP-addresses, port-numbers, H.232 Call-ID, codec, traffic- and QoS-parameters. When receiving the ARQ, the GK-B sends a reservation request to the RM-3. If the reservation was successful before, RM-3

confirms the reservation request to GK-B. After running e.g. an authorization procedure, GK-B may grant access to its zone and reply with an ACF without waiting for a Res_Upd_Conf-message from the RM-system.
If the actual selection of connections and codecs lead to a reduction of resources, RM-3 signals this back to RM-2 and RM-1 by using a reservation update message. This goes in parallel to the H.323 call signaling.

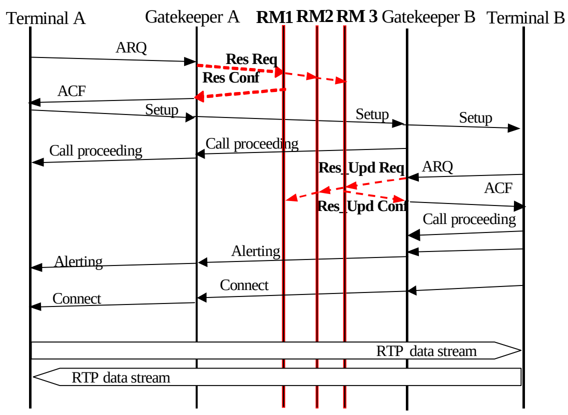

Figure 7: Call-setup with two RMs and two GKs

The main motivation of this approach is to provide end-to-end reservation as early as possible and to reduce the call setup delay by performing resource reservation signaling and H.323 call signaling in parallel. Thus, an initiating Gatekeeper can request a preliminary bidirectional end-to-end connection admission from its RM at the beginning of the call signaling process. The remote terminal does not start ringing before the end-to-end reservation was established successfully. The parallel resource reservation signaling is essential to minimize the reservation setup delay.

6.5. Automatic (Re-)Configuration using Routing Protocols

In larger IP networks, routing protocols are used to compute the traffic routes, particularly in the case of changes. It is possible to use the information computed by routing protocols for automatic configuration and adaptation to network changes. This is possible with advanced routing protocols such as IS-IS or OSPF. These protocols run in a distributed fashion on all routers and compute new, consistent traffic routes.
The main advantage of these link-state protocols over simpler protocols like RIP is that every router has a complete image of the local domain (autonomous system in routing terminology).
OSPF is the protocol recommended by the IETF standards and offers interoperability in comparison to other proprietary protocols.

With routing protocols, link failures can often be handled by routing the traffic over alternative routes. In case of a rerouting, the RM has to be informed about the network changes. This can be done by the routing protocol itself. After new stable routes are established by the routing protocol, the RM has to update its routing table or run topology discovery procedure in order to adapt reservations of still existing H.323 connections over the affected links.
Note that IP routing protocols only consider layer 3 links. Layer-2 devices and reconfiguration (802.1 spanning tree construction) are not considered, since it is assumed that bandwidth is sufficient in layer 2 networks.
These features may require a call release mechanism for the RM, i.e. to request call release from gatekeepers in case of network overload due to network changes.

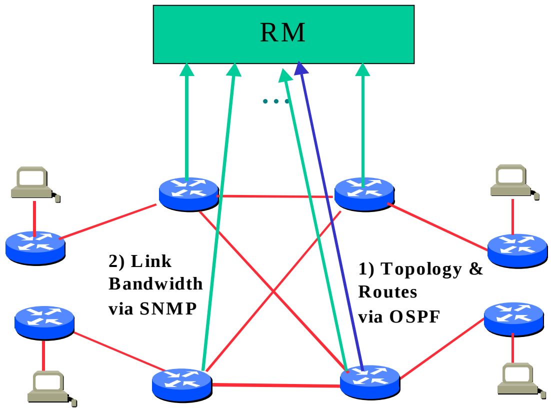

Figure 8: Example for RM with self-configuration and dynamic network adaptation

An example solution with control of all network links (see [Telia-BB] for more details) is shown in Figure 8. The main procedure is as follows:

In the startup phase the RM performs the following:

1) The RM is registered as an OSPF router and hence receives topology and routes from other routers.
2) With the topology information, the RM can request the link bandwidth via SNMP from all routers. A standard for this is in discussion at the IETF [DS-MIIB], but proprietary MIBs might be needed.

With this information, the RM can maintain a complete image of the network and its resource usage. In case of a reservation request for a flow by the gatekeeper, the RM has to determine the route and to check resources at each link.

The RM continuously has to watch topology changes, which are announced via the OSPF routing protocol. In the case of rerouting, the RM has to update its topology and to re-compute the resource usage. If needed, it has to drop calls.

7. CONCLUSIONS

We have presented a complete architecture for scalable QoS for VoIP and have detailed the main new building blocks, including new protocols.
We have shown that the RM concept fits nicely with the H.323 architecture and can be introduced incrementally. In addition, we discussed ways for automatic configuration of the resource management system.
The main advantages of our architecture are:

- Provision of end-to-end QoS in IP-networks.
- Low installation effort.
- Scalability with respect to large networks.
- Automatic discovery and reconfiguration.
- Only a few network instances are involved in the admission process. Therefore the call setup delay can be minimized.
- Call setup delay is further reduced by performing reservation signaling for inter-RM domain calls in parallel to H.323 signaling
- Central QoS control for H.323 connections for monitoring and backend service support.
- Only one new component needed in the network.
- Reservations in advance are supported.

REFERENCES

[I2-BB] Internet2 QBone Bandwidth Broker Advisory Council, http://www.internet2.edu/qos/qbone/QBBAC.shtml

[RSVP] Resource reservation protocol (RSVP)—Version 1, functional specification," IETF RFC 2205, 1997.

[DiffServ] An Architecture for Differentiated Services, IETF RFC 2475, 1998, ftp://venera.isi.edu/in-notes/rfc2475.txt

[802.1D] 802.1D-1998 Media Access Control Bridges, IEEE standard

[MS-QoS] Windows 2000 Quality of Service Technical White Paper, Sept. 03, 1999, http://www.microsoft.com/windows2000/library/howitworks/communications/trafficmgmt/qosover.asp

[Tiphon-Qos] TR 101 329 V2.1.1 (1999-06) - General aspects of Quality of Service (QoS), available from: http://www.etsi.org

[H.323] ITU-T Rec. H.323, "Packet-Based Multimedia Communications Systems," Geneva, Switzerland, Jan. 1998; ttp://www.itu.int/itudoc/itu-t/rec/h (link to substandards H.450.x, H.235, etc.).

[COPS] A. Smith, D. Partain, J. Seligson: "Common Open Policy Service Protocol", Internet Draft, May 2000; http://www.ietf.org/internet-drafts/draft-ietf-rap-cops-client-mib-03.txt

[TB-TS] J. Glasmann, A. Riedl, M. Czermin: "Estimation of Token Bucket Parameters for Videoconferencing Systems in Corporate Networks" to be presented at SOFTCOM'00

[Telia-BB] Olov Schelén, Andreas Nilsson, Joakim Norrgård, Stephen Pink: Performance of QoS Agents for Provisioning Network Resources. In Proceedings of IFIP Seventh International Workshop on Quality of Service (IWQoS'99), London, UK, June 1999. available from http://www.cdt.luth.se/~olov/publications/index.html