X3S33/93-226**R2**

ISO/IEC JTC1/SC6/

1993-06-03

ISO/IEC JTC1/SC6
TELECOMMUNICATIONS AND INFORMATION
EXCHANGE BETWEEN SYSTEMS
Secretariat: U.S.A. (ANSI)

Title: Terminology, Taxonomy, and Conceptual Framework for Multicast

Source: USA

Project(s): JTC1.06.36

Status: US Contribution

1. Introduction

In our review of the proposals that have been made to SC6 for multicast services and protocols, we have found that it is important to use a consistent terminology and set of basic concepts to ensure that we are “all talking about the same thing”, and that comparisons between features in different proposals are valid. We have therefore made an effort to formalize the terminology that we use to discuss multicast, and to create a taxonomy of multicast semantics that is general enough to be used to characterize a very wide range of individual service and protocol proposals with respect to a common set of concepts.

This contribution describes the terminology, taxonomy, and set of concepts that we have developed as tools to support our work on multicast. We believe that this framework can and should be used to facilitate the discussion of proposals for multicast during the meetings of SC6 and its working groups in Seoul in October. As an annex to this contribution, we include a preliminary classification of the multicast proposals that are on the table in SC6 with respect to this framework.

2. Taxonomy

Our taxonomy of multicast recognizes four principal dimensions that define a "solution space", within which any particular multicast scheme can be located by describing where it lies along each of the four dimensions. Each dimension represents a set of differentiating characteristics, such that the important differences between two multicast schemes can be described in terms of the different selection each scheme has made from each set of characteristics. The four dimensions are:

— population characteristics;

— communications discipline;

— transmission characteristics; and

— ordering characteristics.

2.1 Population Characteristics

Population (or "group membership") characteristics describe the way in which multicast groups are defined, established, and maintained. Three distinctions are recognized within this category:

Static vs. Dynamic Population

A **static** group is one in which the population of the group is determined by mechanisms that are independent of any particular group conversation, and does not change during the course of a group conversation.

A **dynamic** group is one in which the population of the group is determined by mechanisms that may operate either independent of any particular group conversation or in the context of a particular group conversation (or both), and may change during the course of a group conversation.

Centralized vs. Decentralized Group

A **centralized group** is one in which a single (designated) member is permitted to send, and all other members are permitted to receive. Depending on the communications discipline, it may or may not be possible for receivers in a centralized group to respond to communications originated by the sender.

A **decentralized group** is one in which any member is permitted to either send or receive (or both).

Known vs. Unknown Population

A group has **known population** when it is possible for any member of the group to determine the identity of every other member of the group. *[Note: In the context of a particular group conversation, "known population" may refer more specifically to the ability of the sender(s) to determine the precise membership of the group, whether or not members who may only receive are able to do so; the important aspect in this case is the ability of a sender to know in advance the complete set of receivers who may be expected to receive a particular PDU]* The "known" population may be either static or dynamic.

A group has **unknown population** when a member of the group is not able to determine the identity of every other member of the group. *[Note: A sender in a group of unknown population may not know in advance the complete set of receivers who may be expected to receive a particular PDU.]*

2.2. Communication Discipline

Communication discipline describes the behavior of the sender(s) in a multicast group.

Send Only

The sender(s) may

Send/Receive

The sender(s) may both originate (send) multicast transmissions and receive responses to them.

[Note: There is an additional case that is not covered by these two alternatives, namely the case in which there is more than one sender in a group, and each sender is permitted to receive the transmissions of other senders, although receivers are not permitted to send responses back to any of the senders.]

2.3. Transmission Characteristics

reliable/unreliable
fully reliable - gurantee of in-order error-free delivery to all members of the active group.
AGI-reliable - defined by the AGI parameter as specified by user
1-reliable -
unreliable -

<<note: this category subsumes Marc’s category of “error control”, and has not yet been fixed......>>

2.4. Multicast Ordering

local

The PDUs generated by a sender are delivered to all of the receivers in the group in the same order the PDUs were sent by the sender. But there is no relation guaranteed in the delivery order among senders of the group.

causal

The PDUs generated by a sender are delivered to the receivers in the group in a partial ordering according to some relation existing among the sending events.

total

The PDUs generated by all senders are ordered and delivered to all receivers in the same order.

3. Concepts and Terminology

<<include terminology from X.6 and 7498 multipeer addendum where appropriate and relevant — work on this section has not been completed>>

3.1. Definitions

Multicast Service: is one in which a single data unit transmitted by a source i received by multiple destinations; it is a one-in, many-out service.[1]

Multicast Group: a set of members participating in the multicast service. The multicast group is defined by a rule (or set of rules) which identifies a collection of members implicitly or explicitly. This rule may associate members for the purpose of participating in a call, or may associate members who do not participate in data transfer but do participate in management, security, control, and accounting for the multicast group.

Active Group: those members of the multicast group that are participating in a particular multicast call at a given instance of time.

Active Group Integrity: specifies whether active group integrity applies to this call. If active group integrity applies, at least a quorum of participants must be joined to this call before any data transfer can occur. If the number of participants joined to the call drops below a quorum, the service will either terminate the call or enforce a cessation of data transfer until a quorum is again achieved.

Note: A variation on active group integrity may use “key member(s)” rather than a “quorum” of members, such that without these key member(s), the call will either be terminated or data transfer will cease.

AGI specifies conditions for the group conversation such as: minimum number of active entities, data transfer reliability, data transfer performance characteristics (delay, throughput), PDU size, and other characteristics that are set for all entities of that group-type.

Partial order

Total order

Causal order

---

[1] A multicast service is intended to mean a service in the sense of services specified in Recommendation X.1/X.2. There is no intended relationship to an OSI service.

quorumcasting - if a multicast group consists of N nodes, quorumcasting sends to a subset consisting of any M (M<=N) of those nodes.

5.3.3.2 bis Multi-Peer Mode

A Multi-Peer Data Transmission (MPDT) is an association established for the transfer of data between members of a group. There may be zero or more peer-entities receiving a single MPDT. This association is established between the members of the group themselves and between each entity and the next lower layer. The ability to begin and end a group conversation and to exchange data within it is provided to the entities in a given layer y the next lower layer as a MPDT service. The use of an MPDT service by peer entities proceeds through six distinct phases:

a) group definition
b) group activation
c) data transmission,
d) group deactivation, and
e) group disband

3.2. Model of Multicast Operation

<<describe the phases of operation>>

<<John — describe as three phases, not five — e.g. enrollment/de-enrollment>>

Multicast operation may be described in terms of three distinct phases: **group enrollment**, **group allocation**, and **multicast data transfer**.

Enrollment
The enrollment phase creates the shared state in the OSIE necessary to permit the allocation of an instance of communication.

This phase makes an object and its capabilities known to the network, any addressing information is entered into the appropriate directories (and routing tables), certain parameters are set that characterize the communication this protocol can participate in, access control rules may be established, etc. This phase creates everything necessary for this protocol to be instantiated for use in a communication. This phase is often characterized by two distinct subphases: the actual registration of the protocol and the activation of the protocol. In general, it is useful to be able to separately control the registration and the actual availablility of the protocol to participate in a communication. Activiation/Deactivation is the traditional operations of taking a facility “off-line” without deleting the systems knowledge that the facility exists. This phase may specify constraints on the participants in the communication.

The De-Enrollment phase is entered when the protocol is either made unavailable for communication (but remains registered), i.e. deactivated, or when the registration is deleted. It will be protocol specific as to what actions are taken with respect to active instantiations if a capability is de-enrolled, i.e. its registration is deleted or the registration is deactivated. In some cases, all active instantiations may be deleted; in others, active instantiations may be allowed to complete normally. The action is not required to be the same for de-registration and deactivation.

Allocation

The allocation phase establishes the shared state amongst the participants based on information established during enrollment of a specific instance of the group conversation for that group to support the Data Transfer Phase.

This phase is entered when the (N+1)-layer or System Management actually initiates a procedure which causes the protocol to allocate resources for a communiation and to enter a state such that it is ready to transfer data. The behavior associated with this phase can range from simply creating bindings with upper and lower protocols to the exchange of protocol to produce an initial shared state with other participants for some subsequent data transfer (connectionless), or with an explicit exchange of initial state infomation (connection-oriented or perhaps a multipeer join). It is during this phase that the specific QoS requirements acceptable to the (N+1)-layer for data transfer to and from this address must be made (or modified) if they were not made during the Enrollment Phase.

The DeAllocation phase is entered when the participant has completed the data transfer phase and terminates any shared state established during the Allocation phase. Enrollment parameters are generally unaffected.

Data Transfer

This phase provides the actual transfer of data. It may include various functions to maintain the QoS, either in terms of error control or resource allocation.

This phase is entered when the actual transfer of data is effected according to the requested Quality of Service among the addresses specified during either of the previous two phases. For application protocols, the Data Transfer Phase will be further subdivided into specialized subphases.

4. Mechanisms and Policies

<<describe the separation of mechanism and policy in terms that make it clear how this distinction facilitates an understanding of the way in which multicast works>>

<<remove the sow's ear....>>

Each protocol is defined as a set of functions that achieve the basic requirements of that protocol. The choice of functions is made based on the operating region in which the protocol is intended to exist and the desired level of service that is to result from its operation. Each protocol function is divided into two fundamental parts: a mechanism and one or more policies. Mechanisms are static and are not changed once the protocol is specified. While, different policies can be chosen and applied to each (N)-association. For any one mechanism, there are a variety of policies that may be applied to it. For example, consider the flow control function, the basic sliding window is one flow control mechanism which is used in many protocols and is part of the protocol specification. Once implemented this mechanism is not modified. However, there are a variety of policies for extending credit and thus controlling the flow that may go with this mechanism: from simply extending new credit on receipt of a PDU, to periodically sending new credit, to high/low watermarks, to slow start, etc.

By separating policy and mechanism and allowing policy to be set at allocation time or to be modified during the Data Transfer phase of hte (N)-association, the operating range of a protocol can be increased and its ability to optimally serve a particular subset of the operating region can be greatly enhanced. The choice of policy depends on the traffic characteristics of the (N)-association and the Quality of Service (QoS) and Nature of Service (NoS). The (N)-PM uses QoS to select the protocol, mechanisms, or policies to match the desire with the reality. Thus, the task of the (N)-Protocol Machine (PM) is to translate these QoS characteristics as supplied by the (N+1)-PM into a particular choice of mechanism and policy based on the NoS expected from the (N-1)-PM.º Cleanly separating policy from mechanism is an important consideration in the design of a protocol that is capable of serving a wide range of QoS parameters while achieving a tighter degree of optimality. This also implies that QoS and NoS parameters should be defined based on their ability to select policy.

[The distinction between Quality of Service and Nature of Service is essentially recognition of the old adage that “you can’t make a silk purse from a sow’s ear,” but maybe we can improve on it a bit. Quality of Service represents a set of characteristics that the (N+1)-PM desires from the (N)-PM for a particular instance of communication (the silk purse). Nature of Service represents the set of characteristics that a (N-1)-PM is actually providing and is therefore likely to be able to provide in the future (the sow’s ear).]

All protocols should support the ability to negotiate policy at least during the allocation phase, if not during the Data Transfer phase as well. The negotiation of policy on a given (N)-associaton is accomplished using the (N)-context. The (N)-context is specified during the Allocation and/or Data Transfer Phases (see below). There are two basic approaches to defining the (N)-context for basic data transfer protocols: the most general approach (and perhaps too general) would be based on a collection of QoS parameters alone, or a more pragmatic approach might identify particular well-known operating regions based on traffic characteristics, e.g. bursty, bulk transfer, voice, etc., mechanisms in a specific protocol, etc. and perhaps parameterize within these “well-known” regions. (If there is any new facility that might be added to the lower layers, it would be the ability to alter context, i.e. policy, during the lifetime of an association.)

<<in the following sections we describe some of the mechanisms and policies that we have identified for each of the phases of multicast operation...but there’s a lot of work that needs to be done to distinguish between statements of definition/description and statements of prescription/proscription.>>

4.1. Enrollment

Group Definition
The set of parameters required to define the characteristics of a group to either the type or instance necessary to initiate the allocation phase.

Group Management
The process by which a group definition and other group enrollment information is maintained and distributed through the OSIE.

Access Control
Constraints on the membership of the group.

4.2. Allocation

Allocation Mechanism

Local Binding
These are the local operations required before a connectionless transmission can be sent or received.

Two-way Handshake
An allocation mechanism consisting of a two-way exchange used to establish the necessary shared state among the communicating protocol entities to support a certain class of mechanisms used during the data transfer phase.

Three-way Handshake
A more robust allocation mechanism consisting of a three-way exchange used to establish the necessary shared state among the communicating protocol entities to support a certain class of mechanisms used during the data transfer phase. This mechanim is used when the Qos of the underlying service is insufficient to ensure the success of a two-way handshake with an acceptable probablity.

<<This is as far as I got as of 6/4/93 AM....>>

6.2. Allocation Policies
TBD

6.3. Deallocation Mechanism
6.3.1. Abrupt
The mechanism to terminate the shared state created by the allocation mechanims. The abrupt deallocation mechanism terminates the connection discarding any data that may have been queued for transmission.

6.3.2. Graceful
The mechanism to terminate the shared state created by the allocation mechanims. The graceful deallocation mechanism terminates the connection after waiting any data that may have been queued to be delivered.

6.4. Context Selection
6.4.1. Isotropic
All members of the group are operating under the same set of policies.

6.4.2. Anisotropic
The members of the group are operating under the different sets of policies. In most cases, this would mean that subsets of the membership would have the same policies.

7. Data Transfer
7.1. Lost and Duplicate Detectcion Mechanism
7.1.1. Mechanism
7.1.1.1. Ack
A mechanism which allows the receiver of PDUs to notify the sender that it will not request further retransmission of the PDUs.

7.1.1.2. Selective Nack
A mechanism which allows the receiver of PDUs to request the retransmission of PDUs with

specific sequence numbers. (Note: this mechanism alone can not guarantee reliable delivery.)

7.1.1.3. Selective Ack/Nack

A mechanism which allows the receiver of PDUs to notify the sender that it will not request further retransmission of the PDUs or to request the retransmission of PDUs with specific sequence numbers.

7.1.2. Multicast Policies

<This list is not considered complete.>

7.1.2.1. All must ack

The most rigid policy is that the sender must receive acks from all of the receivers.

7.1.2.2. Given subset

The sender must receive acks from a specific subset of the receivers.

7.1.2.3. Quorum

The sender must receive acks from a given number of receivers, where the number is less than or equal to the population of the group.

7.1.2.4. Time out

The sender may assume delivery of PDUs after a certain period of time and assume that no retransmissions of those PDUs will be requested.

7.1.2.5. None

QED

7.2. Flow Control

7.2.1. Mechanism

7.2.1.1. Rate-based

A mechanism by which the receiver controls the amount of data the sender may transmit by specifying a rate (usually in units of PDUs/sec or octets/sec) at which PDUs may sent that can not be exceeded.

7.2.1.2. Credit based

A mechanism by which the receiver controls the amount of data the sender may transmit by specifying the number of PDUs may sent without receiving additional credit.

7.2.2. Multicast Policies

<This list is not considered to be complete.>

7.2.2.1. Rigid

All members of the group must maintain the flow control. If one falls behind all are affected.

7.2.2.2. Drop Slow

Members of the group that drop too far behind are discontinued from the group.

7.2.2.3. Average

The flow control is maintained as the average of the rates of the individuals.

7.2.2.4. Quorum

The flow control is maintained based on the rate sustained by the quorum.

7.2.2.5. None

QED

7.3. Sequencing

7.3.1. Policies

TBD

7.4. Distribution

Distribution only occurs in relaying layers (see the Annex) such as the Network Layer.

7.4.1. Mechanism

7.4.1.1. Native

The multicast distribution is an inherent property of the layer (media).

7.4.1.2. Flooding

The multicast distribution is accomplished by each intermediate system that receives the multicast PDU sends a copy on all paths to neighboring intermediate systems.

7.4.1.3. Exploders

The multicast distribution is accomplished by multicast PDUs being sent to a server which maintains (N-1) pairwise connection with the members of the group and sends copies to all members. (This may include mechanisms which have multiple servers and copies are sent to other servers for redistribution. This latter case blurs the distinction between exploders and spanning trees which are non-minimal.)

7.4.1.4. Spanning Trees

The multicast distribution is accomplished by creating a minimal spanning tree based on some metric (cost, delay, hops, etc.). Multicast PDUs are sent on the arcs of the tree being replicated and redistributed in the intermediate systems where the tree bifurcates.

7.5. Data Corruption

7.5.1. Mechanism

7.5.1.1. Forward Error Correcting Code

Error codes are used which detect and correct data corruption.

7.5.1.2. Error Detection Codes

Error codes are used which only detect errors.

7.5.2. Policy

7.5.2.1. Choice of code

Different codes may be chosen depending on the number of errors that may be detected and corrected.

7.5.2.2. How much to check

The policy may determine whether to check the PCI, or selected fields in the PCI or the data or the whole PDU.

7.6. Delimiting

7.6.1. Internal

The delimiting of the PDU is part of the PDU structure, e.g. a length field.

7.6.2. External

The delimiting of the PDU is not part of the PDU structure, e.g. framing bits.

Annex

Characterization of Multicast Proposals

Each individual proposal for a multicast service definition or protocol specification can be described, using the terminology and taxonomy developed in this contribution, as a set of selections from a "menu" of characteristics that define the semantics of multicast in a multi-dimensional "solution space". In this Annex we attempt to do this for each of the multicast proposals that we expect to be discussed during the meetings in Seoul.

Include here the table from 93-221 (Marc Cohn's paper), but with 5 columns labelled with the titles (more or less) of Marc's sections 2.1, 2.2, etc. Below the table, include appropriately re-worded text from each of these sections, and include with that text as much of the material from Marc's section 4 as possible.

For HSTS and TP5, describe them by actually breaking out an additional table, which shows three columns with the intermediate labels (connectionless, stream, and connection-mode, as in the original table 1 from Marc's contribution).

Include a similar table for the protocols — especially for Kevin's work (should be able to cut and paste from his document 93-192).