DUAL–MODE MULTICAST FOR DIS

J. Mark Pullen Elizabeth L. White

Center for Excellence in
Command, Control, Communications and Intelligence
George Mason University

keywords: communications architecture, multicast, scalability

Abstract

To support large-scale DIS exercises effectively, scalable approaches to supporting information flow among the individual distributed simulators must be found. The issue of interest here is whether DIS should be supported by a single homogeneous mode of multicast across all participating LANs and WANs (single-mode multicast) or by separate multicasting schemes on LAN and WAN.

We are currently investigating an approach called Dual-Mode Multicast that uses multicast at the global (WAN) level on a group-per-exercise basis to minimize WAN traffic, while using multicast at the local (LAN) level to limit the number of PDUs delivered to each simulation. This combines the benefits of multicasting with simplicity at the global (WAN) level of the system and is supportable under existing and projected COTS multicast WAN technologies. To evaluate this approach, we have constructed various scenarios and determined PDU rates in various parts of the system for these scenarios. These scenarios have given us insights into the computational demands of this technique, as well as previously described multicast approaches.

1 Introduction

The primary mission of DIS (Distributed Interactive Simulation) is "to define an infrastructure for linking simulations of various types at multiple locations to create realistic, complex virtual worlds for the simulation of highly interactive activities."[DIS94] In the initial architecture to support DIS, developed under the SIMNET program, all simulation nodes received all events broadcast by the other nodes in the simulation and each receiving node was responsible for sorting out what information is relevant to the entities it supported. While this approach was adequate for smaller distributed simulations, as the number of entities increased, the traffic has increased beyond the available bandwidth and therefore this approach has not proven to be scalable. The SIMNET approach has been expanded into a worldwide wide-area network, the Defense Simulation Internet (DSI), to serve as a testbed for distributed simulation. However, this expansion initially used the same broadcast approach and placed unacceptable demands on both the WAN and the individual simulators as the number of simulators grew. In particular the bottleneck represented by encryption for secure exercises permits only static multicast on the WAN. As result, scalable approaches to reducing WAN traffic and simulator input have been sought.

Figure 7 shows what we envision a large exercise looking like. The claim has been made [HCNF94] that even if it were possible to transmit all of the data generated to each participating simulator, 99% of the information would not be used by the receiver. However, simply limiting the amount of data that is sent to the WAN is not a good solution with the increasing presence of wide-area and global viewers associated with exercises. These factors motivate recent approaches to handling issues of scale in DIS environments by implementing a number of information flow management techniques in application gateways[1] (AG) at the LAN/WAN boundary of each participating network(see Figure 1). In terms of scalable

[1] Or in processes associated with the gateways[CaHo94].

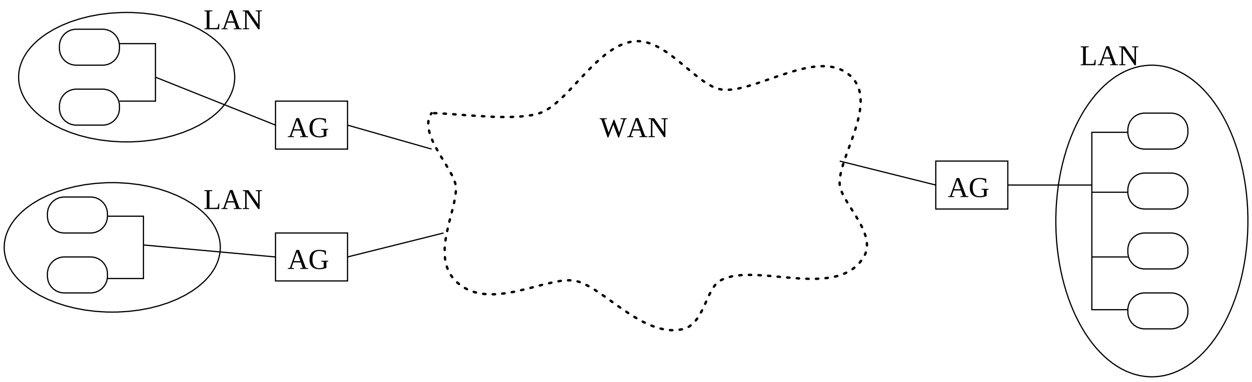

Figure 1: Application Gateways (AG) connect LANs to the WAN.

architectures, the information management technique of greatest interest currently is geographic filtering based on multicast technology.

In this paper we contrast previous geographic filtering approaches based on using multicast technology to support filtering algorithms. Dual-Mode Multicast is an approach that uses multicast technology to support a geographic filtering approach while overcoming some of the problems associated with previous proposed multicast architectures.

2 Geographic Filtering

Geographic filtering is a type of relevance filtering designed reduce the amount of unnecessary traffic received at each simulator. The idea behind geographic filtering is that each entity has some region of interest based upon the inherent sensor limitations of the entity such as viewing range or available sources of input (radar, radio, human vision, etc). Geographic filtering algorithms use information about the limitations to decide where to send messages. Geographic filtering is implemented using multicast technology. The overall terrain is divided into regions, each associated with a multicast group. An entity 'registers' interest in the multicast groups associated with any regions in which it needs to know about events. When an event occurs at some entity located in some square, the information associated with the event is multicast to the group associated with that square. Using this technique, message traffic to the individual simulators can be reduced significantly. Regions may be assigned statically (grid squares or hexagons) or dynamically (clustering). The basic goal is to partition the battlefield such that any entity need listen to at most a few groups to obtain all information within the range of its sensors.

There are several issues that must be considered. Keeping in mind that entities that are "close" in terms of the simulation can be distributed geographically and communicate across a Wide Area Network (WAN), these multicast classes would be at the global level. For a distributed simulation, this may mean maintaining a large number of multicast groups. For a exercise where the terrain is divided into a 30 X 30 grid, there would be 900 different groups. Finer grain divisions could decrease the traffic at individual simulators, but would increase the total number of multicast groups in the exercise. Additional multicast groups (at the WAN level) would exist if there were multiple active exercises.

The DSI currently supports only a small number of static multicast groups, reducing the effectiveness of geographic filtering schemes. An additional issue is managing the assignment of individual simulators to multicast groups. LAN adapters typically support at most 16 groups with hardware filtering; software filtering in the simulators would reduce the benefits of multicasting. Therefore a multicast manager would be needed at each site, to monitor all traffic and coordinate assignment of simulators to groups. Managing the large number of multicast groups at the WAN level is likely to prove difficult.

Another issue associated with multicast geographic filtering is the impact on the amount of bundling that can occur in the system. Some encryptors (including the NES currently used by the DSI) have a limitation on the total packets processed per unit time. Previously, overhead associated with sending (or broadcasting as in the SIMNET architecture) a message to another WAN node was decreased by sending out messages to that node in groups or bundles. However, because the bundling would now have to be done on a per-multicast-group basis, the performance gains associated with bundling of messages would decrease.

Geographic filtering, as previously proposed, tackles the problem of decreasing message traffic at the individual nodes but introduces additional complexity at the global (WAN) level. Given what seems like a very large

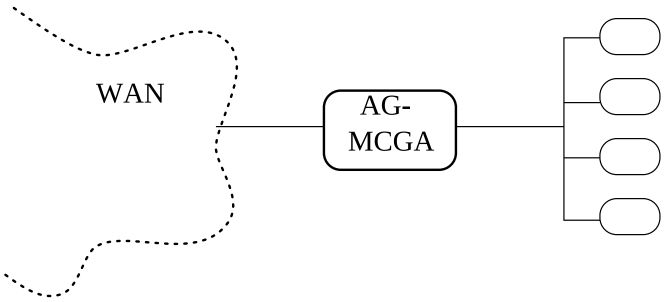

Figure 2: A MCGA is responsible for mapping between the local (LAN) multicast groups, and the global (WAN) multicast groups.

number of multicast groups possible, we are concerned with the feasibility and efficiency of this approach. We believe a global architecture of simplicity comparable to the original SIMNET architecture is required.

3 Implementing Geographic Filtering using Dual-Mode Multicasting

We are currently investigating an approach called Dual-Mode Multicast that combines the benefits of implementing geographic filtering using multicast technology with simplicity at the global (WAN) level of the architecture. What we propose is to use multicast at the global (WAN) level on a group-per-exercise basis to minimize WAN traffic, while using multicast at the local (LAN) level in a manner similar to that described in the previous section, to minimize simulator output.

Just as the previous approach describes an agent that handles the details of the global multicast scheme, we can have an agent per LAN called a MultiCast Group Agent (MCGA) to map between the global multicast layer and the local multicast. This MCGA is associated with a particular exercise and has knowledge of the overall geographic-based scheme and the entities associated with simulators on the LAN. We envision the MCGA to be responsible for receiving all messages associated with a particular exercise and deciding which (if any) nodes on the LAN need to receive each message based on the limitations of the entities associated with the individual simulations. Any messages that arrive and are not needed by any local simulators can be filtered out, or discarded by the MCGA.

At a simulator, event messages are sent by multicasting "to the exercise" (after going through a process of culling[2]). Just as with the previous approach, each entity has a set of multicast groups that it is interested in. Incoming event messages to a node are received as coming from the local or geographic-based multicast groups.

Figure 3 summarizes the algorithm at the application gateway in single-mode multicast and at the MCGA in Dual-Mode Multicast. Both approaches have the same estimated traffic on the LAN, the WAN and at the individual simulators. However, as described previously, the single-mode multicast approach requires support for a large number of multicast groups at the WAN level. An additional similarity can be found the the computational demands of processing entity-state PDUs to determine the appropriate multicast group to be associated with each PDU.

In addition to the advantage of simplicity at the global level, an advantage of the dual–mode multicast approach is that, unlike single-mode multicast approaches, it enables the system to bundle messages on a per-exercise basis. Although bundling is not advantageous in all network technologies (ATM, for example), overhead can be reduce in this manner in some cases.

There are two additional points to note about dual–mode multicast:

- The approach does not preclude other information management techniques such as compression, overload management, and delegated representation.
- The approach is scalable to multiple LANs or subsites, to the degree the MCGA can be placed in a scalable computing base such as a multiprocessor.
- The approach requires all major exercise sites to receive all exercise traffic at their AG/MCGA. However, smaller sites can be connected to the AG/MCGA of a larger site with lower bandwidth links.

However, dual–mode multicast does not come without a price. In a single-mode multicast approach, an application gateway only receives and processes PDUs for multicast groups that are needed on its LAN. In contrast, the MCGA for dual–mode multicast receives all PDUs sent to the WAN, many of which will not be forwarded to the MCGA's LAN.

An important question in establishing the practicality of dual-mode multicast is the cost of hardware to implement. While experimentation with actual hardware will be necessary, an initial estimate of the requirement can be formed as follows. In addition to any other role it has as an application gateway, the system housing the multicast gateway will need to filter each incoming DIS PDU on a basis of distance from virtual entities on the LAN (or at any satellite site), and form IP multicast packets

[2] We mean culling in the same sense as used in the papers referenced here. This process removes messages that can be determined to be of local interest only and messages that are outside the domain of the exercise.

Message from:	Single-Mode Multicast	Dual-Mode Multicast
LAN	Map to local MC group and multicast to WAN	Multicast to WAN using exercise MC group
WAN	if sent by this AG, discard else multicast to LAN	if sent by this AG, discard else map to local MC group multicast to the LAN

Figure 3: Algorithm at the Application gateway in Single-Mode Multicast and at the MCGA in Dual-Mode Multicast.

for them. In very large exercises these functions can be allocated to different processor boards. Based on the small fraction of WAN traffic that is transferred to the LAN in our analysis, we believe the most computationally demanding part of this effort is filtering. Our estimate is that in a 50,000 entity exercise at most 30,000 DIS PDUs per second would be presented to the filter, including the current DIS "heartbeat" minimum PDU rate. Assuming a grid filter algorithm is used, with each grid square labeled as "needed locally" or "not needed locally", a simple filtering algorithm would require less than 100 machine instructions per PDU to implement. This level of computation is well within the capacity of a single processor within multiprocessor platforms such as the Sun Sparc 20 or SGI Challenge. Such systems are affordable to support major sites in a very large exercise. In smaller exercises a uniprocessor workstation such as Sun Sparc 5 or SGI Indigo2 could potentially support the entire AG/MCGA function, interfacing between any available WAN multicast protocol and the IP multicast LAN.

4 Scenarios

To further evaluate this approach, we constructed scenarios and determined PDU rates in various parts of the system for these scenarios. By the term scenario, we mean some mapping of N entities onto a terrain grid. We considered three different N values for our scenarios: 500, 5000, and 50,000. Each different size of scenario has different assumptions associated with all of the related scenarios. For the smallest value, $N = 500$, we assumed all entities could be active in our scenarios. The larger scenarios had smaller activity rates, 50% and 20% for $N = 5000$ and 50,000 respectively. We also started with a given number of sites and distribution of entities across those sites. These assumptions are outlined in Figure 4. Assumptions associated with all scenarios regardless of size included the size of the grid regions (5 KM x 5 KM), the maximum number of entities in a grid square (100 per square KM), and the smallest division of entity groups (50 entities). For each of sizes, we developed many different scenarios and computed PDU rates for the various parts of the system. We were particularly interested in the parts of the system where single-mode multicast and dual–mode multicast differ.

As noted earlier, one difference between single-mode multicast and dual–mode multicast is the number of PDUs arriving from the WAN to be processed by the MCGA. In dual–mode multicast, the MCGA receives all entity-state PDUs associated with the exercise. These PDUs must be processed and the PDUs that are not of interest to any of the local simulator are discarded or filtered by the MCGA. One parameter of interest to us was the amount of filtering done by the MCGA using dual–mode multicast. The degree of filtering depends upon several properties of a scenario including the number of entities on different sites in contact or involved in an engagement[3], the number of sites these entities are distributed across, and the number of separate engagements in the scenario.

We first considered scenarios of 500 entities. Each scenario was constructed by associating groups of entities with terrain grid regions. The PDUs rates for the various parts of the system given this "snapshot" were then computed. Figure 5 shows the average filtering over all sites done by the MCGAs associated with both sites when various numbers of entities were in contact. These filter percentages are shown as ranges giving the minimum and maximum filter rates that we obtained for scenarios of given sizes.

As can be seen in the figure, the filter percent varies depending upon the number of entities in contact. In scenarios with a smaller percentage in contact (30%), the filter rates at the MCGAs are higher because the entity state PDUs of the entities not involved in the engagement are not needed by any other site in the exercise. When more entities are involved in the engagement, the number of PDUs that must be filtered drops, as would be expected.

Figure 6 gives similar information about scenarios of 5000 entities. We give the average filter rates across all of the MCGAs in the exercise along two lines: the % of the entities in the scenario involved in the single engage-

[3] We consider an engagement as a group of entities in sensor range of each other.

Number of Entities	500	5000	50,000
Activity rate	100%	50%	20%
Overall PDU rates (PDU/sec)	100	5500	28000
Total Sites	2	10	50
Distribution of Entities across sites	1 x 100 1 x 400	2 x 2000 1 x 500 3 x 100 4 x 50	

Figure 4: General information about the characteristics of the three scenarios sizes.

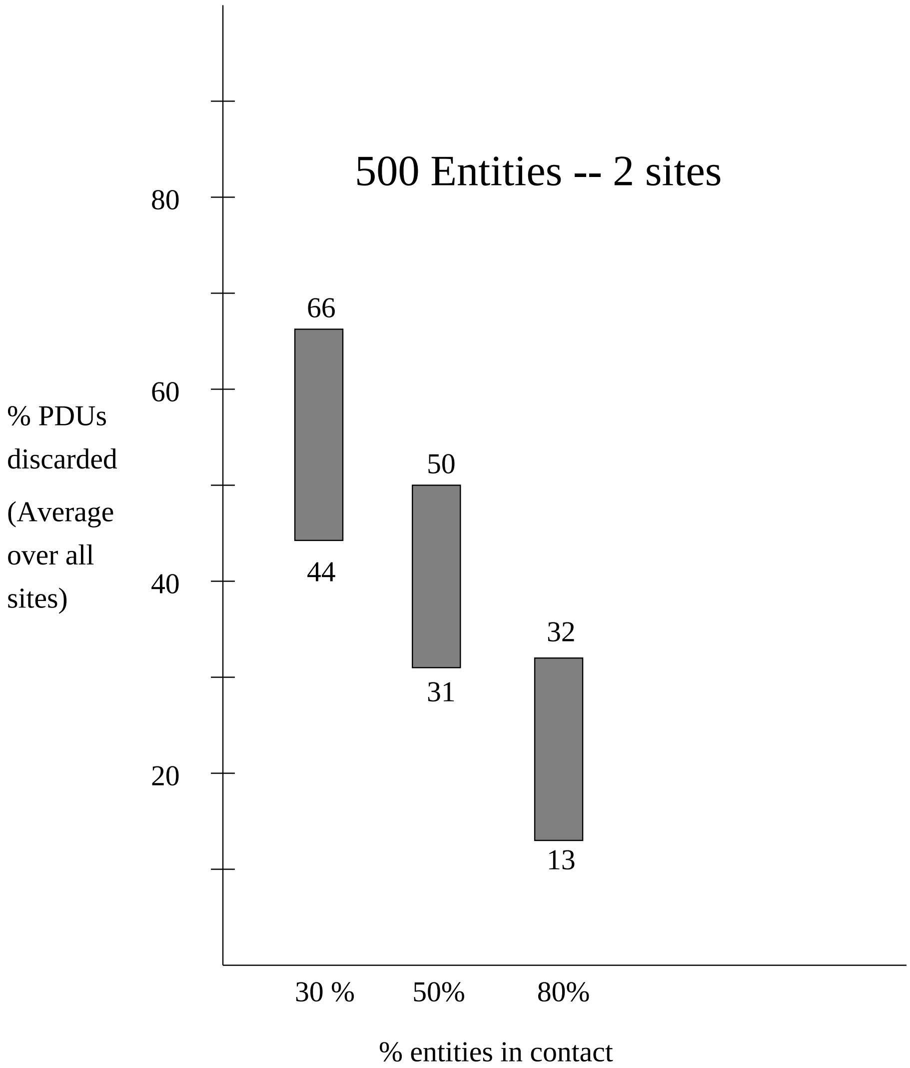

Figure 5: The % of PDUs filtered by the AG in relation to the number of entities in contact.

ment and the number of sites the involved entities were distributed across. Just as with the 500 entity scenarios, the number of entities involved in an engagement partially determines the number of PDUs the MCGA is responsible for filtering. However, this rate varies considerably depending upon the number of sites involved in the engagement. Fewer sites implies a higher average filter rate because many sites discard most of the PDUs.

5 Conclusion

This paper considers the problems of designing a scalable communications architecture for DIS. We describe multicast based approaches that limit the message traffic in different parts of the system. The primary method of interest at the moment is dual–mode multicast, an approach that uses multicast on an exercise level at the WAN level of the architecture and uses geographic-based multicast groups at the LAN level of the architecture. Our initial analysis indicates that the computational requirements of this approach may be satisfied by existing hardware and software.

Our scenarios make broad assumptions about the characteristics of exercises in terms of entity activity rates and distributions across a terrain. Better data, such as that provided by future large-scale exercises, will provide more accurate estimations of the computation requirements. There are also many open questions when considering any multicast-based approach. First we have not factored into our considerations the fact that, if the WAN is purchased as a service, the cost of delivering WAN packets to various sites must be considered in addition to the backbone loading that forms the basis of our analysis. A second open issue is the cost associated with entities moving between grid regions.

References

[CaHo94] Calvin, J. and D. Hook, AGENTS: An Architectural Construct to Support Distributed Simulation, *11th DIS Workshop on Standards for the Interoperability of Distributed Simulations*, September 1994, pp. 357-362.

[DIS94] Dis Steering Committee, The DIS Vision: A Map to the Future of Distributed Simulation, May 1994.

[HCNF94] Hook, D., J. Calvin, M. Newton, and D. Fusco, An Approach to DIS Scalability, *11th DIS Workshop on Standards for the Interoperability of Distributed Simulations*, September 1994, pp. 347-355.

[HoRC] Hook, D., S. Rak, J. Calvin, Approaches to Relevance Filtering *11th DIS Workshop on Standards for the Interoperability of Distributed Simulations*, September 1994, pp. 367-369.

[MZPB94] Macedonia, M., M. Zyda, D. Pratt and P. Barham, Exploiting Reality with Multicast Groups: A Network Architecture for Large Scale Virtual Environments, *11th DIS Workshop on Standards for the Interoperability of Distributed Simulations*, September 1994, pp. 503-510.

[Pull94] Pullen, J. M., Networking for Distributed Virtual Simulation. *Proceedings of INET '94*, paper 243.

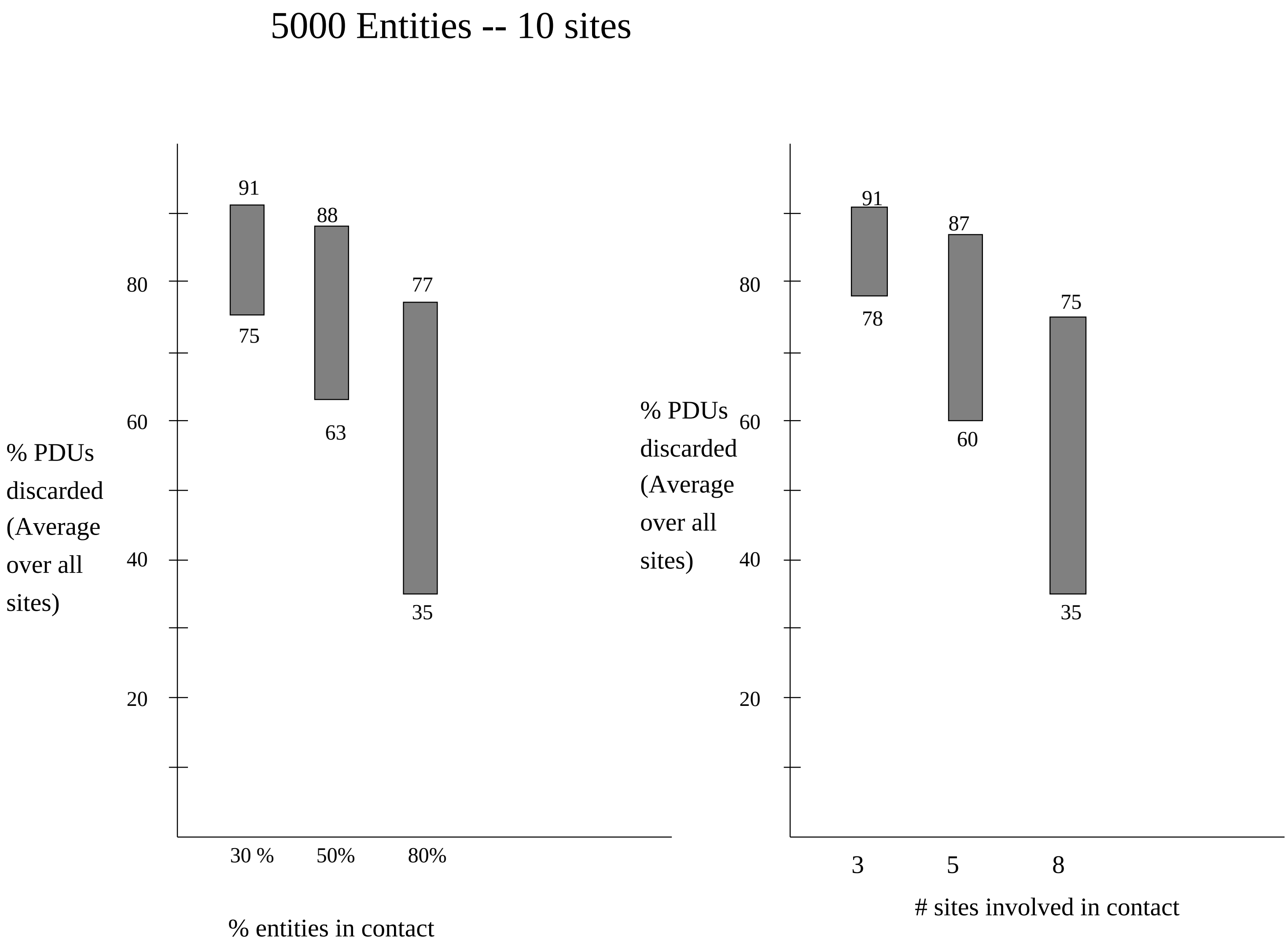

Figure 6: The % of PDUs filtered by the AG.