Relation between Call Control and Resource Control in B-ISDN

Xinli Hou   Ignas G. Niemegeers

Center for Telematics and Information Technology, University of Twente
P.O.Box 217, 7500 AE Enschede, The Netherlands, e-mail: {xinli, niemegee}@cs.utwente.nl

ABSTRACT The functional separation of call control and the control of network resources brings about the flexibility in provision of telecommunication services. This paper addresses the relation between call control and resource control in the B-ISDN signalling. The merits and drawbacks of two complementary interaction strategies, i.e., simultaneous and sequential set-up are discussed. Principles of how to make a choice between the two are proposed.

1. Introduction

The establishment of calls providing B-ISDN services may involve the negotiation among multiple parties and the network on the services to be offered to the call. It also involves the negotiation with the network on the availability of resources (call admission), and the commitment of resources for the call (resource allocation). Resources are meant here in the widest sense and includes bearers and specialized resources such as conferencing bridges. It is therefore useful to separate the aspects related to the service provision and the aspects related to the service implementation. Such a separation provides flexibility and meets the needs of future telecommunication services and network architecture. This issue is known commonly as the separation of call control and bearer control [1], which in the context of B-ISDN should be called more generally separation of call control (CC) and resource control (RC).

The CC is mainly concerned with the end-to-end negotiation on the membership of a call and the services to be provided. The RC interacts with a routing entity to find a best route for each connection, and performs resource allocation on a link-by-link basis. A lot of research has been carried out on the design of signalling protocols supporting multimedia multiparty communications. To name a few examples, ISCP [6] proposes that signalling protocols for B-ISDN be structured according to the OSI application layer structure. A call control protocol at the user network interface (UNI) is presented in [5]. The CMAP protocol at the UNI supports both connection control and call control [3]. RACE project MAGIC [4] aims at investigating long term signalling requirements, architecture and protocols for B-ISDN, resulting in signaling protocols at the UNI and NNI (network node interface)[7]. Apart from B-ISDN, protocols are designed to support multimedia multiparty conferencing over Internet, such as ST-II [9] and RSVP [10] for resource control and MMCC for call control [8].

The functional separation leads to several possibilities for the call control entity to invoke the resource control for resource allocation. However, it seems that the relation between CC and RC is to a large extent overlooked. This paper studies the relation between call control and the control of resources in the B-ISDN signalling. In particular, we study when RC should be invoked and the merits of different strategies.

CC must interact with RC in order to acquire connections and other resources for user information transport. The interaction can be done within the terminal equipment by the user[1] itself [3]. This option allows the most commonality between signalling protocols at the UNI and NNI. The disadvantage of this approach is that it can be a burden to the user because the user has to figure out the resource configuration needed for the communication. The second option [5, 7] is that CC interacts with RC within the network, so that users are relieved from setting up connections and the network can efficiently use resources by, e.g., mapping several calls onto one connection. In this paper, we assume that CC and RC interact within the network. Note that the conclusions are valid for both options.

The paper is organized as follows. Different strategies to invoke resource control are introduced in Section 2. In particular, we focus on two complementary cases, i.e., simultaneous and sequential call set-up and resource allocation. We discuss the factors impacting the choice of strategies in Section 3. Principles of how to make a choice between the two are proposed in Section 4. Section 5 concludes the paper.

2. Strategies to Invoke Resource Control

As can be seen from a typical CC message flow during call set-up shown in Figure 1, RC can be invoked before the call offering so that the call set-up request is offered to the called user(s) only after the resources have been allocated. In this case, the resource allocation and the call set-up occur simultaneously from the called users' point of view.

---

1. A user is everything on the other side than the network of the B-ISDN UNI. Main components of a user are a party, a terminal equipment and possibly a terminal adaptor.

And hence it is called simultaneous set-up. An example message flow is shown in Figure 2.

Invocation of the RC can also be deferred until the call set-up is ready for commitment, i.e., the instant when the originating local exchange (OLEX) receives the READY messages from all the mandatory called users. Call set-up employing this strategy will be called sequential set-up. A message flow is shown in Figure 3 to set-up a two party call.

Note that in Figure 1 only the exchanges involved in call control are shown, i.e., the originating and terminating local exchanges (OLEX and TLEX). Figure 2 and Figure 3 also show the message flow to and from the transit exchanges (TRX) since they are involved in the resource control. Note also that the messages shown in the figures are for illustration purpose. They are not messages of any particular B-ISDN signalling protocol, although we borrowed some message names from signalling system number 7.

For some telecommunication services, e.g., two party communications, RC can be invoked as soon as the called user answers the call set-up request. It is also possible to invoke RC either at the OLEX or the TLEX as soon as the terminating terminal equipment (TTE) indicates, by means of, e.g., an alerting message, that it is suitable and currently able to accept the call. For simplicity, these cases will not be considered in this paper.

The advantage of sequential set-up is that it does not hold resources unnecessarily. However, the called users have to wait for resource allocation after answering the set-up request. Moreover there is a chance that the call must be rolled backed after called users accepted the request, because the network does not have enough resources.

The advantage of simultaneous set-up is that at the moment the called user answers, the desired resources are directly available so that the called users can communicate very soon. The disadvantage is that resources are allocated and thus wasted during the call set-up (including the call processing by the TTE and response times of called parties). In addition, there is a potential waste of resources in case the call set-up must be rolled back. A call set-up fails (partly) when one of the called users rejects (busy) or does not answer the call (absent). It may also happen when TEs are not compatible. The calling user may have to pay for the allocated resources in those cases without having received any service from his point of view; otherwise the network operator will have to pay for the resources.
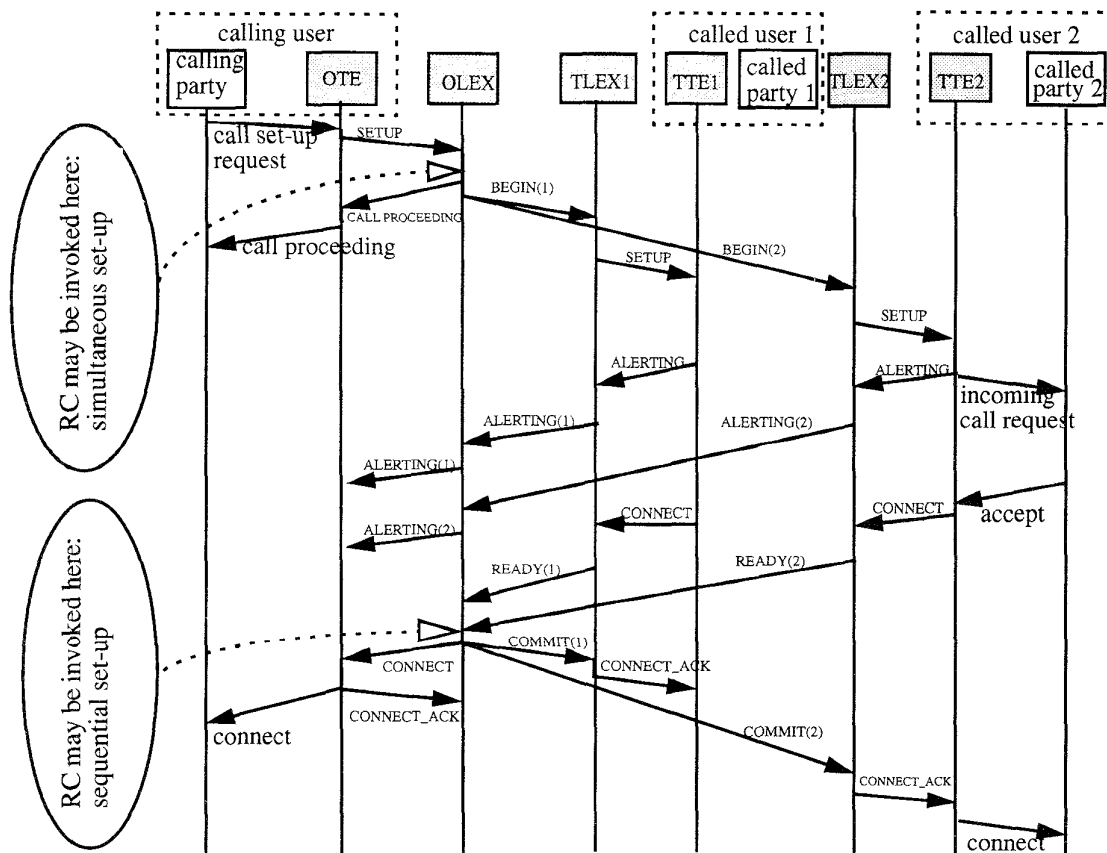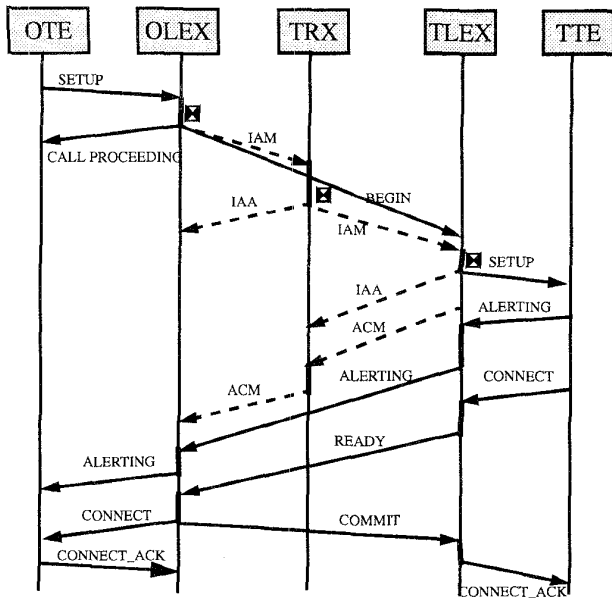

FIGURE 1. Call Control Message Flow for Setting up a Three-party Call

■ resource allocation in forward and backward directions
IAM: initial address message
IAA: initial address acknowledge
ACM: address complete message

FIGURE 2. Message Flow Scenario of Simultaneous Set-up

FIGURE 3. Message Flow Scenario of Sequential Set-up

In multiparty multimedia calls, simultaneous set-up becomes impractical for two reasons. Firstly, the amount of resources to implement a multiparty call can be large. It is more likely that expensive resources such as audio and video mixers as well as conferencing bridges will be needed. Holding these resources even for a short time can be very expensive.

Secondly, the variety of the call set-up success conditions makes it difficult to predict the exact resource configuration of the call before the multiparty negotiation of the call set-up is finished. In case of simultaneous set-up, the maximum resource configuration should be assumed and later on those unnecessary resources are released, leading to the potential waste of a large amount of resources. For example, N out of M called users can be specified as mandatory in a call set-up request. In order to do simultaneous set-up, a resource configuration supporting communications of all the M users must be assumed. It may turn out that only those mandatory called users respond positively therefore resources allocated for communication with the $(M-N)$ optional users have to be released again. It is even worse that the call set-up may have to be rolled back because some of the mandatory called users refused the call set-up request, due to, e.g., their TEs being engaged.

3. Factors Influencing the Choice of Strategies

Invocation strategies can be chosen before or during the service provision. Certain invocation strategy can be specified for a B-ISDN service by the network operator, e.g.,
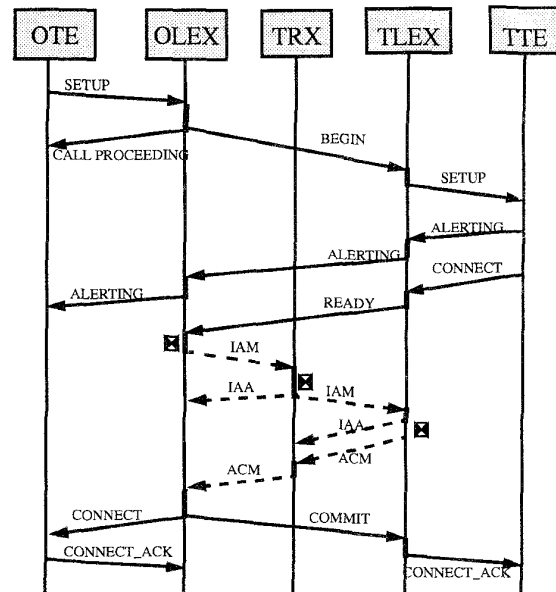
the simultaneous set-up can be specified for simple services, e.g., conventional services which are offered already by N-ISDN such as telephone service, and sequential set-up can be specified for complex services.

A specific invocation strategy can also be determined for each call during the service provision, i.e., during the call set-up to provide a telecommunication service. When a telecommunication service is requested, the OLEX must determine whether simultaneous or sequential set-up should be employed. If the simultaneous set-up is chosen, the OLEX will invoke the RC entity after the OLEX finished the local processing at the call control level. If the sequential set-up is chosen, the OLEX will invoke RC after receiving responses from all the mandatory called users.

The choice of invocation strategies depends on the call parameters which determines the potential waste of resources, QoS requirements and network status, see Figure 4. In the following, we shall discuss the role of each parameter in the choice-making process.

3.1 Call Parameters

1. telecommunication service type

Services offered by current telecommunication networks, such as telephony, will be part of the B-ISDN services. These services, called conventional services, should be provided by B-ISDN, from the user's point of view, in the same way as they are provided in current telecommunication networks. We call this requirement "service compatibility" requirement.

Provision of conventional services is currently done by signalling systems without CC/RC separation. Call and
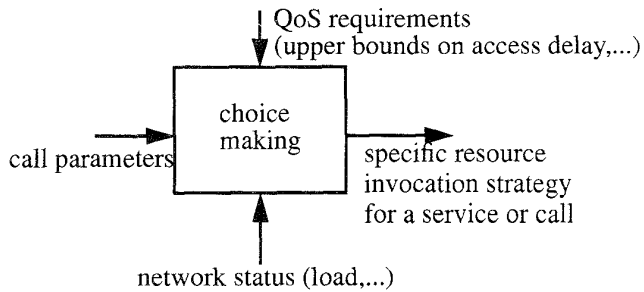
FIGURE 4. Input to the Invocation Strategy Choice-making

bearer are set up simultaneously so that a called party can communicate immediately after the call set-up request is accepted. If sequential set-up is applied in the provision of conventional services, the service compatibility will not be guaranteed anymore.

2.   expected media

The media (or service modules or service components) to be supported determine the transmission capacity needed by the call. Some media such as HDTV are resource demanding. Others, e.g., a connection for remote log-in, need a little guaranteed bandwidth.

3.   expected number of users and their attributes

They determine the resource configuration of the call and the need for special resources. For example, communications between two users do not need a conferencing bridge even if the call was assumed as a conferencing call.

A user can be mandatory or optional for the call set-up. If a call set-up request contains an optional user, the resource configuration cannot be determined until receiving replies from all the called users.

A user is allowed to access certain media. Such a permission determines the resource configuration. If one user is allowed to transmit information to a medium and the others can only receive information, a point-to-multipoint connection is sufficient for this medium. If all the users can transmit to and receive from a medium, either bi-directional point-to-point connections and a conferencing bridge, or a multipoint-to-multipoint connection is needed to support the medium. The latter can be implemented by multiple point-to-multipoint connections.

4.   size of the resource configuration

Depending on the location of each called party, the size of the resource configuration of the call can change and so does the amount of resources claimed for the call. On the one hand, the bigger the resource configuration is, the greater the amount of resources is needed to implement the call. On the other hand, the bigger the resource configuration is, the worse the QoS, in terms of access delay, should be expected by users and can be provided by the network.

We do not consider another call parameter, the expected call duration because its role is twofold. On the one hand, the longer the call is, the more the call is charged and thus users may not care the small portion of the extra cost due to simultaneous set-up. On the other hand, the longer the call is, the more likely users do not care of waiting a little bit longer in the beginning due to the sequential set-up.

3.2 Network Status

The load situation of the network determines the call blocking probability and thus has an impact on the choice of invocation strategies. The more heavily the network is loaded, the more probably the call set-up will be rolled back due to the lack of the resources. In this case, it is better not to choose the sequential set-up because it is likely that the called party is requested first to participate in a call, and is informed later of the fact that the call cannot be set up due to a lack of resources.

On the other hand, saving a small amount of resources for a short period of time can be significant in case that the network is heavily loaded. It may help to reduce the call blocking probability.

3.3 QoS Requirements

In the call set-up phase, one concerned QoS parameter is the access delay. For a calling party, we define the access delay as the interval which starts when the call set-up request is issued, and ends when the calling party is allowed to communicate if the calling party intends to participate in the communication. For a called party, the access delay is defined as the interval which starts when the called party accepts the call set-up request, and ends when the called party is allowed to communicate.

In simultaneous set-up, the resource allocation can be done in parallel with the call set-up request delivery so that the access delay of the calling party is reduced. The access delay of a called party is very small because resource allocation has finished before the call offering. In sequential set-up, delivery of the call set-up request, acknowledgment of the call acceptance and resource allocation are done serially, leading to a longer access delay.

Some calls are restricted to access delay, e.g., distributed computing where the calling party (i.e., a computer process) is hanging up to wait for the reply from the called party. To set up access delay sensitive calls, the simultaneous set-up is preferred because the call and resource control processing is done in parallel. Some calls are not too sensitive to the access delay. For example, parties in a conferencing call may not care if the access delay is a few seconds longer.

4. Principles for choosing the Invocation Strategy

The objective is to choose a proper invocation strategy so that B-ISDN services can be provided with a satisfactory QoS as efficiently as possible. The objective is vague in the sense that it can be difficult to judge to what extent the

Table 1: Service Characteristics, Network Status and Invocation Strategy

service characteristics/network status	simultaneous set-up	sequential set-up
expected media	non-resource demanding	resource demanding
expected number of mandatory called users	one	greater than one
expected number of optional called users	zero	one or more
size of the resource configuration	small	medium or large
access delay	sensitive	non-sensitive
network status	lightly loaded	medium or heavily loaded

service provision satisfies a user of the service, in particular in case human beings are involved (note that the QoS is subjective rather than objective [2]). Nevertheless, the following guideline may be established.

During the choice-making process, we must take into account the service compatibility issue, i.e., from a user's point of view conventional services will be provided in the same way as they are offered now. Most of conventional services, such as telephony etc., are point-to-point mono-medium communications. Only a limited amount of resources are needed to implement these services. Due to these considerations, we suggest that simultaneous set-up be used for calls offering conventional services.

For new services to be provided by B-ISDN, we assume that users will be getting used to the way in which B-ISDN services are provided. For instance, called users will not be surprised that a call set-up is rolled back after they all accepted the invitation to join the call. We also assume that intelligent terminal equipment will be used to offer the new B-ISDN services so that calling and called parties will get information on the progress of the call set-up procedure. Based on these assumptions, principles are proposed to determine either simultaneous or sequential set-up is suitable for a new B-ISDN service or a call providing such a service, as shown in Table.1.

5. Conclusions

Due to the separation of call control and the control of network resources, RC can be invoked in different moment of time in order to allocate resources for a call. We have compared two complementary cases, namely simultaneous and sequential set-up. For simultaneous set-up, the access delay can be very small and it is ensured that once the called users accepted the call set-up request, communication can start. However, resources are allocated and thus wasted during TE compatibility checking and the user response time. In addition, resources should not have been allocated when the call is refused by the called user. The sequential set-up enables resources to be allocated only if necessary, although the access delay is longer. It is also possible that the call is cancelled after called users accepted the set-up request, due to the lack of resources.

Choice of different invocation strategies depends on the call parameters, QoS requirements and network status. It

has been suggested that simultaneous set-up should be employed only for calls offering conventional services, and access delay sensitive, non-resource demanding calls.

References

[1] P. Blankers and K. Keskin, "The Separation between Call Control and Connection Control", IFIP Transaction, C-4, A. Casaca (ed.), North-Holland 1992, pp. 125 - 134.

[2] CCITT Recommendation E.800, "Quality of Service and Dependability Vocabulary", Blue Book, Melbourne, Nov. 1988

[3] J. DeHart, M. Gaddis and R. Bubenik, "Connection Management Access Protocol (CMAP) Specification", Technical Report of Computer Science Department, Washington University, WUCS-92-01, Feb. 1992.

[4] C.M.W.Gabriel, J.J.M. Heffels, X.Hou and H. Ouibrahim, "Heading Towards an Advanced Signalling System for Multimedia, Multiparty Services in Broadband ISDN", Electronics & Communication Engineering Journal, Vol. 5, No. 2, April 1993, pp. 103 - 112.

[5] S. Minzer, "A Signalling Protocol for Complex Multimedia Services", IEEE Journal on Selected Areas in Communications, Vol.9, No.9, Dec. 1991, pp.1383-1394.

[6] A. Paglialunga and M. Siviero, "ISCP (ISDN Signalling Control Part): The Best Candidate for the Target Broadband -ISDN Signalling Protocol", European Transactions on Telecommunications and Related Technologies, Vol.4, No.2, March-April 1993, pp.193-200.

[7] RACE II/MAGIC Deliverable R2044/BTL/DP/DS/P/005/ b1, "Signalling Protocol for B-ISDN and Interworking", Nov. 1993.

[8] E.M. Schooler, "Case Study: Multimedia Conference Control in a Packet-switched Teleconferencing System", Internetworking: Research and Experience, Vol.4, 99-120(1993).

[9] C. Topolcic, S. Casner, C. Lynn, P. Park and K. Schroder, "Experimental Internet Stream Protocol, Version 2 (ST-II)", RFC 1190, CIP Working Group, October 1990.

[10] L. Zhang, S. Deering, D. Estrin, S. Shenker, and D. Zappala, "RSVP: A New Resource ReSerVation Protocol", IEEE Network, Sept. 1993, pp.8-17.