Spatial Scope and Multicast in Large Virtual Environments

Chris Greenhalgh[1]

Technical Report NOTTCS-TR-96-7
Department of Computer Science
The University of Nottingham
Nottingham, NG7 2RD, UK
September 25, 1996

Email: c.greenhalgh@cs.nott.ac.uk

Abstract

This paper considers the use of multicasting and spatial scoping in creating large-scale multi-user virtual environments. The asymmetry of network multicasting with respect to producers and consumers is discussed. Five approaches are described for using multicast groups in systems with spatially scoped interaction: one group per object, one group per world, location-based groups, content-associated groups and dynamic groups. Some illustrative aspects of the design of MASSIVE version 2 are described. Specific strengths and weaknesses of each approach are identified together with appropriate areas of application. Finally some generalisations of multicast-style functionality are suggested.

1. Introduction

Over the past two years I have been developing a multi-user virtual reality systems called MASSIVE [4,5]. Its primary application is tele-conferencing and an important distinguishing feature of it is the use of the spatial model of interaction [1] as a basis for controlling interaction between participants and the virtual world. An important goal of the spatial model, and also of MASSIVE is to support users in interacting with very large virtual worlds, for example with large crowds of people, or large database visualizations. The second generation of MASSIVE is well under way; an important element of the evolution from the first version to this has been the integration of multicast communications to facilitate scaling to considerably more people that could be supported by the unicast (UDP) based first version (predicted scalability of MASSIVE version is reported in [5]). This paper draws out my experiences with integrating a multicast-suitable style into the spatially mediated framework of the spatial model.

1. This work was funded by the UK Engineering and Physical Sciences Research Council through their PhD studentship program.

The most closely related research is that of the Naval Postgraduate School's NPSNET [8] system - a DIS-compliant military simulator capable (currently) of supporting a few hundred simultaneous participants. They have been developing and first published the location-based architecture which is one of those discussed in this document. The DIVE [2] system, version 3, from the Swedish Institute of Computer Science has moved its networking to use multicast but lacks any spatial scoping - it represents the one multicast per world case also discussed here.

The relatively limited take-up of multicasting to date reflects a number of factors: virtual reality research as a whole is relatively immature and a lot of effort is still (correctly) being devoted to more "basic" things than building very large systems; it is very difficult or expensive to get hold of really large number of systems; and network-level multicast technology is itself immature and under constant development, presenting something of a moving target. All three of the systems mentioned above use IP multicasting [3,6] which has been a relatively recent (1989 onwards) and very important enabling technology for the development of multicast-based applications and systems.

The structure of this paper (and my argument) is: to examine the nature and benefits of multicast communication; to argue that it puts greater demands (in terms of resources) on consumers of information than on producers and that as a consequence consumers bear the principle responsibility for selecting the information which they desire and can cope with (rather than the producers selecting appropriate recipients for their data); four types of scoping are identified which can drive and inform the selection of information; the uses of spatial scoping are examined and justified; five approaches to using multicast groups with spatial scoping are identified and compared to identify appropriate applications for each; finally some generalisations of multicast-style functionality are suggested.

Firstly, I consider the nature and benefits of multicasting.

2. Multicast

Multicasting is the form of communication whereby a block of information can be sent in a single operation to a set of destinations. This is contrasted with unicast communication (e.g. as in point-to-point connections) in which a single send operation causes the block of information to reach at most one destination (exactly one destination, if the communication is reliable and neither party fails).

2.1 Benefits of multicasting

Multicasting can be a useful programming abstraction, but it has the additional benefit that bus-based physical communication systems can support multicasting at the hardware level in an inherently parallel manner. In an inter-network in which the routers support multicasting the cost to the sender of multicasting can be virtually the same as a single unicast send: the responsibility for copying and multiple forwarding of the data is assumed by the network which can make use of physical parallelism in bus architectures to perform multicasting efficiently and appropriately for the topology of the network. To create the effects of multicasting with only unicast sends requires that the same data be unicast to each destination individually. This has two disadvantages compared to the network multicast case:

- the sender's effort has an $O(N_{\text{destinations}})$ component relating to the repeated sends; and
- the network load (on shared segments and routers) is increased, especially close the sender, since the same data must be handled as many as $N_{\text{destination}}$ times in the same place. With multicasting, the data should only be present on any network segment once (and should arrive at each router once).

The argument in favour of using network multicast is compelling, especially as the number of destinations increases.

2.2 Costs of multicasting

As usual there is no such thing as a free lunch. The costs of multicasting are more subtle that the benefits (and unlikely to cause an outright rejection of multicasting). The costs are the control messages to handle multicast routing and group membership, and the group-related state which must be maintained by the multicast routing nodes. For reliable multicasting there will also be some overhead to ensure reliability in the form of acknowledgements and resend requests. For a multicast group this reliability overhead will be at least $O(N_{\text{destinations}})$.

2.3 The asymmetry of multicasting

Physical multicasting makes being a producer very inexpensive; more specifically it makes being a big producer (with a large number of consumers) no more expensive than being a very small producer. However the cost to each consumer is unchanged.

Current networking technologies generally place similar limits on both the rate at which information can be sent and the rate at which it can be received. With unicast communications this is appropriate: the quantity of information sent is the same as the quantity received (ignoring lost packets). With multicast, a consumer which wishes to draw on a number of producers must divide its receiving capabilities between them, allowing each a fraction. But a producer which wishes to send identically to a number of consumers can use all of its sending resources and allow the network to copy it to each consumer. So it will be easy to multicast a description of a virtual world which will be too complicated to render at any satisfactory speed.

It is this imbalance of power between producer and consumer as a result of multicasting which places the onus for scoping on the consumer of information. This also echoes, as noted before, a characteristic of computers which might be informally summed up as "computers are much better at copying than at thinking". A fundamental concept of the spatial model of interaction

[1] is that awareness (the measure of interaction) should be derived by *negotiation* between the producer and the consumer. However, while this is important at the application and interaction level, it is not directly reflected in multicast-based communications.

3. Information in a virtual environment

As a foundation for the rest of the discussion of multicasting and spatial scope I introduce here some of the principal characteristics of shared objective virtual worlds. This paper considers only *objective* virtual worlds, i.e. worlds in which the same information is (in principal) available to all participants; this is the style adopted in almost all current VR systems and is in accordance with peoples' experiences of the everyday world. However subjectivity may be important in some situations (see for example [10]).

Normally a virtual world is composed of a number of identifiable objects or artefacts. These typically have a location, a visual appearance, an audible character, and may have other characteristics as well such as an associated video image or information for tactile feedback. The artefact and each of its characteristics are represented or encoded as values of some appropriate and agreed types. Some of this information reflects persistent state while some represents transient events. An artefact may be decomposed into a number of information sources, one for its location, another for its appearance, and so on (i.e. applying functional scoping). Each information source has a double character: viewed at any instant in time it has some current state; but viewed over a period of time it is seen to emit transient messages. Some of these transient messages describe changes in the persistent state and so are reflected by the state while others have no persistent reflection (unless it is within those who view the event).

It is characteristic of objective virtual worlds and also one of the most significant characteristics of computers in general that any number of copies can be made of this information and distributed to other computational processes or objects. In virtual world terms, we expect each artefact to have many simultaneous viewers and each viewer to view many artefacts simultaneously.

4. Scoping

In an ideal world, with unlimited network bandwidth and computational resources a computer (which may in turn represent a human user) would have simultaneous access to all available information, and be able to organise that information to extract all relevant material. So every viewer of a virtual world would be fully aware of every artefact in all its forms (and, indeed, of every virtual world). However, the everyday world is somewhat less than ideal in this respect. Computers (and their users) must, of necessity, lower their aspirations. The greatest single challenge which this raises (especially in the light of the asymmetry of multicasting) is of consumers using the available finite resources in the best way: finding the subset of the available information which is most relevant and useful. To this end we employ the concept of scoping. The scope of a consumer of information describes a (logical) boundary around the information which they wish to obtain and which they believe (by their own deduction or from recommendations) to be the most desirable. Developing from the list in [7] (due to Macedonia et al.) I highlight four kinds of scope; these are briefly described below.

4.1 Worlds

Many virtual reality system (e.g. DIVE, AVIARY [9], MASSIVE) have a concept of multiple worlds. Each world is a disjoint and notionally infinite space in which artefacts may be situated and viewed. A user typically interacts with a single world and has no awareness of artefacts in other worlds. This is perhaps the most basic form of scope used in virtual environments. In non-spatial terms it might correspond to disjoint domains of activity, e.g. in independent companies. It may also correspond to more semantically defined scopes, for example, a news group, which is defined by a common interest and has hard boundaries.

4.2 Functional

The second kind of scope is functional. To create and use information in a digital context requires knowledge of its type (structure and meaning). In a virtual world we can identify many different types of information which are independent. For example, the graphical appearance of objects is described in one way and is only relevant to graphical renderers, while the audible character is described in a different way and is relevant only to audio-capable processes.

4.3 Quality

Even within the same world and for the same purposes information can be made available with a range of qualities of service. For example: different levels of detail for object geometries, different resolutions and frame rates for video, different requirements for reliability, timeliness and consistency. So it may be appropriate to make the same underlying data available in a range of qualities which viewers can chose to suit their requirements and resources.

4.4 Space

The final kind of scoping and the focus of this paper is spatial scoping, i.e. scoping which relates to the spatial relationships between artefacts and viewers in the virtual world. This is developed in more detail in the next section.

5. Ideal spatial scope

The nature of virtual environments, analogous to the everyday world, introduces three factors which can be exploited in the selection of spatial scopes for information and interaction:

- rendering and perceptual limitations, e.g. the use of perspective projections results in more distant objects being seen with less detail and the field of view is typically limited;
- analogy with distance attenuation in the everyday world - Electromagnetic radiation (light, radio, etc.) drops off distance as $1/d^2$, and imperfect media (such as air) introduce noise; and
- interference between artefacts, e.g. a closed box (or a room) may cut off the artefacts within in from the artefacts without.

These factors tend to make perceptual and interactional spaces relatively compact and structured, and allow for the progressive abstraction (at decreasing levels of detail) of more distant artefacts. The first two factors suggest scopes based on distance from the located observer, possibly with some consideration of the direction in which the viewer is facing. The third factor introduces a direct dependence on the current contents of the virtual environment within this larger scope.

Current systems which use such spatial scoping include: MASSIVE [4], in which artefacts' auras define spherical scopes; AVIARY [9], in which the visual renderer scopes its interaction with the world through a projection of the clipping volume (plus a pre-fetch margin) in the virtual space; and NPSNET [7] which is developing a system with approximately circular scopes based on the real-world characteristics of military engagement.

When considering a user perceiving and interacting with a virtual world the ideal use of spatial scoping would ignore (exclude) only artefacts which could not be perceived or influenced in any case and would view artefacts at reduced qualities (e.g. lower level of detail) only when such reduction was hidden (indistinguishable) because of limitations of the renderer or the user's perceptual system. In practice, this may not be a sufficiently small scope, and resources must be allocated only to the more significant or important subsets of this ideal scope. This will result in visible (or audible, etc.) discontinuities as artefacts enter and leave the scope within perceptual range.

In the sections which follow I look in detail at the ways in which spatial scoping can be realised.

6. Spatial scoping

To implement spatial scoping the system must translate a viewer's ideal scope of interaction into the corresponding set of artefacts and make use of appropriate communications facilities to monitor those sources/artefacts as illustrated in figure 1.

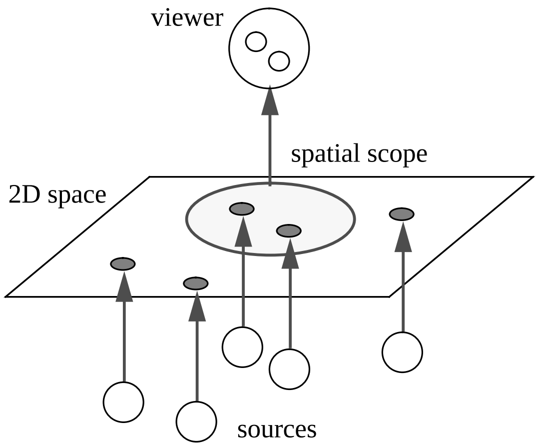

FIGURE 1. illustration of spatial scope in 2D

Based on the discussion so far it is clear that the appropriate communications facilities will be some form of multicast, conveying information about a source to all interested parties. Based on the factors determining spatial scope in section 5 we can identify five approaches for relating sources (artefacts) to multicast groups:

- each source can have its own multicast group;
- all sources can share a single multicast group;
- space can divided up independent of content with sources allocated to multicast groups according to location;
- multicast groups can be associated with objects in the space (e.g. rooms) to reflect their structuring of the space; and
- multicast groups can be added and removed dynamically in response to the movement of sources (e.g. picking out large groups of objects and giving them new multicast groups).

In all cases information relating space to sources and thence to multicast groups must be available to the viewer or to some process or processes acting on their behalf. Also, where a source's multicast group can change as a result of movement (all bar the first case) there must be a mechanism whereby sources are made aware the need for (and given the information to make) such changes.

6.1 Implementing spatial scoping

We will consider the principle events in a spatial scoping system to assess the benefits and costs of these various approaches.

Each source is located within the shared spatial frame. A multicast group includes one or more sources. Depending on the approach taken, allocation of sources to multicast groups may be based on source identity or on its location in space. In all cases the multicast group is associated with a region of space, either directly, because it is defined that way (and sources are allocated to it accordingly), or indirectly, because the sources which are allocated to it are located in space. A viewer's ideal spatial scope is defined in terms of the common spatial framework of the world, possibly adjusted according to local artefacts such as rooms and walls. To be aware of all of the sources within its scope the viewer must join (as a consumer) all of the multicast groups which overlap with, contain or are contained within its spatial scope.

As viewers move, their spatial scopes moves with them and so they move out of the influence of some multicast groups and into the range of others. Equivalently, as sources move their multicast groups may move, leaving some spatial scopes and entering others, or their multicast groups may remain fixed in which case the sources must move into new multicast groups appropriate to their location.

6.2 Costs of spatial scoping

There are three main events in the process of spatial scoping. These are outlined below in the context of the demands they make on the system and the parties involved.

6.2.1 source/scope collision

The cost of a source falling within a new spatial scope is that the source must send its current state to the new viewer which must receive and handle it. Where this occurs frequently for a single source this cost can be mitigated (for the source) by collecting new viewers over some (relatively short) time period and multicasting the state to them all. This approach is adopted in MASSIVE version 2 (see section 8)

6.2.2 source/multicast collision

The cost of a source moving to a new multicast group is that of being made aware of the group and of the need to change, of leaving the old group and of joining or forming the new group and of announcing its current state (via the group) to any viewers already present. For single-source multicast groups this costs is incurred only when the source is first created. For the other approaches it is incurred repeatedly as source move relative to multicast groups. Note however that static objects (typically buildings, scenery) will incur no repeated cost on this basis.

6.2.3 scope/multicast collision

The cost of a spatial scope overlapping a new multicast group is that the viewer must join the multicast group as a consumer. It may well need to know what sources are within the group. Its effective scope may also now be considered to include the whole of the region covered by the group, in which case all of the sources in the group would need to send the new viewer their current state (as above).

In all approaches mobile viewers will cause scope/multicast group collisions. With mobile multicast groups (single group per source or dynamic groups) collision may occur even for stationary viewers.

7. Evaluating the alternatives

Having identified the components of spatial scoping I now use these to assess and compare the effectiveness of the five alternative approaches introduced in section 6 and to identify the situations in which they are appropriate.

7.1 One group per source

When each source has its own multicast group source/scope collisions and scope/multicast collisions obviously coincide. The advantage of this approach are that (unlike the others) a viewer's effective scope (describing the sources they are in touch with) is the same as their ideal scope. I.e. a viewer is never left with sources they are not interested in just because they share the same multicast group as sources which they *are* interested in. As a result, the number of sources in the effective scope and the rate at which they change are both the minimum possible for that ideal scope.

The difficulties with this approach are the need to consider all sources and all viewers individually when checking for overlap, and the need for one multicast group per source (two if a separate group is to be used for multicasting state data to new viewers). This would not scale well to very large systems. The handling of multicast groups is more demanding on routers than unicast addresses because there is no hierarchical organisation of multicast groups tied to sub-networks as there is with unicast addresses: the point of multicast groups is that they are destination independent, but that limits the number of groups which a network can support.

The other approaches are all different approaches to sharing multicast groups between sources so that fewer multicast groups are needed for a given world. They trade a certain amount of selectivity for (typically) 1-3 orders of magnitude reduction in the number of multicast groups required.

7.2 One group per world

With this approach there is really no spatial scoping at all within a world: all sources send on the same multicast group and all viewers monitor it. Joining and leaving occur only as sources and viewers enter and leave the world. This clearly requires that the viewers be able to handle the aggregate traffic for the world. If they handle the data then this approach is appropriate and efficient. However there is nothing that can be done when the world grows too large (except leave). So this approach is best used where world size can be limited and enforced. This is the approach used in the current versions of NPSNET [8] and DIVE version3.

7.3 Location-based groups

With this approach the world is divided into areas which tile the world. Sources belong to a single multicast group based on their location, and move between multicast groups as they cross the boundaries. Viewers monitor each multicast group which overlaps their scope; this overlapping set changes as the viewer moves about. The situation is effectively the same as taking all the sources within a region and placing them together in the centre of the region, and then increasing the size of the viewers scope to compensate for the spread of the multicast regions. So, as with all shared multicast group schemes, the effective scope of each viewer is extended and the viewer is made aware of a number of sources which they are not actually interested in. For example, a visual renderer in a virtual world would be required to render additional irrelevant (or at least relevant) artefacts.

Making the groups larger makes use of fewer multicast groups but at the expense of increasing the effective scope of all viewers and therefore reducing the supportable ideal scope. Also, a given viewer will be able to support fewer larger groups at once, so that each change of multicast group will have a greater impact producing less continuity of experience.

As well as reducing the number of multicast groups required the other advantage of this and other shared multicast group approaches is that collision detection need only be done between sources or scopes and the multicast groups, rather than against each of the scopes or sources within it.

To use this approach it must be possible to pre-determine a suitable tile size according to the expected (and characteristic) use of the world. This is the approach under development for a future version of NPSNET [7]; being used for a specific application (military simulation) with interaction governed by vehicle capabilities taken from the everyday world it is possible to determine a suitable cell size in advance. However this may not be possible where the use of the space is subject to unanticipated change, such as extensive growth or elaboration (building). It may also be inappropriate when the interaction style is less constrained by the characteristics of the real world. For example, if users are able to suddenly take off and fly at any speed they like then a general choice of cell size will be difficult.

7.4 Content-associated groups

This approach is similar in many ways to fixed tiled approach above, except that the groups are related to the spatial organisation (the architecture) of the virtual world. For example, a room might have its own multicast group, reflecting the tendency for interactions within closed spaces to involve the contents of the space, and to be noticeable throughout that space.

The world description will need to provide some explicit cues to the distribution system about the suitability of a region for use as a group whereas a location-based approach can be overlaid on an arbitrary space. Also, clearly, this approach cannot be used where there is no appropriate structural content to the space. So unless we force worlds to comprise only connected rooms or similar spaces than this approach can only be used in addition to the next or the previous one. This is the main style of organisation being developed for MASSIVE version 2 (section 8)

7.5 Dynamic groups

The previous two shared-group approaches were static and unchanging. This final approach is dynamic, responding to the disposition of sources and viewers within the space. One example of such an approach would be analogous to an adaptive subdivision scheme as used in the formation of a spatial quad-tree to structure the objects in the space. So a world might start out with a single multicast group. Then as the population increased that group could split into multiple groups, each covering part of the space (as with fixed tiled groups). If one of these groups became crowded in turn then it could split again (c.f. proposed EDB behaviour in AVIARY [9]).

This approach may be considered as the complement to the fixed-tiled approach for worlds which are unpredictable and changeable, and not subject to constraining “physics”. A location-based world may be designed within its constraints to be less distinguishable from an infinite spatial scope (which is in some sense the ideal). With dynamic groups on the other hand, “paging” artefacts are likely to be more apparent as new multicast groups come into scope or are split and recombined; also the effective scope is more likely to vary in size, but it will be easier to limit the number of source viewed, even in chaotic situations, so that the performance of the viewer (and hence the human user’s interface) can be maintained at an acceptable level.

8. MASSIVE version 2

Having considered the broad range of alternative approaches which may be adopted this section briefly outlines some of the pertinent aspects of the design of MASSIVE version 2 to illustrate some of the foregoing discussion and in the hope that other researches will find food for thought.

MASSIVE 2 supports two styles of communication: connection-based communication as in MASSIVE 1 which include RPCs, attributes and events; and reliable multi-sender multicast groups which support attributes and events (but not RPCs). Multicast groups are used to form virtual worlds, as discussed above. The multicast groups are implemented over (unreliable) IP-multicast, with each reliable group making use of three IP multicast groups: one group carries the normal transient messages to all viewers, one group is used to multicast initial state to new viewers and one group is used for group management and to support reliable delivery.

The multicast sources have an explicit model of state (a set of attributes); when a new viewer joins a MASSIVE multicast group it listens on both the update IP group and the state IP group. After some delay (to allow other new viewers to join) each source multicasts its current state, reliably, on the state group. Once a viewer has received the current state for all sources in leaves the state IP group so that it doesn't see those messages when subsequent viewers arrive.

The organisation of multicast groups in MASSIVE 2 is a mixture of world-based, location-based, and content-associated approaches. Each world has a multicast group. Within this there can be other multicast groups which cover sub-regions (location-based). Sub-region multicast groups can also be associated with an artefact in the super-region; for example a box (in one region) might enclose its own sub-region (with a separate multicast group) which includes the contents of the box. Except for having no super-region, a world acts like any other region. So regions can be organised into arbitrary hierarchies. The regions are self-describing, allowing viewers and sources to more between regions in a consistent manner.

The use of a hierarchy of multi-source groups allows an additional technique to be used to enhance scalability: group-abstraction. Consider a dynamic group which is formed around a spontaneous gathering of people; from a distance a largely undifferentiated crowd is visible. The group itself could abstract from the sources within it to a high-level view of its contents - a kind of group level-of-detail - and express this group-view as an artefact in its superior multicast group. Then a viewer could take only the abstract group-view, or might choose to "open up" (join) the group to get full details on its constituents.

9. "Aggrecasting" and active network applications

Before concluding I would like to develop the subject of the asymmetry of multicasting. I suggest that multicasting is just one example of a *class* of application that naturally live within the fabric of the network itself.

Multicasting - sending the same thing to many destinations - is an example of an activity that is very sensitive to the complications of networks: which links are really there and how fat and how fast they are. It is also concerned with distributing information, and as such is an ideal candidate for distribution itself. So the multicasting facility is spread across the whole network (in the routers) where it has direct access to links, and can watch data on its way. The end result is that multicasting data becomes a much less painful activity.

But there is no reason why such distributed activity should be limited to duplicating the input as with multicast. One can imagine, for example, an "aggrecast" - the opposite of a multicast; instead of taking one input and producing several outputs an aggrecast would take a number of inputs and generates a composite output. For example, statistics from all nodes could be combined as it flowed through the network in a natural and appropriately parallel manner; this would be capable of scaling indefinitely with the network, since extensions to the network need new routers and those routers exactly handle the additional work.

Responses (e.g. acknowledgements or replies) to multicast messages are a prime candidate for distributed collation within the network. Such a facility could give reliable multicasting at a very similar “cost” to current unreliable multicasts.

In the context of virtual reality the network functionality could be a little more complicated. Each artefact could make available a much larger amount of information about itself (at a range of levels of detail) that would normally be appropriate for multicasting. At the network nodes, all of the inputs could be assessed and selected at the appropriate levels of detail to fit with the constraints of the various outgoing links. This would be continuing simultaneously all across the network. The net effect would be to tailor the delivered information according to the capacities and characteristics of the intervening links. So for example, machines sharing a high-capacity network would get plenty of detail (e.g. textures, video) from each other without that detail swamping the narrower links of the network.

I describe these types of application as “active network applications” since they live within the network itself and performs a more active role than is normally visible from routers and the like. Three major stumbling blocks for such a scheme are: security - allowing new applications to migrate onto critical network facilities; standardisation - all routers in the network may need to support them; and resources - routers don’t often have much effort to spare. Perhaps specific standardised functions could be added, such as the statistics gathering example above or selective forwarding.

In many ways the hierarchy (also a network) of regions within a world in MASSIVE 2 could act in the same way, filtering and abstracting the information within a region (multicast group) for external consumption and also representing the “outside world” within the group. Or perhaps the most productive deployment of such a scheme would be in a special purpose parallel computer: a world on chip, anyone?

10. Summary

In this paper I have discussed the nature of multicast communication and its undoubted savings when sending identical data to a number of destinations as is the normal pattern in objective multi-user virtual environments. I have also argued that it puts greater demands (in terms of resources) on consumers of information than on producers. As a consequence of this, the consumers bear the principle responsibility for selecting the information which they desire and can cope with (rather than the producers selecting appropriate recipients for their data). I have outlined four types of scoping which can drive and inform the selection of information: world, function, quality of service and spatial location.

I have developed the uses of spatial scoping and justified it on the basis that interaction in 3D space tends to be localised because of: perceptual and rendering limitations (field of view, resolution); attenuation of energy and information over distance; and the interference of objects (close objects concealing distant or contained objects).

Given that multicasting is an appropriate choice for an objective virtual world with multiple observers (both people and applications) I have identified five approaches to using multicast groups with spatial scoping. First:

- Multicast group per source (with at least one source per artefact) produces little wasted information but requires the most demanding collision management and the most multicast groups, with no way to reduce this number.

The other four approaches are all alternative approaches to sharing multicast groups between numbers of sources. They all approximate the ideal spatial scope of viewers (introducing additional sources), thereby reducing the useful capabilities of a viewer. On the other hand they require fewer multicast groups and less demanding collision detection (based on multicast groups which each represent potentially many sources).

- Group per world is the basic scope-less case, which is efficient for small worlds but does not scale to large worlds.
- Location-based groups are seen to be suitable for real-world derived situations where density is limited, movement constrained and the scope fixed in size (distance).
- Content-associated groups require explicit support from the creators of the virtual world but can make use of the structure of the world's contents to form "good" groups, i.e. groups which are likely to be wanted as a whole and introduce little unwanted information.
- Dynamic or adaptive groups are seen to more resilient to unpredictable restructuring or use, and to less "realistic" interaction, but at the expense of additional complexity and a greater likelihood of the effects of scoping being visible.

Clearly, there is no trivial choice between alternatives: each is appropriate in different situations. Real experience is needed with these different approaches (e.g. NPSNET, MASSIVE 2 as described in section 8). As these implementations (and others) mature their applicability can be more rigorously determined.

Finally I suggested some generalisations of multicast-style functionality which invite further development.

11. References

1. Benford, S., Bowers, J., Fahlén, L., and Greenhalgh, C., "Managing mutual awareness in collaborative virtual environments," in the *Proc. of the ACM conference on Virtual Reality Software and Technology* (VRST'94), Singapore, August 1994, ACM Press.

2. Carlsson, C., Hagsand, O., “DIVE - A Multi-User Virtual Reality System”, in Proc. VRAIS’93, IEEE Virtual Reality Annual International Symposium, pp. 394-400.

3. Deering, Stephen, “Host Extensions for IP Multicasting”, IETF *RFC 1112* (Aug. 1989).

4. Greenhalgh, C., and Benford, S., “MASSIVE: a Distributed Virtual Reality System Incorporating Spatial Trading,” in Proc. *IEEE 15th International Conference on Distributed Computing Systems* (DCS’95), Vancouver, Canada, May 30 - June 2, 1995, IEEE Computer Society.

5. Greenhalgh, C., and Benford, S., “MASSIVE: a Collaborative Virtual Environment for Tele-conferencing,” to appear in *ACM Transactions on Computer-Human Interaction* (ToCHI), in press (autumn 1995).

6. Macedonia, Michael R., and Brutzman, Donald P., “MBone Provides Audio and Video Across the Internet”, IEEE *Computer*, April 1994.

7. Macedonia, Michael R., Zyda, Michael J., Pratt, David R., Brutzman, Donald P., and Barham, Paul T., “Exploiting Reality with Multicast Groups: A Network Architecture for Large-scale Virtual Environments”, in *Proc. 1995 IEEE VRAIS’95*, 11-15 March, 1995, RTP, North Carolina.

8. Macedonia, Michael R., Zyda, Michael J., Pratt, Donald P., and Barham, Paul T., Zeswitz, Steven, “NPSNET: A Network Software Architecture for Large-Scale Virtual Environments”, in *Presence*, Vol. 3, No. 4, Fall 1994, pp. 265-287, MIT Press.

9. Snowdon, D., and West, A., “AVIARY: Design Issues for Future Large-scale Virtual Environments”, in *Presence*, 3(4), 1994, MIT Press.

10. Snowdon, Dave, Greenhalgh, Chris, and Benford, Steve, “What You See Is Not What I See: Subjectivity in Virtual Environments”, to appear in *Proc. FIVE’95*, 18-19th December 1995, Queen Mary & Westfield College, University of London, UK.