INTERNET-DRAFT Carsten Bormann
Expires: May 1996 Universitaet Bremen
Joerg Ott
TU Berlin
November 1995

MMUSIC/ITU Interoperability Scenarios |
draft-bormann-mmusic-itu-interop-01.txt |

Abstract

This memo is a rough summary of potential scenarios where |
teleconferencing systems based on ITU standards (H.320, T.120)
interoperate with teleconferencing systems based on RTP and MMUSIC
style (``Internet'') standards. Version 01 is a minor update mainly |
based on ITU progress up to, but not including the November 1995 |
Geneva SG15 meeting (which extends two days beyond the I-D deadline). |
Change bars are provided relative to version 00.

This memo is a submission to the IETF MMUSIC working group. Comments should be addressed to the confctrl@isi.edu mailing list.

1. Introduction

Within the ITU (formerly known as CCITT), a number of ``recommendations'' (ITU name for standards) have recently been generated that cover audiographic and video teleconferencing over telephone (ISDN) lines. These recommendations are commonly subsumed by the names of the two overview recommendations, H.320 (narrow-band visual telephone systems and terminal equipment, 1993) and T.120 (data protocols for multimedia conferencing). Products conforming to these recommendations are appearing on the marketplace rapidly. Work |

is in progress to accomodate these standards to other transport media |

than ISDN, including analog telephone lines, LANs, and ATM circuits. |

With the increasing interest in multicast based teleconferencing based on standards being developed by the AVT and MMUSIC working groups of the IETF, it seems prudent to examine the potential of interoperability between systems conforming to each of the two protocol suites. Since the two suites are significantly different not only in protocol details but also in fundamental approach and assumptions, we propose to first examine the possible scenarios under which such interoperation would occur.

In this memo, we will assume basic knowledge of the work of the IETF working groups, and only provide some text to explain a few basics of the ITU teleconferencing work.

2. Terminology

This memo will use a mixed terminology, with some ITU terms and some terms as they are used in the Internet world. As some readers will not be familiar with ITU terms, this table provides a reference.

ITU Term	Equivalent term(s)	
recommendation	standard	
terminal	host, end system (including video telephones)	
MCU (multipoint control unit)	(application) gateway, intermediate system	
PSTN (public switched telephone network)	POTS (plain old telephone service)	
LAN	LAN (possibly with bridges and routers), (small i) internet	
application	media agent	
conference	session, conference	||
session	group of peer media agents	||
conference profile	session description	

3. State of standardization

As of now, ITU has standards for ISDN interconnection of pairs of H.320 systems, as well as for ISDN interconnections of multiple H.320 systems (terminals) via intermediate systems called MCU (multipoint control units). Extensions of these standards for PSTN (POTS) interconnection, for operation over LAN protocols as well as over ATM are in preparation (according to the current state of discussion, |
even in LANs, special Multipoint Controllers (MCs) are likely to be |
used as rendezvous points when more than two participants are |
involved). The T.120 family of standards defines conference control

and ``data'' applications for these environments, based on point-to-point multicasting trees defined by MCS (T.122/T.125).

In the ITU context, using IP implies a (possibly bridged or routed)

LAN (so far); the assumption is that wide area traffic will be circuit-switched via ISDN with H.221 (a frame based bit allocation protocol) being used for multiplexing. It has been decided that the LAN-based multiplex (H.22Z) will use RTP over LANs, probably augmented by ITU-specific profiling and by special setup protocols (most likely based on Q.931). The use of reservation protocols within a LAN currently is considered a local matter -- only a mechanism to request bandwidth from some (MCU-style) well-defined entity is being defined.

The IETF has standards for AV multicast (RTP and RTP payload data formats) and is working on control (MMUSIC). These standards do not explicitly differentiate between LAN and WAN applications; they were designed with WAN considerations in mind.

4. ITU Basic Assumptions

T.120 conferences are tightly coupled. The general assumption is that all participants know about all other participants, as well as their characteristics such as the set of applications available to them and the applications' capabilities. This knowledge is kept consistent throughout the course of the conference by a conference management system (GCC, T.124) using a reliable multicast transport (MCS, T.122/T.125).

5. ITU conference model (T.121)

The ITU model of the way applications interact in a conference is defined in recommendation T.121. As applications themselves are outside the scope of ITU standardization, this recommendation defines the term ``application protocol entity'' (APE) as those parts of applications that engage in the horizontal protocols defined by ITU. One or more APEs are in an ``Application Protocol Session'', i.e., they form a group of peer entities engaged in a single instance of the horizontal protocol.

T.121 defines several types of such sessions within a conference:

a) Default sessions, which are not used for actual application activity, but as a placeholder for information about applications not currently in actual sessions as well as a mechanism to maintain registries for sessions whose parameters have not been standardized.

b) Static and dynamic multicast sessions, which differ only in whether their MCS parameters have been standardized or are maintained in a registry. Both have lifespans independent of any particular member and are open to any member of the conference.

c) Dynamic user-id sessions. This type is similar to the previous |
type, except that the lifespan of a dynamic user-id session is |
bound to the presence of a specific identified creator. This |

can e.g. be used for centralized applications. |

d) Dynamic private sessions. This type is similar to the previous |
type, except that admission to a session of this type is |
controlled by its creator.

6. ITU Interconnection Models

[The following text is a slightly edited quote from one of the authors' previous contributions to SG15, AVC-797.]

Figure 1 shows a complex scenario how terminals may be interconnected through WANs and LANs, either in a point-to-point call or in a multipoint conference.

```
                         _______                      ________
      +-+      +-+      /       |   +-+      +-+     /        |   +-+
      | |      | |------  WAN #1 ---| |      | |-----  WAN #2  ---| |
      +-+      +-+      |_______/   +-+      +-+     |________/   +-+
       |        |         /          |        |          |   |
     --+---+----+---   +-+        ---+----+---+---      +-+   +-+
           |           | |                |             | |   | |
          +-+  LAN #1  +-+        LAN #2 +-+            +-+   +-+
          | |                            | |
          +-+                            +-+
```

Figure 1: Interconnection Models for LANs and WANs

This figure is a generalization of the following possible scenarios:

a) WAN only terminals (listed here for completeness)

b) LAN terminal(s) connected to WAN terminal(s) through a gateway

c) LAN terminals within a single LAN only

d) LAN terminal(s) connected to other LAN terminal(s); the LANs are interconnected by a WAN

e) WAN terminal(s) connected to WAN terminal(s); different WANs
are used; the different WANs are interconnected through a LAN |
[this scenario is somewhat out of focus for ITU work]

The design of any transport for T.120 data information should consider the existence of all the above scenarios. This means that any extension of the T.123 protocol stacks has to be able to interwork with all other T.120 terminals that do not implement this extension. As a corollary, the service offered by the T.122/T.125 Multipoint Communication Service must not be affected.

[End of quote]. The latter comment obviously also applies to audio and video streams.

Note that each of the ``LANs'' in the ITU scenarios could be an internet in the interoperability case; appropriate gateways would be used for bridging. A LAN-to-WAN gateway would need to perform at least the following functions:

- conversion from ISDN multiplexing (H.221) to a format more suitable for LANs (H.22Z, which is based on RTP)

- conversion of audio/video encoding formats (e.g., deletion of BCH envelopes for H.261 to obtain RTP payload data formatting), as required

- filtering of data streams to keep only those absolutely necessary (e.g., the LAN could use ``continuous presence'' of all participants by their video streams, while on the WAN only the streams of the speaker and the previous speaker are retained).

- transport layer gatewaying (e.g., X.224/RFC1006/TCP/IP to X.224/Q.933/Q.922)

7. Types of interoperation

Based on these interconnection scenarios, the following scenarios for interoperation between ITU and IETF conferencing systems could be addressed:

1) T.120 ISDN terminal users ``phone in'' to a classical IETF-style WAN internet multicast session (e.g. an IETF broadcast).

   1a) Actually, not just one terminal but a whole T.120 conference network is built on the T.120 side.

   1b) The internet WAN session becomes more controlled than a ``classical'' session -- more information needs to be relayed to the T.120 session control. (This, obviously, depends on what kind of session control is used on the Internet side.)

   The assumption here is that the IETF style conference is the one ``in control'' and ``phoners-in'' are accepting some semantic lossage. E.g., the T.124 (GCC) conference roster (attendance list) could be incomplete, it might not be possible to perform certain actions (such as addressing single participants), etc.

   Note that for a conference in which #apps applications (such as whiteboard etc.) are used, MCS/GCC runs into a hard limit of 64535/(#apps+1) participants (or less than that -- the denominator may actually be higher).

2) LAN-wide internet multicast sessions are used behind a local T.120 MCU (i.e., LAN systems don't speak T.120 but support classical IETF sessions only)

2a) Internet multicast sessions with additional T.120 consciousness are used behind a local T.120 MCU (different from 2?). In the simplest case, they would have to be able to take part in and make use of the T.124 conference roster generation process; applications could announce their capabilities in the application roster, etc. |

2b) Additional LAN participants just listen in to multicast traffic on their LAN, these don't take an active part in the T.120 protocols. | |

The assumption here is that T.120 is ``in control'' and the LAN group has to cope. |

3) A group of internet WAN participants and a group T.120 WAN participants are joined by a gateway/MCU. Both parts get the illusion of a homogeneous conferencing environment.

The ``gateway/MCU'' would be a much more sophisticated form of the same gateway referred to above. Achieving a homogeneous conferencing environment certainly would require a high degree of semantic compatibility of the IETF conference control protocol with those of the ITU. |

8. Technical implications

For all these scenarios, special consideration must be given to the following aspects.

[Note: These items must be sorted into those relevant specifically to MMUSIC and those relevant only for a broader discussion.]

8.1. Type of mapping within a gateway

A gateway may attempt to map a semantic feature of one domain into an equivalent feature of the other domain and vice-versa (bidirectional mapping). Alternatively/additionally, it may attempt to tunnel information only supported by one domain through the other domain in A-B-A configurations (e.g., it could attempt encoding the T.120 application capabilities in an RTCP text attribute).

8.2. Agreement protocol vs. conducted mode behavior

The ITU-T conference control distinguishes two different modes of operation: a conducted and a non-conducted mode. In conducted mode, a single participant largely controls the conference requiring the others to query for permission to perform certain actions (which actions are affected is defined in the session description as well as the respective recommendations for conferencing applications). In non-conducted mode no such restrictions are imposed.

These two modes represent the two extremes that can be thought of when using the MMUSIC agreement protocol; they could be modeled by |

specific voting rules in the MMUSIC agreement protocol, which allows |
other styles of voting rules as well. Within the ITU-T conference |
control no such intermediate modes are defined.

8.3. Resource control (bandwidth management)

The current approach pursued by SG 15 is to limit the number of AV connections gatewayed into a LAN.

In addition, possibly, recoding will be required between high and low bandwidth environments.

8.4. Addressing

Participants will have to be addressed by POTS/ISDN numbers (generally E.164) as well as by addresses from internets and the Internet. This is confounded further by ITU embracing IPX as well as IP.

8.5. Session description

In the ITU model, a session is ``described'' by participants that update roster information and that actually start applications based on the capabilities in that roster information. Currently, only a small static information base may be configured at conference startup time (part of which remains unchanged throughout the course of the conference). This information base describes the conference (e.g. the conference name) and defines some attributes of the conference (conducted or not, some authentication mechanism, e.g. a password in
the simplest case, etc.). A more detailed a priori description of |
the conference will be defined in the new ``T.RES'' advance |
reservation work.

In the classical IETF model, the session description is broadcast beforehand; it cannot be changed during the session or adapted to the capabilities of the participants. Other uses of the IETF session description language SDP are being considered; note that currently multicast address allocation (see also below) is intertwined with session description broadcasting.

8.6. Authentication

Internet applications generally will use cryptography based end-to-end authentication and confidentiality.

MCS does not use authentication within the conference; instead, unwanted participants cannot obtain transport connections to the MCS domain (data part of the conference) at all. The T.120 conference control protocol GCC currently allows for a challenge-response mechanism for authentication to the MCS domain. Confidentiality can

be achieved using H.233/H.234 by enciphering the entire transport |
stream, i.e., hop-by-hop based enciphering, possibly separately for |
audio, video, and MLP (``data''). This requires trusted MCUs

(proposals for operations with non-trusted MCUs are being made).

8.7. Use of Multicasting

Given the intent of ITU SG15 to generate a draft standard by November | 1995 (to be voted on in 1996), complications such as multicasting | have a relatively low priority. It seems unlikely that SG8 will come | around quickly to extending MCS to incorporate multicast subtrees (based on multicast transports such as MTP-2 or RMP). Multicast will | be an option for ITU's usage of RTP, but note that ITU needs a handle | on IP multicast address allocation for this to become real (see next | item).

In any case, for operational use of multicasting in environments that may or may not have multicast capable routers (or operating systems, or protocol stacks) it must be possible to use point-to-point meshes as a fallback. This fallback should be automatic; manual configuration is unlikely to be workable. One solution currently being offered within the ITU environment is to start a conference as a point-to-point mesh and to allocate a multicast address and to start testing multicast connectivity simultaneously. Terminals that | do have multicast connectivity withdraw (partially) from the point-to-point mesh.

8.8. Multicast address allocation

In IETF conferences, the allocation of multicast address is done administratively (by applying for an address at IANA) or by global broadcasting of address claims. Administratively scoped multicast may alleviate the problem for conferences confined to a site only. For operational use, an address allocation mechanism must be found that scales to large numbers of conferences and avoids conflicts quite reliably. Note that conferences that must be protected from denial-of-service attacks will need a form of authentification that might make conflicts less of a problem.

9. Security Considerations

Any interoperation between ITU-based systems and Internet-based systems must take care to preserve the point-to-point link based security model underlying the ITU standards. In T.120, much of the access control relies on being able to reject the attempt to join a conference via an ISDN connection to an MCU. See also ``Authentication'' above.

10. Authors' Addresses

Carsten Bormann
Universitaet Bremen FB3 MZH 5180
Postfach 330440

D-28334 Bremen
GERMANY
cabo@informatik.uni-bremen.de

Joerg Ott
Technische Universitaet Berlin FR 6-3
Franklinstr. 28/29
D-10587 Berlin
GERMANY
jo@cs.tu-berlin.de

Appendix: Pertinent standards bodies

ITU-T SG8: T.120 standardization (MCS, application protocols, conference control)

ITU-T SG15: defines LAN-WAN gateway

IMTC CNAG: defines LAN-WAN interworking |

IETF AVT WG: defines real-time transport and payload data formats

IETF MMUSIC WG: defines conference control