

Motivation of an end-to-end regulation of bandwidth in intra-networks: The ROBIN concept

M. Frank, P. Martini

University of Bonn, Institute of Computer Science IV

Römerstraße 164, D-53117 Bonn, Germany

Tel.: +49-228-734550, Fax: +49-228-734571

E-mail: {matthew, martini}@cs.uni-bonn.de

Abstract

The variety and heterogeneity of different approaches for Quality of Service (QoS) support on different protocol layers result in the need of a solution with an end-to-end overview and regulation of data flows in a network. The ROBIN (**R**egulation **O**f **B**andwidth in **I**ntra-**N**etworks) concept introduced in this paper is such a mechanism which controls the bandwidth used by data flows in heterogeneous networks in case of network congestion. In a project at the University of Bonn, the concept of ROBIN is designed, a prototype is implemented for laboratory measurements and a performance analysis by simulation is carried out. The paper presents an overview of the concept and reports on the state of practical work on ROBIN.

Keywords

Bandwidth regulation, congestion control, QoS support, relative quality of service, rate-based flow control, transport layer extension.

The original version of this chapter was revised: The copyright line was incorrect. This has been corrected. The Erratum to this chapter is available at DOI: [10.1007/978-0-387-35388-3_42](https://doi.org/10.1007/978-0-387-35388-3_42)

1 INTRODUCTION

“The race to run time-sensitive traffic over packet-based LANs has spawned a host of approaches ... but which will emerge as the technology of choice is anybody’s guess” (Roberts, 1996). This quotation from the magazine “Data Communications International” both mentions the highly relevant issue of transmission of real-time traffic over packet-based LANs and describes the state of the art on finding an effective solution for this problem. Although this statement is already more than one year old, it has not lost a bit of relevance.

The situation has been caused by an evolutionary development of global and local networking in recent years. The technical capabilities of networks are improving rapidly, in particular the available bandwidth is rising especially in local networks (e.g. Fast Ethernet is available and becoming cheaper, Gigabit Ethernet is on its way to the desktop). Additionally, the requirements of applications change and a variety of new applications appeared on the scene bringing requirements that were not considered when classical communication protocols were designed.

Today, there are several new real-time applications that have a demand for a minimum guaranteed bandwidth (e.g. real-time transmission of audio- and video-data) and additionally may have a low upper delay bound (e.g. in a live video conference). Nevertheless, classical communication forms (like filetransfer ftp, Hypertext Transport Protocol http, News, Electronic Mail, Network File System NFS, remote login, etc.) are still alive and growing in terms of data volume. This results in a situation of competition of data flows of different applications with different requirements in a network with limited resources.

In many places, researchers and developers of networking hardware and software work on this issue. The suggestions vary from new approaches in network technologies, in particular strategies in medium access, to improvements, design or redesign of upper layer communication protocols. The efficient support of QoS (Quality of Service) properties in a heterogeneous network requires an excellent cooperation or matching of concepts, which is a very crucial point with the variety of existing or proposed approaches addressing several protocol layers.

This paper introduces an extension of the transport layer (**ROBIN - Regulation Of Bandwidth in Intra-Networks**), which improves the effectiveness of QoS support based on the end-to-end knowledge about application data flows and a bandwidth regulation at the sending side. ROBIN has been designed for campus or intra-networks like a typical autonomous system e.g. at an university, research institute or larger company. The goal of ROBIN is to establish and maintain an (almost) optimum overall provision of the available network bandwidth to the application data flows which at the same time takes care of individual requirements of different applications. The activities concerning the ROBIN concept, realisation and evaluation are covered by a research project at the University of Bonn.

The paper is organised as follows: Section 2 presents a short review of existing and future approaches to QoS support and leads to a motivation for an end-to-end regulation of bandwidth to support different application requirements in a heterogeneous networking environment. Section 3 introduces our end-to-end

ROBIN approach with its characteristics and functionality. Section 4 presents the state of our work on ROBIN, both describing the activities on a prototype implementation and a performance analysis by simulation of ROBIN. Finally, section 5 concludes the paper and gives an outline on future work related to ROBIN.

2 MOTIVATION FOR END-TO-END REGULATION

As mentioned above, there are different suggestions from researchers and developers working on the issue of integrated QoS support. Figure 1 shows a selection of different classical and new QoS approaches or protocols to different layers of the OSI reference model.

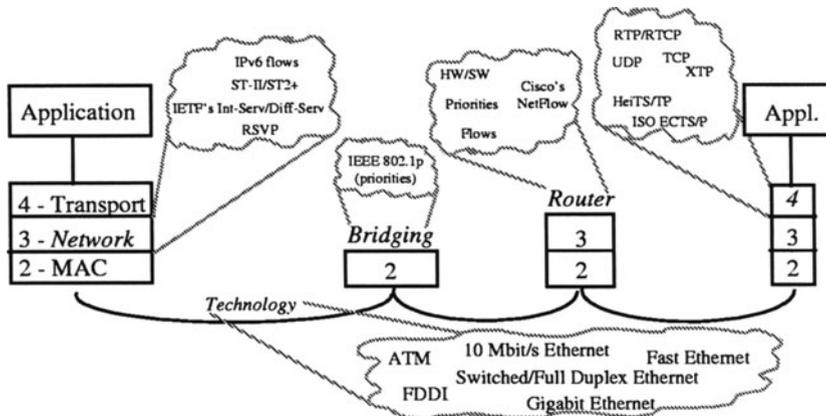


Figure 1 Overview of mechanisms/protocols with or without QoS support.

In both the transport-oriented layers of the OSI-RM (transport and network layer) and the data link layer (in particular the medium access) there are several approaches to support communication services with specific QoS requirements. In this paper we cannot discuss all of these in detail. Instead we emphasise important characteristics of the approaches and the variety of these. As important representatives of networking technologies we point at ATM (Asynchronous Transfer Mode) and Ethernet. With its connection-oriented concept ATM is well-suited for the support and guarantee of different QoS properties for different communication flows. In contrast to this, Ethernet in its basic form does not allow for any service guarantees. The development of Ethernet from classical 10 Mbit/s to 100 Mbit/s Fast Ethernet, Switched Ethernet and Full-Duplex Ethernet up to future Gigabit Ethernet includes possibilities for future QoS support. However, there is still much work to be done in this direction.

The desire to support service integration in the Internet has strongly influenced the Internet Engineering Task Force IETF in the design of the new version of the

Internet Protocol IPv6. The real use of new functions is still vague and criticism on IPv6 arises (Hutchison, 1997: "... the semantics of communication offered by IPv6 is still the well-known best-effort ..." and "...we may suspect that the support of QoS has been deliberately evicted of IPv6's design"). Additional protocols like the Resource ReServation Protocol RSVP (cf. Braden, 1997) or ST-II (cf. Delgrossi, 1996) are required to make guaranteed QoS a reality in the Internet. An important step towards RSVP is the support of resource reservation within the network connecting elements (like bridges or router) and within the networks themselves. A comprehensive overview of reservation protocols for multimedia communication may be found in (Delgrossi, 1996).

At the end of last year a new working group (Differential Services for the Internet – Diff-Serv) of the Internet Engineering Task Force (IETF) split from the Integrated Services (Int-Serv, cf. Baker, 1996) working group. The differential or differentiated services (both terms are in use) intend to introduce a service discrimination in the Internet without the need for per-flow state and signalling at every hop (cf. Nichols, 1998) and without requiring admission control for flows using this service. Several proposals for service categorisation and frameworks for deployment of these services are under discussion (e.g. Kalyanaraman (1998) or Nichols, 1997). Without having direct relation to activities of Int-Serv and Diff-Serv working groups, Braden (1998) indicates some recommendations on active queue management and scheduling algorithms for congestion avoidance of best-effort traffic in the Internet. However, no consensus solution is existing to control congestion caused by unresponsive or not sufficiently responsive flows (not adapting to indications of congestion like TCP does). Also Floyd (1997) addresses these problems and also emphasises the negative impact of non-controlled best-effort traffic on the Internet in terms of "extreme unfairness ... to the potential for congestion collapse". They suggest per-flow scheduling and router support for end-to-end congestion control.

Independent of the underlying transport and network layers, the Real-time Transport Protocol RTP (cf. Schulzrinne, 1997) has been designed by the IETF to support real-time communication. RTP and RTCP (RTP Control Protocol) provide identification of payload type, sequence numbers, timestamps and report feedback about the network status allowing the application to do the synchronisation and media scaling. However, there is no negotiation of QoS parameters prior to the data transmission and in case of congestion no active actions of RTP are performed. Instead, the application is expected to take care of these problems. Applications typically run RTP on top of UDP, which offers no functions of QoS support that RTP could use. Furthermore, with RTP/RTCP it is not possible to estimate the impact of other (possibly malicious) protocols operating in the same network.

The title "Internet Protocol Quality of Service Problem Statement" of the Internet Draft (Bradner, 1997) already points at the problem about QoS in IP: Bradner (1997) lists a number of essential but extensive requirements for effective support of QoS in the Internet Protocol. Some of the approaches are questionable already because of intrinsic problems. The situation becomes much worse in larger campus

or intra-networks consisting of up to several hundreds of hosts, as it is characteristic for universities, research institutes or larger companies. Typically, in such environments we find heterogeneous interworking of different network technologies and topologies, connected either by hosts operating as software routers or by hardware components (bridges, routers, switches and hubs). We expect that this situation will remain typical for such networks since large local and campus networks usually have heterogeneous structures resulting from incremental growth and replacements in specific parts of the network. A complete move to a homogeneous technology would probably bear too much cost to afford.

The provision of QoS properties in such a heterogeneous network requires cooperation or matching of concepts, which is a very crucial point with the variety of existing or proposed approaches covering several protocol layers. These approaches are far from being intuitively combinable. The argument of heterogeneity, in particular on layers 2 and 3 of the OSI RM, leads to the conclusion, that the control mechanism to support QoS should be moved to or added above the network layer with an end-to-end knowledge of different transport protocols, their active data flows and the current network status. The traditional end-to-end view of transport protocols is not enough for this, because the congestion reaction scheme (e.g. as with TCP) does not know details (e.g. bandwidth requirements or actual bandwidth use) of other communication flows, neither of the same protocol nor of other transport protocols.

The idea to make use of global or end-to-end knowledge on active data flows and network state for an integration of QoS support has already been considered at the University of Virginia: TReg - Transport Regulation (cf. Liebeherr, 1995) enhances applications on top of the TCP/UDP protocols to control the application data flows. With the TReg concept of Liebeherr (1995), all application processes using the service of TCP or UDP have to be enhanced by adding TReg components (so-called "TReg-Stub" and "TReg-Daemon"). These components have a global view on all active TReg flows in the network. They know about the utilisation of network resources and are directly able to control the source of each flow within the application process before having access to the transport protocols. However, the TReg approach suffers from the possibly malicious interference of other applications and transport protocols that do not implement the TReg mechanism. A closely related work (cf. Akyildiz, 1996) implements the bandwidth regulation within the network layer, which obviously implies an adaptation of the network protocols including the inter networking units (possibly difficult to realise in bridges or hardware routers).

The Lancaster QoS-Architecture (QoS-A, cf. Campbell, 1994) defines an "integrated and coherent framework that incorporates QoS interfaces, management and mechanisms across all architectural layers" (Campbell, 1994). QoS-A has a focus on an integration of QoS properties in both end-systems and the network with a support of quantitative end-to-end QoS. The ROBIN approach with its end-to-end knowledge on active data flows very well matches the architectural model of QoS-A to represent one possible option of QoS support on the transport layer. This way, the multi-layered approach of QoS-A may help ROBIN in achieving its

goals. However, the ROBIN approach is free from depending on a quantitative specification of QoS properties prior to the transmission and can already work without support from the network and its hardware/software elements.

The working group ISSLL (Integrated Services over Specific Link Layers) of the IETF presently discusses a framework for providing integrated services on OSI-RM layer 2 over shared and switched IEEE 802 LAN network technologies (cf. Ghanwani, 1998). The goal is to control the resources of network segments and bridges/switches at the data-link layer with a "Subnet Bandwidth Manager" (SBM, cf. Yavatkar, 1998) to support integrated services and quality of service provision. The approach of SBM is designed to work without dependencies on particular upper layer protocols like RSVP or IP. Within the framework and architecture of SBM on OSI layer 2 and our ROBIN concept on OSI layer 4 there are several related aspects: Ghanwani (1998) introduces several possibilities of implementing the bandwidth manager (centralised, distributed or semi-distributed), which has to manage the resources of the entire subnet. Aspects like soft state with periodical refresh and good scalability are essential requirements. As with ROBIN, the flow separation and scheduling is an important issue. The documents of the ISSLL group on SBM are Internet Drafts documenting the state of work in progress and there is no information on practical experience available yet. However, the progress of this work has to be studied carefully to make use of solutions on common issues with our ROBIN work.

The next section presents our approach to extend the transport layer to support QoS properties by bandwidth regulation of different transport protocols and their data flows.

3 THE END-TO-END CONCEPT OF ROBIN

ROBIN is operating as an extension of the protocol stack between transport and network layer protocols. ROBIN regulates the bandwidth by a rate control at the source of individual data flows. Due to the end-to-end operation of ROBIN, no active cooperation of network protocols and inter-mediate network nodes is required. The ROBIN mechanism is not connection-oriented in the sense that a negotiation of QoS parameters for each flow must be done prior to the data transmission. Instead, the goal of ROBIN is to establish and maintain a nearly optimum overall utilisation of the available network bandwidth by the application data flows which at the same time takes care of individual requirements of different applications.

Basic ideas on the improvement of QoS support with ROBIN are as follows:

- In a campus network that does not support any QoS functionalities at network or lower layer protocols, ROBIN is able to considerably improve this situation by providing a relative QoS based on the different requirements of different applications.
- ROBIN is able to control the sending bandwidth of flows unresponsive to congestion and thus is able to improve the fairness of network bandwidth distribution between all active flows.

- ROBIN is able to make use of or cooperate with different QoS approaches on layers 2 and 3 with a mutual benefit and increase of effectiveness: E.g. reservation protocols or priority schemes may support ROBIN in providing its end-to-end QoS and in turn ROBIN will extend the efficiency of other schemes with its end-to-end knowledge of the data flows in a campus network.

3.1 Basic concept of ROBIN bandwidth regulation

The basic concept of ROBIN bandwidth regulation has been published in (Frank, 1997). The feedback and experiences gained within the first steps of implementing the ROBIN concept as a prototype and for simulation study have already led to some improvements which are included in the following description of the ROBIN concept.

ROBIN is operating as an extension of the protocol stack between transport and network layer protocols in the operating system of each station of the network. Each instance of ROBIN has knowledge about the complete network topology, the capacity of network elements and the individual bandwidth requirements of “own” flows leaving this station. This “own” information is distributed to the peer ROBIN instances within a frequent state exchange message. Within this, all ROBIN instances approximately have the same knowledge of the complete network state and active data flows. In case of overload of particular elements in the network ROBIN distributes the bottleneck bandwidth between the competing flows by regulating the sending bandwidth in an appropriate way. The overhead for state exchange and additional calculation limits the area of ROBIN regulation to campus or intra-networks like a typical autonomous system e.g. at a university.

With the calculation of the sending bandwidth of individual flows, ROBIN achieves a nearly optimum utilisation of the bandwidth available in the network for all active data flows across this network. All existing local data flows of transport protocols like UDP, TCP and special purpose protocols are considered. Therefore, ROBIN is able to include congestion unresponsive flows (like UDP or other transport protocols) in its bandwidth distribution. The available bandwidth of the network (in particular parts of it being a bottleneck) is shared according to the principle of relative QoS as used in (Liebeherr, 1995): Different applications are classified according to their requirements in terms of minimum bandwidth (e.g. for video) or relative priorities (e.g. video is more important than filetransfer). The available network bandwidth for each class is equally shared by all active communication flows of this class (e.g. all active video connections). However, the reserved share of the total bandwidth differs between classes according to their priority. The definition of classes, mapping of applications to classes and the reservation of bandwidth for each class in all network components is specified in advance by the system administrator.

The main advantage of sharing the bandwidth with this principle of relative QoS is that this approach satisfies the bandwidth needs better than a best-effort network service without additional knowledge about flow requirements and thus sharing the bottleneck bandwidth evenly between the competitors, if the latter is met at all.

E.g. the unfairness problem of TCP is well-known and several approaches have been proposed to improve these drawbacks (cf. the proposal of Satyavolu (1998) or related work cited there).

Where possible, ROBIN also supports the negotiation of a certain transmission bandwidth during connection set-up. Applications that are able to specify their bandwidth demand in advance (with similar parameters as for RSVP set-up), may reserve bandwidth from ROBIN. If the connection is accepted then these flows are given higher priority in the bandwidth allocation (class 0 flows, cf. Table1).

Table 1 Possible classification of applications for relative QoS

<i>Class</i>	<i>Share of bandwidth (%)</i>	<i>Applications type</i>
0	Absolute fraction of bandwidth	All applications that are able to negotiate desired bandwidth in advance
(1-4	Classes 1-4 share the remaining bandwidth)	
1	up to 50	Audio, video, multimedia streaming
2	up to 30	WWW, ftp
3	up to 15	E-mail, news, software distribution
4	up to 5	Telnet, rlogin

Table 1 shows an example of mapping of applications to classes and the reserved bandwidth share of a bottleneck in the network (here with a total of 5 classes). The reserved bandwidth shares may be selected individually for particular network segments, e.g. in network elements close to a video server the fraction of class 1 may be higher than in other parts of the network.

If 4 video flows (class 1) and 10 ftp-flows (class 2) have to share the bandwidth in a 10 Mbit/s Ethernet segment, with the principle of relative QoS each video flow gets 1,75 Mbit/s (70%/4) and each ftp-flow 0,3 Mbit/s (30%/10), where the unused reservation of classes 3 and 4 has been used by class 1. Without such a regulation by ROBIN, all flows had to share the available bandwidth equally (10 Mbit/s/14 = approx. 0,7 Mbit/s for each flow), which would lead to a considerable degradation of quality for the video flows.

Figure 2 depicts the components of a ROBIN instance and the information flow between these components. The functionality of each component is explained in the following.

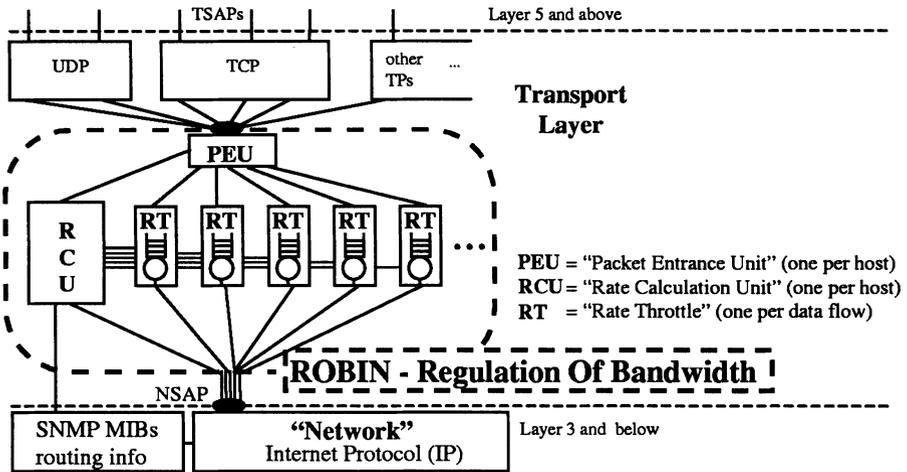


Figure 2 ROBIN architecture.

Packet entrance unit (PEU)

Each data segment passing from transport to ROBIN is handled by the PEU and according to its data flow identification is directed to the corresponding rate throttle (RT). Flows and their application classes are identified by source and destination IP addresses, port numbers and/or transport protocol identifier. This identification of data flows within a protocol below the transport layer is an adoption of ideas from the Integrated Services Architecture ISA of the IETF (cf. Baker, 1996).

If there is no RT existing for this flow (i.e. a new transmission started), the PEU generates a new RT, passes the segment to the new RT and also notifies the RCU about the new flow. Within the IP-like access to ROBIN via the PEU we achieve a transparency of ROBIN from the transport protocol's point of view. Correspondingly, the integration of ROBIN into an existing stack of protocols in the TCP/IP family is rather simple. Furthermore, with this approach the data flows of all transport protocols are controlled by ROBIN bandwidth regulation, even those of connection-less protocols like UDP.

Rate throttle (RT)

There is one rate throttle for each active data flow leaving the ROBIN instance. The main task of the RT is to measure the throughput of data sent from the application and transport protocol of the corresponding flow. If activated by the RCU, the RT limits the bandwidth available to this flow with a rate control according to the parameters provided by the RCU. Due to the automatic generation of the RT by the PEU (without any connection or flow set-up), the RT does not learn explicitly about the end of transmission for a flow. Instead, the RT is removed after a certain inactivity time-out.

The limitation of sending bandwidth within the RT is established by a rate control with parameters <rate> and <burst> (as with rate control in XTP, cf. XTP-Forum, 1995). This scheme is comparable to the classical token bucket algorithm as used for traffic shaping, where our sending RT is able to save up permission to send <burst> bytes without constraint. In case of suspended transmission, the RT may apply a blocking operation to give backpressure to the transport protocol, or may even work in non-blocking mode with queuing and scheduling in a fair queuing manner. Currently, different realisations of RT modes are under discussion and will be evaluated in further studies.

The actions of the rate throttle to regulate the bandwidth by delaying or in the worst case even discarding packets may cause an interference with mechanisms of transport protocols. These will be studied in detail to avoid a reduction of the efficiency of those protocols. Nevertheless, this interference may be used for a passive cooperation of ROBIN with transport mechanisms: e.g. TCP's congestion control is sensitive to round-trip delay and packet loss and may adapt its sending bandwidth to the available bandwidth coming from ROBIN regulation. Well-known approaches for adaptation of queue management like Random Early Detection (RED, cf. Braden, 1998 or Floyd, 1997) or new proposals for TCP support with per-flow queuing (cf. Suter, 1998) to improve performance of reactive end-to-end congestion control mechanisms will be considered for buffering in the rate throttles of ROBIN as well.

In addition, an active cooperation with transport protocols is useful, when the protocols are able to use information from ROBIN: The calculated sending rate from ROBIN may be used by the transport protocol to scale down its data flow accordingly or to pass this information to the application to react (e.g. similar to what RTP/RTCP does by its QoS monitoring).

Rate calculation unit (RCU)

The RCU is the most important component of ROBIN: It maintains a data structure representing the topology of the network, the bandwidth capacity of all network elements operating in the network, and the classification of application flows and per class bandwidth reservation for realising the relative QoS. Each RCU knows the details (application class, presently consumed bandwidth, bottleneck location) of the "own" flows leaving the host where it resides. It distributes this information to all peer RCUs in frequent intervals of time by a multicast or broadcast state message. By this, all RCUs have almost the same global knowledge of the number and properties of all active flows in the network and are able to determine bottlenecks and to limit the bandwidth of flows involved to avoid congestion.

The network topology is automatically detected by a tool working with SNMP (Simple Network Management Protocol) and internally represented as a graph. The routes of flows in the network are internally stored after using information from routing protocols. We assume a static routing in IP in use within the network under control of ROBIN. This assumption is no limitation, because static routing is very often used in campus like networks.

The algorithm to calculate the sending rate of the RTs for each flow works in two phases: In the first phase, the algorithm follows all routes of own flows from the source to the destination within the internal network graph. For each network element with an overload caused by own and foreign flows passing through, the bandwidth shares of all flows are calculated according to the principle of relative QoS. It also considers reserved bandwidth of empty classes, which is immediately distributed to classes with active flows. In the same step, the bottleneck of each flow is determined when following the routes of the flows and the sending rate is adapted to the bottleneck bandwidth. After this phase, there may be a situation in single nodes with several flows, where some of these flows have their bottleneck at this node and flows having a different bottleneck node. Thus, the latter do not use their calculated share completely and bandwidth is left unused. Additionally, some flows may not completely use their calculated share of the relative QoS principle without having a bottleneck at all. In both cases, the unused amount of bandwidth should be re-used by bottlenecked flows in these nodes.

Phase 2 tries to re-use possibly unused bandwidth in all bottleneck nodes of own flows. Unused bandwidth of particular flows is first distributed to bottlenecked flows of the same class and second to flows of other classes (with an order of reuse or priority from classes 1,2, ... down). With all flows belonging to the same class, the ROBIN regulation very well meets the requirements of MaxMin Fairness (cf. Simcoe, 1994). The bottleneck rate for a flow is achieved, unused bandwidth shares are divided equally among bottlenecked flows and after a change in network state, the sending rates are adapted to the new situation accordingly. But also the classification of flows according to relative QoS with different amounts of bandwidth reserved for each class introduces no unfairness and does not violate the MaxMin fairness.

3.2 Critical issues on ROBIN operation

Data flows to/from outside the intranet

The area of ROBIN regulation is limited to a campus network with limited geographical extension. Obviously, there are connections or data flows running to and from the network located outside the ROBIN intranet. The flows leaving the intranet are under control of ROBIN on their route from the source to the border of the network (e.g. the router or gateway to the WAN) and the sending rate may be limited to adapt to overload within the campus network.

At the sending station of flows entering the campus network no ROBIN instance is participating in the bandwidth regulation to control these flows. In most cases the connection of a campus network to a WAN (like the Internet) is of a rather small bandwidth compared to the bandwidth within the local network. Thus, a dedicated ROBIN instance at the "entrance" to the campus network could passively quantify the incoming external traffic and consider this amount within the calculations on the route from the entrance router/gateway to the destination subnetwork and host. Additionally (or in particular with a faster connection to the outside world), an active component could perform a traffic shaping of incoming

data flows. Commercial products are available for this purpose in the IP-protocol family (cf. "Packetshaper", Bruno, 1996), which are able to control incoming flows of particular applications, destination subnetworks or even individual hosts.

Scalability of ROBIN

The first version of the ROBIN concept was based on the assumption that each ROBIN instance distributes its information about own flows frequently to all other ROBIN instances, which register the data and have the knowledge of the complete network present all the time. Although ROBIN usage is limited to a local area, the overhead for distribution of state information and the calculations based on the complete knowledge may already limit the scalability to the order of typical intra-networks.

To improve the scalability of the ROBIN approach, we follow different solutions which may be combined. Up to now, the state exchange and re-calculation of sending bandwidth was done after fixed time intervals. An estimation of communication overhead with 2000 active data flows showed for the worst case (large state exchange packets, large overhead for lower layer protocols as IPv6 + FDDI, state of each flow sent in a separate packet) that an exchange interval of one second results in an additional network load of 1 Mbit/s. In classical 10 Mbit/s Ethernet segments, this is an overhead of 10% for realising the ROBIN concept. To ease this situation, a dynamic determination of points in time for distribution of state information and for re-calculation reduces the overhead. A compromise between "as little as possible" and "as much as necessary" has to be found: Only after relevant changes in the system (such as additional flows, flows leaving, change of parameters within pre-defined tolerance), rates will be re-calculated and a state packet will be sent.

An additional approach would be a hierarchy according to the principle "divide and conquer": The total network is divided into several ROBIN-islands with only local knowledge about island-internal topology and internal flows. The state information is exchanged on two levels: Firstly, within an island for all internal flows, secondly between islands only regarding the flows between islands, where one island does not know all details of other islands. This approach reduces overhead for state exchange and complexity of calculations and additionally allows for a connection of several remote ROBIN islands (e.g. two remote company network sites connected by a fast direct link).

These extensions of the basic concept are presently under preparation and will be covered by detailed studies of performance and scalability within a comparison of different versions of our concept.

4 IMPLEMENTATION OF ROBIN

Within the scope of the project the concept of ROBIN is designed, specified and implemented as a prototype. The implementation aims at verifying the functionality of ROBIN in a local network environment. Due to the fact that performance measurements giving sound results in a prototype system are only

possible with some limitations, a performance analysis with modelling and simulation of the ROBIN system will also be done. This simulation study will evaluate the impact of certain parameters on the ROBIN functionality and will be able to analyse the ROBIN performance in larger scenarios than possible with the prototype implementation.

The ROBIN prototype system has been specified in SDL (Specification and Description Language) and automatically implemented with our proprietary tool SDL2C (translating SDL specifications to C-code). The prototype is running on a UNIX platform. The modelling and program implementation for the simulation study is done with the help of our proprietary simulation tool OOSIM (Object-Oriented SIMulation Library), a library to support discrete, asynchronous, event-driven simulation. A detailed description of the simulation model for ROBIN studies has been presented in (Frank, 1998). For the experimentation with our concept of ROBIN we follow an integrated approach of specification and implementation for the prototype realisation and the simulation study of ROBIN. The interfaces to the network and transport protocols, the PEU and the RTs have specific implementations for the prototype and for simulation. The SDL specification with processes, internal signals and interface functionalities had to be transformed into the simulation model of PEU and RTs and to corresponding simulation events. The RCU, the data structures for network topology and flow characterisation and in particular the algorithm to calculate the bandwidth shares for each flow have been integrated directly in C and thus can be used for prototype implementation and object oriented simulation without further adaptation. This allows for a close mapping between prototype implementation and simulation activities and a simple exchange of adaptations to the concept between the two of these.

The objective of the prototype measurement presented here was to demonstrate the principle of relative QoS and the effectiveness of ROBIN to achieve this principle. Each flow belongs to a different class (1 .. 4) and generates a load of 60% of the bottleneck bandwidth (in an Ethernet network segment). Thus, with 2 or more flows active, there is a state of overload. Without any control mechanisms in the transport protocol (like with UDP) the flows are unresponsive to indications of congestion. The flows are starting at regular intervals of 15s. The reserved bandwidth for each class is equal to the example classification given in Table 1. Figure 3 shows the bandwidth shares of the four flows resulting from measurements on the prototype. The left graph depicts the shares of sending bandwidth calculated by ROBIN and the right graph shows the actual throughput of the flows (after a possible regulation), both displayed over the time.

With only flow 1 active (0 - 15 s), there is no rate limitation of ROBIN and there are no plots in the left graph yet. From 15 - 30 s, flows 1 and 2 are active which results in overload. The initial share is calculated from the bandwidth reserved for each class with active flows: Classes 3 and 4 have active flows and have a reservation of $15\% + 5\% = 20\%$ of the total bandwidth. The reservations of classes 1 and 2 are unused and are immediately distributed to classes 3 and 4: Class 3

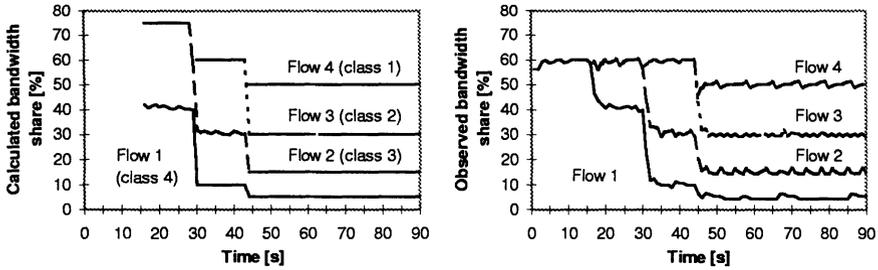


Figure 3 Prototype measurement: (a) calculated, (b) measured bandwidth shares.

receives $15/20 = 75\%$ and class 4 receives $5/20 = 25\%$ of the total bandwidth. The actual demand of flow 2 (of class 3) is only 60%, thus the 15% of unused bandwidth are reused by flow 1 (resulting in a regulation of $25\% + 15\% = 40\%$). Nevertheless, a rate limitation of flow 2 with a value of 75% is activated for safety reasons. From 30 - 45 s, the shares of the three active flows are $30/50 = 60\%$, $15/50 = 30\%$ and $5/50 = 10\%$ for classes 2, 3 and 4 respectively. After 45 s all flows are active and the calculated shares equal the class reservations as indicated in Table 1.

The relative QoS principle with ROBIN bandwidth regulation is clearly visible: The bottleneck bandwidth is purposely shared unequally between flows from different application classes. Also the dynamics of the rate adaptation is illustrated with starting flows: The plots of calculated rates achieve the new bandwidth share value instantaneously after a change in the network. In contrast to this, the measured bandwidth share reflects the impact of the ageing function used to smooth the measured throughput.

Due to performance limitations with the prototype implementation, it is not possible to make a fair comparison of performance in a scenario with and without ROBIN operation. Thus, this comparison has been performed by the simulation study: Figure 4 shows the measured bandwidth for the previous scenario with and without ROBIN bandwidth regulation.

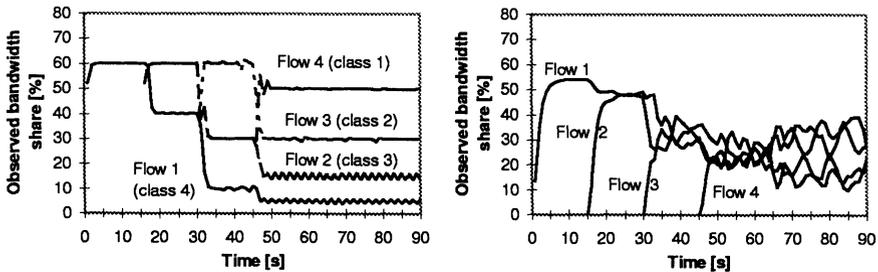


Figure 4 Simulation result: Bandwidth shares (a) with, (b) without ROBIN.

Figure 4 (a) is almost identical to Figure 3 (b), which demonstrates the correct functionality of ROBIN in both the prototype measurement and the simulation study. Without ROBIN operation, the measured throughput on the receiving sides oscillates rather heavily and there is even an unfairness between the four flows visible after about 65s. Again it has to be noted that we used no transport layer mechanisms for congestion control to exclude an interference with ROBIN operation in our first studies. Thus, the situation without ROBIN operation is caused by an overload of the network as it is the case with several unresponsive UDP data flows sending to the same network segment.

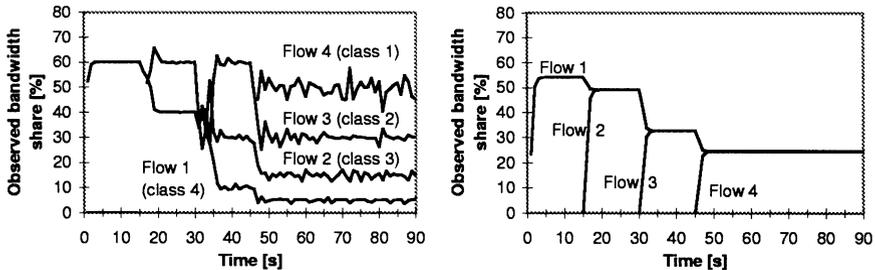


Figure 5 Bandwidth shares in (a) Ethernet with ROBIN capacity 90 Mbit/s), (b) FDDI without ROBIN regulation.

The capacity of the FastEthernet segment for ROBIN bandwidth regulation was selected to 80 Mbit/s (in Figure 4 (a)), which is shared amongst all active flows. A higher capacity of 90 Mbit/s already caused frequent oscillations of the measured throughput due to collisions on the medium (cf. Figure 5 (a)). Thus, the capacity of particular network elements as an input parameter for ROBIN as target value for maximum network utilisation has an important impact on the performance with the bandwidth regulation. Furthermore, the achievable utilisation depends on the type of the network, as it was seen for an Ethernet type where the collisions on the medium restricted the maximum utilisation.

This is completely different for a network with a deterministic medium access like FDDI. Figure 5 (b) shows the FDDI results on measured throughput without ROBIN bandwidth regulation (as Figure 4 (b) did for an Ethernet network). At all points in time, the available network bandwidth is equally shared between all active flows. The high aggregate throughput of 96,5% is achieved by the implicit flow control of the FDDI medium access. Thus, with a bandwidth regulation with ROBIN, a target network utilisation of more than 90% is achievable without degradation of the overall throughput. However, when selecting the values for the capacity of particular network elements and network segments under control of ROBIN, the delays with a high target network utilisation and additional delays coming from ROBIN's limitation of the sending rate have to be considered carefully. Our studies will be continued soon to investigate these aspects.

5 CONCLUSIONS AND FURTHER WORK

Already our short review on existing and future approaches to QoS lead to the conclusion, that the effective support of QoS in a typically heterogeneous network requires close cooperation or matching of concepts. With the variety of existing or proposed approaches, this is a very crucial point and presently hampers an effective support of QoS in such networks. Our paper introduced the new ROBIN concept of end-to-end regulation of bandwidth in local or campus area networks to control the sending bandwidth of application data flows with respect to a relative quality of service principle. In addition, we presented the state of our practical work on the implementation and performance evaluation of the ROBIN concept.

Altogether, the overall objectives of first prototype measurements and simulation runs have been met: The main goal of providing a relative QoS concept by the ROBIN mechanism is achieved and the RCU is determining a bandwidth bottleneck and the corresponding shares correctly. The benefit of the ROBIN approach and its relative QoS was underlined by a comparison to the operation without ROBIN in different network types without further adaptive flow or congestion control (as with connection-less UDP flows or other unresponsive protocols). The approach of ROBIN has the same goals as Braden (1998) suggesting active queue management and per-flow scheduling in routers for preventing unfairness and congestion collapse in the presence of unresponsive flows. In contrast to these approaches using mechanisms in the network, ROBIN indirectly meets the recommendations of Braden (1998) by adding some functionality to the sending source hosts. ROBIN establishes its principle of bandwidth scheduling on an end-to-end basis with a per-flow regulation at the sending hosts.

Further measurements with the prototype implementation will be carried out, probably leading to new iterations of concept enhancement, implementation and experimentation. These measurements in particular will focus on the evaluation of different RT modes (as mentioned above) and on the effectiveness of re-use of bandwidth of empty classes or of unused bandwidth of single flows with bottlenecks in different networks.

In parallel, the simulation studies of ROBIN performance will be extended: The performance analysis will be extended to larger and more complex scenarios as to be studied with prototype measurements. These simulation studies will in detail evaluate the impact of certain ROBIN parameters on performance and overhead and intend to investigate bounds on scalability of the ROBIN approach. The inclusion of typical LAN traffic with many short-lived flows into our experiments is an important step and may have considerable impact on ROBIN performance and efficiency of its bandwidth regulation. These aspects have to be evaluated carefully as well. However, a dynamic distinction and separation of packets belonging to long-lived and short-lived flows with a regulation of the aggregate bandwidth of short-lived flows will allow ROBIN to achieve its goals under these circumstances. The continuation of our studies probably will give hints for further improvements of the concept.

6 REFERENCES

- Akyildiz, I., Liebeherr, J., Sarkar, D. (1996) Bandwidth Regulation of Real-Time Traffic Classes in Internetworks. *Computer Networks and ISDN Systems*, Vol. 28, No. 6.
- Baker, F. (1996) Real-Time Services for Router Nets. *Data Communications International*, Special Issue May.
- Braden, B., Zhang, L., Berson, S., Herzog, S., Jamin, S. (1997) Resource ReSerVation Protocol (RSVP) -- Version 1 Functional Specification. *RFC 2205*.
- Braden, B., Clark, D., Crowcroft, J. et al (1998) Recommendations on Queue Management and Congestion Avoidance in the Internet. *RFC 2309*.
- Bradner, S. (1997) Internet Protocol Quality of Service Problem Statement. *Internet Draft <draft-bradner-qos-problem-00.txt>*, IETF.
- Bruno, L. (1996). The Internet Gets a Traffic Cop. *Data Communications International*, Vol. 25, No. 17.
- Campbell, A., Coulson, G., Hutchison, D. (1994) A Quality of Service Architecture. *ACM Computer Communications Review*, Vol. 24, No. 2.
- Delgrossi, L. (1996) Design of Reservation Protocols for Multimedia Communication. Kluwer Academic Publishers, 1996
- Floyd, S., Fall, K. (1997) Router Mechanisms to Support End-to-End Congestion Control. *Technical Report LBNL Network Research Group*.
- Frank, M., Martini, P. (1997) Practical Experiences with a Transport Layer Extension for End-to-End Bandwidth Regulation. *22nd Annual Conference on Computer Networks*, LCN '97, Minneapolis.
- Frank, M., Martini, P. (1998) Performance Analysis of an End-to-End Bandwidth Regulation Scheme. *Sixth International Symposium of Modeling, Analysis and Simulation of Computer and Telecommunication Systems*, MASCOTS '98, Montreal.
- Ghanwani, A., Srinivasan, V., Smith, A., Seaman, M. (1998) A Framework for Providing Integrated Services Over Shared and Switched IEEE 802 LAN Technologies. *Internet Draft <draft-ietf-issll-is802-framework-04.txt>*, IETF.
- Hutchison, D., El-Marakby, R., Mathy, L. (1997) A Critique of Modern Internet Protocols: The Issue of Support for Multimedia. *Proceedings of 2nd European Conference on Multimedia Applications, Services and Techniques - ECMAST '97*, Milan.
- Kalyanaraman, S., Harrison, D., Arora, S., Wanglee, K., Guarriello, G. (1998) A one-bit feedback enhanced differentiated services architecture. *Internet Draft <draft-shivkuma-ecn-diffserv-01.txt>*, IETF.
- Liebeherr, J., Tai, A. (1995) A Protocol for Relative Quality-of-Service in TCP/IP-based Internetworks. *Proceedings of 3rd IEEE Workshop on the Architecture and Implementation of High Performance Communication Subsystems*, HPCS.
- Nichols, K., Jacobson, V., Zhang, L. (1997) A Two-bit Differentiated Services Architecture for the Internet. *Internet Draft <draft-nichols-diff-svc-arch-00.pdf>*, IETF.

- Nichols, K., Blake, S. (1998) Differentiated Services Operational Model and Definitions. *Internet Draft* <draft-nichols-dsopdef-00.txt>, IETF.
- Roberts, E. (1996) Changing the Lay of the LAN. *Data Communications International*, Issue October.
- Satyavolu, R., Duvedi, K., Kalyanaraman, S. (1998) Explicit rate control of TCP applications. *ATM Forum Document 98-0152R1*.
- Schulzrinne, H., Casner, S., Frederick, R., Jacobson, V. (1997) RTP: A Transport Protocol for Real-Time Applications. *Internet Draft* <draft-ietf-avt-rtp-new-00.ps>, IETF.
- Simcoe, R.J. (1994) Test Configurations for Fairness and Other Tests. *ATM Forum Document 94-0557*.
- Suter, B., Lakshman, T.V., Stiliadis, D., Choudhury, A.K. (1998) Design Considerations for Supporting TCP with Per-flow Queueing. *Proceedings of IEEE INFOCOM '98*, San Francisco.
- XTP-Forum (1995) Xpress Transport Protocol Specification, XTP Revision 4.0. *XTP-Forum Document*.
- Yavatkar, R., Hoffmann, D., Bernet, Y., Baker, F., Speer, M. (1998) SBM (Subnet Bandwidth Manager): A Protocol for RSVP-based Admission Control over IEEE 802-style networks. *Internet Draft* <draft-ietf-issll-is802-sbm-06.txt>, IETF.

7 BIOGRAPHY

Matthias Frank graduated in computer science at the University of Paderborn in 1994. His thesis dealt with the design and evaluation of new flow control mechanisms for high speed transport protocols. From 1994 to 1996 he has been with the University of Paderborn working in the European RACE project COMBINE and from 1995 in the European ACTS project AMUSE. In these projects he gained experiences in interactive multimedia system modelling and in the evaluation of service related field trial experiments. In October 1996, he changed to the University of Bonn when his Ph.D. supervisor Prof. Dr. Peter Martini became head of the Institute of Computer Science IV. His current research focuses on end-to-end bandwidth management and control.

Dr. Peter Martini received the M.Sc. and Ph.D. degrees in computer science from the Aachen University of Technology, Germany, in 1986 and 1987, respectively. From 1986 to 1990 he was with the Institute of Computer Science IV at the Aachen University of Technology, Germany, from 1990 to 1996 he was professor of operating systems and computer networks at the University of Paderborn, Germany. Since 1996 Dr. Martini is head of the Institute of Computer Science IV at the University of Bonn, Germany. His research includes performance analysis of heterogeneous networks, protocol specification and implementation, network management, digital libraries, network security and multimedia communication.