ATNP/WG 1/IP
ATNP/WG 3/IP

September 23, 1998

AERONAUTICAL TELECOMMUNICATION NETWORK PANEL

WORKING GROUP 2 (Internet Communication)
Bordeaux, France, 30 September – 02 October 1998

MULTICAST ARCHITECTURE

Agenda Item:
Prepared by: James Moulton
Presented by: James Moulton

SUMMARY

At the previous meeting of WG 2 (Utrecht), members of WG 2 requested further information on the architecture underlying multicast communication. This paper presents an overview of multicast and distinguishes multicast from broadcast and presents terminology useful in describing multicast communication.

1 Introduction

The following paragraphs are lifted from an early draft of the proposed modifications to the OSI Basic Reference Model. The text as originally constructed is somewhat difficult to read since it is written in standard terminology. For that reason, the text has been modified and added to for clarity.

2 Description of multipeer

Note: The term used with the OSI Reference Model for multicast is "multipeer" and the new section of the OSI Reference Model is entitled "MultiPeer Communication Architecture". These terms are descriptive of the breadth of communication possibilities based on a relaxation of the point-to-point communication.

Note: Multi-Peer Communication Architecture (MPCA) describes the combined service and protocol architecture needed to support multi-peer communication. Mulit-peer as defined below is limited to the service view of communication among more than two participants. Protocols defined to support multi-peer services in the MPCA may be described using terms such as:

- multicast: where a single service invocation results in the (attempted) delivery to a group of recipient service-users;
- broadcast: where a single service invocation results in the (attempted) delivery to all peer service-users in the OSIE;
- distributed peers (as in Transaction Processing or chaining in Directory Service Agents): TBD;
- producer consumer model (such as in Time Critical Communication Architecture): TBD.

Editor's Note: The above text was driven by the Editor's Instructions to include a description of different types of MPCA. In terms of the architecture defined in the remainder of the Standard, these terms and their distinctions become special cases of the general architecture

2.1 Definitions

MPCA-data transmission : a transmission from a single source to multiple destinations.

(N)-group-association : a cooperative relationship among (N)-entity-invocations.

(N)-group-connection : an association requested by an (N+1)-entity for the transfer of data among a group of (N+1)-entities. The association is established by the (N)-layer and provides explicit identification of a set of (N)-group-data-transmissions and agreement concerning the (N)-group-data transmission services to be provided for the set.

(N)-group-connection-endpoint : a terminator at one end of an (N)-group-connection within an (N)-service-access-point

2.2 Description

For information to exchanged among (N+1)-peer entities, a group-association is established among them in the (N)-layer using an (N)-protocol.

The rules and formats of an (N)-protocol supporting multi-peer communication are instantiated in an (N)-subsystem by an (N)-entity. An (N)-entity may support one or more (N)-protocols including both MPCA and non-MPCA protocols. (N)-entities when supporting multi-peer communication maintain the binding of (N)-group-connections to the appropriate (N+1)-entities at the appropriate (N)-SAPs.

Within MPCA, multi-peer communication has the following fundamental characteristics:

a) it involves establishing and maintaining a three or more party agreement concerning the transmission of data between the group of peer (N)-entities concerned and the (N-1)-layer providing the service;

b) it allows the negotiation between all the parties concerned of the parameters and options that will govern the transmission of data;

c) it provides group identification by means of which the overheads involved in address resolution and transmission can be avoided on data transmissions;

d) it provides a context within which successive units of data transmitted to the (N)-group from a peer (N)-entity are logically related, and makes it possible to maintain sequence and provide flow control for those transmissions.

The characteristics of MPCA are particularly attractive in applications that require a single transmission to be delivered to a set of group members.

Within MPCA, multi-peer communication only supports communication where the same information is sent to multiple recipients. It does not include those cases where an (N)-entity needs to send different information to different recipients. The physical transmission of the information may be serially (that is where the information is individually sent to each recipient) or at the same time (using an (N-1)-MPCA-service. MPCA may also be supported through the use of (N)-relays that support MPCA services.

3 Concepts of groups

3.1 Definitions

(N)-Group, (N)-multicast group : a set of (N)-entities that abide by the rule(s) for belonging to a group able to utilize multicast services.

(N)-enrollment group : a set of (N)-entities which belongs to the multicast group and meets the group membership criteria established during enrollment. It may not be possible for any (N)-entity to determine the members of the enrollment group, i.e., the rule defining the enrollment group is known but it may not be feasible to determine all entities that satisfy the rule.

(N)-enrolled group : a set of (N)-entities which belongs to the enrollment group and has successfully completed the enrollment phase. In the context of the enrollment phase, the enrolled group is those (N)-entities which have completed the registration and the activation procedures. It may not be possible for any (N)-entity to determine the members of the enrolled group, i.e., the rule defining the enrolled group is known but it may not be feasible to determine all (N)-entities that satisfy the rule.

(N)-active group : a set of (N)-entities which belongs to the enrolled group and has entered the data transfer phase.

Note - A multicast group is the highest level in the group hierarchy; it defines the universe for the characterization of the different group types. Multiple enrollment groups may be defined from the multicast group. Either a group name or a set of group names are assigned to an enrollment group. From an enrollment group, a single enrolled group bound to a single group address can be formed.

3.2 Group Description

The use of the generic term group and the qualified terms such as enrollment-group define a hierarchy of group types. The most general group type is the (N)-enrollment-group and consists of the (N)-entities which abide by the rules defining the group. The (N)-enrollment-group is identified by the (N)-group-name. The (N)-registered-group is a proper subset of the (N)-enrollment-group and consists those (N)-entities which have announced to the OSIE their ability to participate in (N)-group-connections. There can be only one (N)-registered-group for any (N)-enrollment-group. The (N)-enrolled-group is a proper subset of the (N)-registered-group and consists of those (N)-entities which have announced to the OSIE their willingness to participate in an (N)-group-connection. There can be only one registered group for any single (N)-group-connection. At this point, the (N)-enrolled-group:

- consists of the set of (N)-entities which may participate in an (N)-connection
- is identified by the (N)-group-name
- is a proper subset of the (N)-enrollment-group and (N)-registered-group.

When an (N)-entity wishes to participate in an instance of communication by sending or receiving data over an (N)-group-connection, it joins the (N)-active-group corresponding to the (N)-group-connection. An (N)-active-group is a subset of the (N)-enrolled-group and consists of those (N)-entities which have joined an (N)-group-connection. There may be more than one (N)-active-group defined for an (N)-enrolled-group. Each (N)-active-group is identified in the OSIE by an (N)-group-address. Note: The (N)-group-address may consist of the (N)-group-name along with other additional information such as a Connection-endpoint-identifier (CEPI) needed to identify individual (N)-group-connections within the scope of the (N)-enrolled-group as identified by the (N)-group-name.

4 Group Characteristics

MPCA encompasses 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;

— dialogue control characteristics; and

— reliability characteristics.

Note: SC 6 N9161 introduces the concept of a Group Association to embody some of these concepts. However, Group Association and the associated limitations on conversations is deemed too narrow for an architectural document.

4.1 6 Population Characteristics

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

Static vs. Dynamic Population

A **static** group is one in which the population of the group does not change.

A **dynamic** group is one in which the population of the group may change.

Known vs. Unknown Population

A group has **known population** when it is possible for members of the group to determine the identity of one or more other members of the group.

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.]*

4.2 Communication Discipline

Communication discipline describes the allowed group behavior of the sender(s) and receivers in a multicast group with respect to the transfer of data.

<u>Send Only/Receive Only</u>

Within a group a sender(s) may originate (send) multicast transmissions that are not associated with any response(s) from recipients. Within the same group receivers may only receive transmissions.

<u>Send/Receive</u>

Within a group all members may either send or receive messages.

4.3 Dialogue Control

<u>Centralized vs. Decentralized Group</u>

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.

The designation of a member as sender may be dynamic and change from one member to another as long as only a single member is allowed to send at any given time.

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

Note: It is possible for restrictions to be placed on individual members of the active group on whether they or permitted to send data. That is, in the decentralized group, only a subset of the active members may be authorized to send data.

4.4 Reliability Characteristics

This category incorporates the concepts of reliability and error control as they are applied both to instances of multicast transmission and to the maintenance of multicast group integrity. "Reliability" spans a range from **fully reliable** (in-order error-free delivery is guaranteed to all members of an active group) to **not reliable** (no guarantees concerning the order of delivery or the presence of errors are made.) Reliability may also be expressed with respect to the Active Group Integrity characteristics, which allows for variations such as "fully reliable with respect to a specified subset or an unspecified quorum of the members of the group, don't care beyond that", etc.

4.5 Naming and Addressing

4.5.1 Definitions

(N)-group-name : a name, unambiguous within the OSIE, which is used to identify a set of (N)-entities which identifies the group.

(N)-group-address : a name, unambiguous within the OSIE, which is used to identify a set of (N)-addresses which identifies members of the group.

4.6 Description

The invocation and provision of the (N)-MPCA service makes use of the following:

- an (N)-group-address
- a list of (N)-addresses and/or (N)-group-addresses

Some uses of MPCA involve a transmission to an unknown or partially known group. This can be done using (N)-entity-group-titles or (N)-group-addresses.

4.6.1 Naming

An (N)-group can be identified by its (N)-group-name. The (N)-group-name must be unambiguous within the OSIE. The (N)-group name must be unambiguous to ensure that entities registering for a particular (N)-group can unambiguously identify which group it wishes to join. The (N)-group name identifies only the group and does not identify entities, or the mapping of entities belonging to the group and addresses. The (N)-group-name is used for the purposes of identifying the existence of the (N)-group to the entities in the OSIE.

An (N)-group-name is assigned and managed by an appropriate naming authority under the auspices of System Management.

An (N)-group-connection-id is a name used for identifying specific (N)-group-connections within the scope of a single group. The combination of an (N)-group-name and an (N)-group-connection-id unambiguously identifies an instance of communication within the OSIE. The (N)-group-connection-id must be unambiguous within the scope of the (N)-group-name. The assignment and management of (N)-group-connection-ids is by a designated authority assigned by System Management to an entity responsible for managing the (N)-group.

4.6.2 Addressing

4.6.2.1 Group-Addresses

An (N)-group address is a name unambiguous within the OSIE which is used to identify a set of (N)-addresses that are associated with multi-peer data transmission. When an (N)-group address is used to send an (N)-PDU, all members of the list receive the (N)-PDU. Thus an (N)-group address stands for a list of (N)-addresses.

(N)-group addresses are constructed in the same way as (N)-addresses. An (N)-group address of the upper four layers is constructed from a Network address and (N)-group-selectors. However, there are some additional restrictions on (N)-group addresses and (N)-group selectors. An (N)-ga and an (N)-gs must be unambiguous, at least among the members of the open systems group. This means that the values cannot be determined unilaterally.

Specifically,

- an (N)-group-selector is unambiguous within the scope of an (N-1)-group-address
- each (N-1)-group-address identifies a distinct (N)-entity group.

As a result, an (N)-entity must be able to distinguish between an (N)-group-address and a two-peer (N)-address. There are four ways that this could be achieved:

- coding of MPCA (N)-PCI such that MPCA (N)-PDUs can be distinguished from other (N)-PDUs. Any (N)-address can then be identified as (N)-group address or two-peer (N)-address according to the type of (N)-PDU in which it occurs.
- in some layers, the (N)-entities handling (N)-MPCA protocols may be separate from those handling two-peer (N)-protocols. In this case an (N)-entity will always process the (N)-address according to the type of protocol it handles.
- in some layers, (N)-group-addresses may be encoded distinctly from two-peer (N)-addresses.
- the (N)-entity may have local knowledge of which addresses are (N)-group addresses and which are (N)-addresses.

4.6.2.2 Explicit addressing

For small groups or very short-lived groups, it may not be economical to go through the process of establishing (N)-ga. Thus, explicit lists of (N)-ga or (N)-gs may appear in (N)-PCI. Combinations of (N)-ga's and explicit lists of (N)-addresses at different layers is permitted. Because of the requirement to combine (N)-ga's with explicit lists of (N)-addresses and the fact more than one (N)-selector may be associated with an (N-1)-address (i.e., more than one member of the group in the same system), it will be necessary to provide algorithmic definition or proper syntactic encoding to ensure the proper (N)-selectors are associated with the correct (N-1)-addresses.

When an end system receives an (N)-group-PAI in an (N)-PDU, it must consult a list to determine to which (N)-address(es) the PDU should be delivered. When an intermediate system receives an (N)-group-PAI in an (N)-PDU, it must relay the PDU to those addresses in the list which are reachable from this intermediate system.

4.6.2.2.1 Selection of (N)-group selectors

(N)-group selectors are only unambiguous with respect to their associated (N-1)-group-addressses. It is the responsibility of each system receiving an MPCA transmission to map or translate the (N)-group selectors to whatever is required locally on a layer by layer basis using locally stored directory information and a reference to this specific group derived from recognition of the (N-1)-group-address. This local reference is passed between layers as a service parameter with the SDU.

In the same way that non-MPCA selectors are chosen by the addressed system as a local matter, the assignment of values to (N)-group selectors is local to the combination of systems who are addressed by the associated group network address. The mechanism for agreeing on the values could be one of the following:

- a prior agreement by manual means or using a special administrative application
- layer protocol negotiation

Note that if there were only one multi-peer application allowed per group network address then there would be no need for selectors at all since local tables could define what is required for a given group network address. (N)-group selectors allow for having more than one multi-peer application using the same group network address and also to aid with internal considerations caused by several units making up the end systems. It is also possible through internal tables to have one group-P-selector connect to several end users on this system.

Of course, if lists are used instead of group addresses then group selectors are not involved.

4.7 Ordering

Ordering is concerned with the following two aspects:

1. how PDUs of a single sender are presented to the receivers
2. how a single receiver gets PDUs from the sender(s).

In the case of a single sender, ordering if needed ensures that the data units generated by the sender are delivered to each receiver in the active group in the same order as they were sent. In the case of multiple senders, ordering determines the relative sequencing of data received from multiple sources. The ordering relationship defines the arrangement or interleaving of data from the multiple senders. The ordering relationship can be: local, partial, causal, or total. Note that when there are only two participants in the active group, local ordering, causal ordering, and total ordering are the same.

The properties of ordering apply at the service level and at the protocol level. At the service level, the service provider is required to provide guarantees regarding the order in which SDUs are delivered to the receiving service users. There are applications that do not require all of these (or even any of these) ordering properties: the service provider delivers the SDUs received from the sending service users to the receiving service users in no particular order. There are applications that require the service provider to deliver the SDUs to the receiving service users in the same order or in a different order (e.g., last SDU received by the service provider is first delivered to the receiving service users) in which the SDUs were received by the service provider from the sending service users. At the protocol level, PDUs are ordered to achieve the ordering property required by the service.

The following notation is used to describe the ordering relationships.

Notation: Let `S_k(m)` be the local event of sending data unit m at site `k(=1,2,..,N)`.
Also let `A_l(m)` be the local event accepting the data unit m at a site `l`. The symbol
"∅" is "happened before" relation.

4.7.1 No ordering

No ordering applies when the data units are delivered to the receivers in the active group in no particular order regardless of the number of senders. Particularly in the case of multiple senders, no ordering relationship exists for the arrangement or interleaving of data units from the multiple senders.

Using the notation specified above, no ordering means:

S_p(m) ∅ S_p(m’) =/ A_i(m) ∅ A_i(m’)
for every p, i and for every (m,m’) pair

4.7.2 Local ordering

Local ordering orders the data units sent by a particular sender according to the order they were sent by that sender. Local ordering does not establish an ordering relationship among the data units sent by multiple senders within an active group. Each receiver in the active group may receive the data units sent by multiple senders in a different order, however the data units generated by a particular sender are delivered to all the receivers in the active group in the same order in which the data units were sent by that sender.

Using the notation specified in 6.5.4 above, local ordering is defined as:

S_p(m) ∅ S_p(m’) A_(i(m) ∅A_i(m’)
for all p, i and for all (m,m’) pairs

and

S_p(m) ∅ S_q(m’) =/ A_(i(m) ∅A_i(m’)
for all p, q, i and for all (m,m’) pairs

4.7.3 Causal ordering

Causal ordering orders the data units generated by all the senders in the active group according to the causal dependency relationship among the sending events. A causal dependency relationship is established between two sending events, a and b, if the following applies: (1) a happens before b if a and b are sending events generated by the same sender and a is sent before b, and (2) a happens before b if a and b are sending events generated by two different senders and the data unit generated by the event a by one sender is received by the other sender before it generates the event b. A causal dependency relationship is established among more than two sending events as follows.

If it is established that a happens before b and that b happens before c, then it can be established that a happens before c. A causal dependency relationship cannot be established between the two sending events a and b if it cannot be established that a happens before b and that b happens before a.

S_p(m)∅S_q(m’) A_i(m) ∅ A_i(m’)
for all p, q, i and

- for all (m,m’) pairs that satisfy the temporal relationship: A_q(m) ∅ S_g(m’) when p q, or
- for all (m,m’) pairs where p = q.

4.7.4 Total ordering

Total ordering orders the data units generated by all the senders in the active group according to the absolute timing relationship existing among the sending events. Each receiver in the active group receives the data units from all the senders in the same order.

Using the notation specified above, total order means:

$S\_p(m) \varnothing S\_q(m') \quad A\_i(m) \varnothing A\_i(m')$

for all p, q, i and for all (m,m') pairs.

4.7.5 Partial ordering :

Partial ordering orders the data units generated by all the senders in the active group according to an arbitrary rule. If the data units are ordered according to an arbitrary rule chosen by the service provider and applicable to all the receivers, then each receiver in the active group receives the data units generated by all the senders in the same order. If the data units are ordered according to an arbitrary rule determined by the receiver, then each receiver may receive the data units in a different order.

Using the notation specified above, total order means:

$S\_p(m) \varnothing S\_q(m') \quad A\_i(m) \varnothing A\_i(m')$

- for all i but for some p, q and for some (m,m') pairs, if the arbitrary ordering rule is set by the service provider for all receivers, or
- for some p, q, i and for some (m,m') pairs, if the arbitrary ordering rule is
set independently by each receiver.

4.8 Model of Multi-Peer Communication

4.8.1 Model of Multi-Peer Association

The primary reason for MPCA is the support of Applications needing MPCA services. This leads to the need for the definition and invocation of groups consisting of Application Layer entities. The definition of an Application-group is through the registration of the group definition in the OSIE. At this stage, only the rules for the Application-group may only exist in the data base of the group manager, and the entities comprising the group may not be identifiable. However, any Application-entity that may potentially satisfy the rules for the group may be considered, in a less restrictive sense, as a member of the group. The registration of the (N)-group definition in the OSIE is the initial part of the enrollment phase. Hence, the (N)-enrollment-group is a set of entities which potentially satisfy the rules used to define the (N)-group. In the specific case of the Application Layer, the Application-enrollment-group is the set of Application-entities which adhere to the rules for membership in the Application-group.

As soon as the (N)-enrollment-group is reported to the group manager, an (N)-group-name is assigned to the group. Hence, an (N)-enrollment group can be identified by its (N)-group-name in following actions by potential (N)-entities. The (N)-group-name identifies an (N)-type as defined by the rules defining the group. For example in the case of the Application Layer, the Application-group-name identifies an Application-group whose characteristics are defined in the group rules.

After a (N)-group is defined, potential members may register themselves as members at the group manager by using the group-name already given for the enrollment phase and publicized. Any (N)-entity that wants to receive or send group messages must complete registration. The group formed through registration is call the (N)-registered-group.

It is important to note that, without registration, there is no way for an (N)-entity to receive MPCA messages over an (N)-group-connection. Registration of an (N)-entity creates a binding between an (N)-group-address and the (N)-address of the (N)-entity.

When an instance of communication is desired, an (N)-entity communicates with the group manager to allocate resources to the communication. At this point, an (N)-group-connection is instantiated. The instantiation of the (N)-group-connection creates an (N)-group-connection-id that identifies the instance of communication to the (N)-entities in the (N)-registered-group. A particular (N)-entity may choose to join an (N)-group-connection through the appropriate mechanisms based on the (N)-group-name and (N)-group-connection-id. For the case of the Application Layer, an Application-group-association is created with the complete state information about the Application-group-association available to the Application-entities.

In establishing an Application-association, Presentation and lower layer connections are established.

4.8.2 Model of Communication

There is no restriction on the topology of multi-peer communication in MPCA. The general topology consisting of all active group members being able to send and receive information at the same time, over the same (N-1)-connection is inherent in all multi-peer communication.

The basic component of all multi-peer communication is the (N)-group-connection (or at the Application Layer the Application-association). An (N)-group-connection can generally support the sending and receiving of data at the same time. No inherent restriction is placed on which entity can send and receive.

From the general MPCA model of an (N)-group-connection, restricted cases of communication can be generated. For example, if the (N)-group definition defines the model of communication as a single sender, then the (N)-group-connection may be defined with a data token (like in the half-duplex communication mode). Only the (N)-entity with the data token may send data. Further, this may be expanded to allow for passing the token (roving single sender); or further to allow multiple tokens (up to the number of active (N)-entities.)

Editor's Note: In reviewing some of the contributions from SC 6 (N9161) considerable effort is given in defining restricted sub-categories of (N)-connections that are based on a single sender model. After defining the model for the general, unrestricted model, each of the restricted modes can be generated using group definition rules, AGI, and QoS. In the definition of the architecture, it is important to define the general underlying case rather than the restricted "special cases" to ensure that the special cases do not create future situations which cannot be modeled. As has been described before, the general multi-sender case CANNOT be modeled as a set of single sender connections. The set approach fails to consider the problem with ordering, reliability, QoS, and other state information considerations across single sender connections. It is not possible to co-

ordinate multi-connection data since the characteristics (i.e., failure modes) cannot be completely specified.

However, the restricted single sender case can easily be generated from the general multi-sender case; as can any other special purpose (N)-connection.

5 Conclusions

Multi-peer communication involves a dramatic shift in the way communication is viewed. In a point-to-point (unicast) case, the basic parameters of reliability, sequencing, and QoS are single points in a discrete spectrum. In a multi-peer case, all of the different parameters are continuums based on the perceived needs of the group rather than of individual receivers.