Improving Scalability of Large Federations that use UDP Multicast

Mr. Donald A. Washburn, Dr. Paul Wiegand,
Mr. Sergio Tafur, & Dr. Brian Goldiez
UCF, Institute for Simulation and Training
Orlando, Florida
dwashbur, wiegand, stafur, bgoldiez @ist.ucf.edu

Mr. Jaime E. Alfaro
Joint Forces Command, Orlando, FL
Mr. Mark S. Williams
Joint Forces Command, Suffolk, VA
jaime.alfaro, mark.williams @att.jfcom.mil

ABSTRACT

The University of Central Florida (UCF) Institute for Simulation & Training (IST) led a team investigating the impact of User Datagram Protocol (UDP) multicast for conveying information between United States Joint Forces Command (USJFCOM) Joint Live, Virtual, and Constructive (JLVC) simulation federates. This paper represents a summary of recommendations for reducing or eliminating the dependence of JLVC federates on UDP multicast. This work was funded by USJFCOM under N00189-08-D-Z020 Delivery Order 0009. Other large federations have similar configurations, so we believe our recommendations will be of general use. We decomposed the problem space into three areas: Run Time Infrastructure (RTI), simulations, and network. Our points stem from our own investigations, discussion with USJFCOM, as well as feedback from a workshop we hosted on UDP multicast issues affecting JLVC operations, attended by 38 JLVC stakeholders.

Since its inception, the number, size and complexity of these exercises have increased, lending to the increased use of UDP multicast, and causing network traffic and configuration difficulties. These difficulties will only continue to grow. Further, true resolution of these difficulties requires addressing fundamental scaling troubles. Our recommendations are divided into two scopes.

First, we provide some simple cost effective solutions to reduce short-term difficulties. These include network configuration changes to use sparse mode multicast and thorough RTI Initialization Data (RID) file optimization. Secondly, we provide longer-termed avenues that address scaling challenges more foundationally. We recommend alleviating complexities of traffic management by using software routing technologies within current RTI. More critically, we believe that long-term scalability of the JLVC Federation hinges on improving the effectiveness of Data Distribution Management (DDM) by incorporating more sophisticated mechanisms for declaration management—namely, incorporating entity dynamics. Our view is that sustainability and scalability of JLVC exercises are not addressed by the short-term options.

ABOUT THE AUTHORS

Mr. Donald A. Washburn is a Senior Research Scientist at the Institute for Simulation and Training (IST), University of Central Florida (UCF). During his 25 year career, Don has held Systems Engineering, Project Leadership, Software Engineering, and Principal Investigator positions developing Modeling and Simulation (M&S), military, and commercial products. Don received his B.S.E. in 1982 and his M.S.E. in 1984 both in Industrial Engineering from UCF.

Dr. R. Paul Wiegand is on the research faculty at the Institute for Simulation & Training and holds a joint appointment in the Department of Electrical Engineering & Computer Science at the University of Central Florida. Dr. Wiegand has conducted research at the Navy Center for Applied Research in Artificial Intelligence at the Naval Research Laboratory, the Evolutionary Computation group at George Mason University, and the Algorithm Efficiency and Complexity group at Dortmund Universität. Currently, he directs the Natural Computation & Coadaptive Systems lab at UCF.

Mr. Sergio Tafur is the systems administrator for the STOKES HPCC at the Institute for Simulation & Training and a PhD Candidate in the Department of Physics at the University of Central Florida. Mr. Tafur has previously served

as systems and network administrator to the Nanoscience Technology Center and Department of Physics. He currently conducts research and develops theoretical and parallel computational models describing the nonlinear optical response of matter and quantum-electrodynamic response of nanoscale quantum dots from first principles. Tafur received his B. Sc. in 2005 and M. Sc. in 2007 both in Physics from UCF.

Dr. Brian Goldiez is the Deputy Director at the Institute for Simulation and Training. During his 35 years in M&S Dr. Goldiez has conducted and lead research in quantifying and improving performance of various types of simulation. He has also served on various conference and journal committees and editorial boards. Goldiez received his Ph.D. in Modeling and Simulation from UCF.

Mr. Jaime E. Alfaro is Chief of USJFCOM-J7 Forward in Orlando. Mr. Alfaro is the previous Program Director of the Joint Development Integration Facility in Orlando. Other assignments included the Warrior Preparation Center in Germany, Headquarters US Army South in Panama, and Modern Aids to Planning Program in the US Southern Command-J5.

Mr. Mark S. Williams is the Deputy Chief of the TSD-JATTL Branch Solutions Division Training Development Group of US Joint Forces Command.

Improving Scalability of Large Federations that use UDP Multicast

Mr. Donald A. Washburn, Dr. Paul Wiegand,
Mr. Sergio Tafur, & Dr. Brian Goldiez
Institute for Simulation and Training at UCF
Orlando, Florida
dwashbur, wiegand, stafur, bgoldiez @ist.ucf.edu

Mr. Jaime E. Alfaro
Joint Forces Command, Orlando, FL
Mr. Mark S. Williams
Joint Forces Command, Suffolk, VA
jaime.alfaro, mark.williams @att.jfcom.mil

INTRODUCTION

The University of Central Florida (UCF) Institute for Simulation & Training (IST) investigated the impact of current use of User Datagram Protocol (UDP) multicast for conveying information between Joint Live, Virtual, and Constructive (JLVC) simulation federates. This paper presents a summary of our recommendation on how to reduce or eliminate the use of UDP multicast for conveying information between JLVC simulations federates. This research was performed for the United States Joint Forces Command (USJFCOM) JLVC exercises using the Joint Training and Experimentation Network (JTEN).

The JLVC Federation is a large-scale distributed simulation architecture constructed to actualize a variety of different joint training exercises hosted by USJFCOM. It includes live, virtual, and constructive simulation applications, though it focuses primarily on constructive simulations. Additionally, though the JLVC Federation primarily makes use of the High Level Architecture (HLA) for distributed simulation to handle communication and integration between different simulation applications, it also includes programs that make use of the Distributed Interactive Simulation (DIS) standard, as well as a few that use the test and Training Enabling Architecture (TENA). Typical JLVC exercises involve the execution of these distributed simulation programs on different subnets across the network, integrating the simulation of hundreds of thousands of entities operating within dozens of simulation applications at near real-time speeds. Since its inception, the number, size and complexity of these JLVC exercises have increased and will continue to do so. This growth has led to network traffic and configuration difficulties and, without some action, these difficulties will only grow. Other large federations have similar configurations, so we believe our recommendations will be of general use.

Chief among the immediate traffic and configurational challenges faced by USJFCOM arise from an increased use of multicast groups. IST was specifically tasked with investigating how to reduce JLVC Federation dependency on UDP multicast, and in particular to investigate how several newer technologies might help alleviate such problems. For example, there were questions as to whether the combination of Source Specific Multicast (SSM) and Internet Group Management Protocol Version 3 (IGMPv3) might provide a pathway for reducing multicast dependencies. Additionally, it was envisioned that emerging approaches, including the use of Internet Protocol version 6 (IPv6), may provide improved performance. We found that, while there were several advantages to these new technologies, they will not resolve the fundamental problems within the JLVC Federation, either in the short or the long term. Since in our view the core scaling problems require longer-termed solutions, while some of the configuration and traffic issues faced by USJFCOM are immediate, we provide several recommendations, addressing federation scaling issues in both the short and long term

The details of this research including the survey results were documented in the Comprehensive Roadmap Document (Institute for Simulation and Training at UCF, 2010) delivered to USJFCOM. In addition, this document included detailed steps to implement each of our four recommendations.

ANALYSIS

Our effort decomposed the problem space into three areas: the Run Time Infrastructure (RTI), simulations, and network. We performed an RTI assessment that included 57 different RTI criteria. We received responses from 13 suppliers who described 18 RTI versions. The simulation assessment included 14 criteria regarding more than 40 HLA, DIS, and TENA based live, virtual, and constructive applications. We obtained survey results for 20 simulation applications used within the JLVC Federation currently or potentially in the near future. Network information was obtained from the literature and meetings with USJFCOM staff. We obtained additional feedback by

holding a workshop on UDP multicast issues affecting JLVC operations on August 10-11, 2010 in Orlando, FL. The workshop, entitled, "Use and Development of User Data Protocol Multicast in the Evolving Joint Training Live, Virtual, and Constructive Environment", was hosted by IST and attended by 38 JLVC stakeholders, including experts in networking, HLA runtime infrastructures, and modeling and simulation applications. Participants included military, academic, and industry personnel with experience in operations for Army, Navy, and Air Force services.

Below we briefly discuss some of our findings in the three components of our analysis. Since in our view, the issues we explored are particularly critical where the RTI and simulation applications meet (interest management), we also include a separate, more detailed discussion of this topic.

Network Considerations

First, let us briefly describe several relevant methods of communicating data within a network during JLVC Federation operations. In the unicast method of communication, data is delivered uniquely between two common nodes or destinations. In the multicast mode of communication data, delivery occurs within a preselected group of destinations. An extension of the multicast method of communication, Source Specific Multicast (SSM), allows data delivery from a source to a sequence of destinations requesting interest in traffic generated by that host. The broadcast method of communication provides the delivery of data to all destinations that are part of a network[1]. Finally, anycast communication will yield the delivery of traffic to topologically "best" destinations. A topologically "best" Rendezvous Point (RP), or node in the network is one common to many sources and destinations.

Currently two popular modes of operation have been defined for use with multicast traffic: "Sparse Mode" and "Dense Mode". Though it is unclear how current technology performs, in the past USJFCOM had difficulties using multicast Sparse Mode (SM) with occasional losses of short duration. Consequently, JLVC operations currently make use of dense mode (DM) multicast. One of the major challenges presented by UDP multicast traffic as used by JLVC operations is a large volume of replicated traffic. Additionally, DM

---

[1] Technically, IPv6 has no "broadcast"; however, it has the ability to send all-local-host and all-host multicast messages, which are forms of broadcast.

periodically floods and prunes (every 3 minutes), which generates considerable traffic.

In addition to explicit hardware-level communication mechanisms, we investigated the use of Collective Communication (CC) as an alternative to current common communication schemes, which will be discussed later. Collective communication should not be confused with software routing. While software routing is essentially a "software version of hardware routing," collective communication involves data sharing between more than two tasks, which are often specified as being members of a common group, or collective. This implies that collective communication is significantly more diverse and robust than multicast communication. In fact, collective communication schemes are able to take advantage of multicast communication. A common implementation of collective communication actively being developed, maintained and deployed by the high-performance computing community (where networking applications consistently demand low latency and high scalability) is the Message Passing Interface (MPI). MPI is a language-independent communications protocol that supports both point-to-point and collective communication among nodes (here nodes refers to its most abstract definition as points on any general mathematical network).

We also examined the use of various network protocols. The network protocols and hardware implementations evaluated can all be used independently or in conjunction to facilitate the most effective/efficient means of communication and delivery of meaningful information between hosts and clients. When directly comparing Internet Protocol version 4 (IPv4) to IPv6 it is clear that there exist no real disadvantages with some advantages associated with transitioning to IPv6. Though IPv6 is not backward compatible with IPv4, there exist transition strategies and mechanisms that allow hosts in using either protocol to communicate with one another. This is also true of hardware implementations where an additional layer of communication and routing allows traffic from one to be routed to the other. In IPv6, the space of addresses starting with FFxx::/8 are intended for multicast communication [RFC 3306 (Internet Engineering Task Force, 2010)]. This subset includes source specific multicast addresses defined within the range of FF3x::/96. An added feature of IPv6 is that the addresses of routers that act as rendezvous points (RP) can be embedded within the IPv6 address itself. This scheme is formally termed "Embedded rendezvous points" or "Embedded RPs" [RFC 3956 (Internet

Engineering Task Force, 2010)]. These Embedded RPs can be used, in the following way:

FF7x:y40:2001:DB8:BEEF:FEED:: Multicast addr. w/Embedded RP
2001:DB8:BEEF:FEED::y   Router Address.

In contrast to IPv6, which relies on Embedded RPs for routing multicast traffic, IPv4 requires additional standardized protocols, such as Multicast Source Discovery Protocol (MSDP) [RFC 3618 (Internet Engineering Task Force, 2010)] for defining RPs. When considering transitioning from IPv4 to IPv6 multicast traffic, Embedded RPs can therefore take on the role of protocols such as MSDP in defining rendezvous points for routing purposes.

When discussing the communication paradigm that should be central to addressing the scalability of "network" topologies and their operators in computer networks, it is still important to consider collective communication (CC). The CC paradigm should be considered for network and topological definitions of network subsets and operators because it is part of the standard required to define optimization of performance benchmarks and scalability benchmarks.

The JLVC simulation federations can be viewed as such a group of "network" topologies and subsets with their corresponding operators. We therefore recommend working in the CC paradigm when addressing their scalability. We recommend working in the CC paradigm not only because it abstracts away the complexities of protocol specific communication, but because it inherently defines compatibility between protocol versions, implementations, and the optimization of performance and scalability benchmarks. We discuss some of these ideas below in the *Interest Management Considerations* section.

RTI Considerations

There are three different HLA/RTI specifications that include requirements for the RTI. The first was HLA 1.3, which was released in 1995. The second was IEEE 1516-2000, which was released in 2000. IEEE 1516-2010, which is also called HLA Evolved, is the third and newest specification. Below are some key HLA Evolved features as described in 08F-SIW-064 (Möller, Morse, Lightner, Little, & Lutz, 2008) that would benefit JLVC federations. Although there are many advantages of these newer specifications, we were unable to determine how they would substantially reduce the JLVC federation dependence on UDP multicast.

- **Modular Federation Object Models (FOMs) and Simulation Object Models (SOMs)** make it possible to maintain HLA object models in a more flexible and scalable way, to separate local extensions from standardized models, to support partial federation agreements and even to gradually extend the object model of long running federations.
- **Web Services support** through the new Web Services Description Language Application Program Interface (WSDL API) gives users of Web Services communication frameworks access to the full HLA functionality over wide area networks while providing features like authentication and encryption. It also provides support for a large number of programming languages.
- **Fault tolerance support** gives federations a well-defined semantics for handling unreliable or crashing federates and network links that go down.
- **Smart update rate reduction** makes it possible to subscribe to the same information with higher or lower update rates. This gives the ability to focus on certain entities or to reduce network traffic on wide area networks.
- **Dynamic link compatibility** makes it easier to switch between different RTI implementations without modifying, recompiling or re-linking federate code.

From our evaluation of the many RTI versions available, we determined they all have strengths and weakness. There is not an RTI version that is clearly superior. In addition, there are some roadblocks to migrating to a different RTI provider. Most RTIs are not backward compatible or interoperable between vendors. For example, changing standards (i.e. 1.3, 1516-2000/2010) can cause all the Application Program Interfaces (APIs) to change. These changes would impact all the simulations in the federation. Even if Dynamic Link Compatible (DLC) API was followed by a federate, many of the new RTI's improvements could require a significant redesign and resulting cost of change for each federate. This is especially true when changing between different HLA specifications. Several of the RTI products have tools that support converting a federate when changing HLA specifications. However, when these interface tools are utilized the federate will not utilize the new features, for example of HLA Evolved. Therefore, each federate would need to be updated in order to take advantage of the new specification features. This could require significant effort to re-integrate and re-test the federation under the new RTI version. So the benefits of these new features need to be considered in comparison to impacts of these changes.

Simulation & Application Level Considerations

Outside of interest management, there is very little individual simulation applications can do to address the scaling problems with the Federation as a whole. To some extent, the Department of Defense's interest in large-scale distributed simulation technology is not that different from modern popular, Internet-oriented games. Massively Multiplayer Online Games (MMOGs) also involve a number of interacting entities operated by software located across a large network of networks. Indeed, MMOGs must also find ways to control the amount of data that is exchanged by managing the interests of these software components.

While some MMOG distribution and load balancing questions are similar to large federation needs for large-scale federation operations, the actual routing and delivery of content faces quite different obstacles. Instead, the MMOG community makes use of communications mechanisms that are commonly referred to as application level multicasting[2]. Application level multicasting methods are software-based, self-configuring solutions that attempt to provide efficient delivery for multipoint communication across heterogeneous networks (Yeo, Lee, & Er, 2004).

Such methods are particularly useful for implementing distributed systems that make use of collective communications (Chen, Huang, & Lei, 2005), such as those employed by, and optimized for, very well-known systems for parallel distributed processing, such as MPI (Thakur, Rabenseifner, & Gropp, 2005). This field is fertile ground for future research for distributed simulation systems since there's been a great deal of work investigating methods for optimizing them in a variety of ways (Chan, Heimlich, Purkayastha, & Van de Geijn, 2004)—particularly for situations in which a large volume of small messages must be sent between a large number of group members (Bruck, Ho, Kipnis, & Weathersby, 1994).

Given the flexibility and configurability of collective communications mechanisms, there may well be situations in which hardware-based Internet Protocol (IP) multicast solutions cannot sufficiently capitalize on efficiency advantages over application level

---

[2] We note that strictly speaking, these are *not* multicast from a network engineer's point of view; however, the term "*application level multicasting*" is accepted and widely used in some communities. We use it here to remain consistent with the literature we are discussing.

multicasting. That is, though there are obvious latency questions with software-based routing solutions, hardware solutions are not universally superior in terms of latency for large-scale distributed, collective operations. More importantly, software-based solutions leave the burden of filtering and focusing information within a collective system in the software implementing such solutions, as opposed to sharing that configuration burden with hardware that is typically agnostic to the content of the data being routed.

Interest Management Considerations

In order to accommodate such large-scale, rapid communication, it is necessary for the federation to "manage the interests" of individual simulation applications with respect to shared data. That is, since it is impractical and unnecessary for all simulations to receive all information about all entities in every time step, there must be ways to filter and focus data flow. Amongst HLA applications, interest management is primarily handled via Data Distribution Management (DDM) services offered by the RTI of the HLA, which in turn makes use of hardware level multicast (among other things) to help deal with traffic on the network. Interest management of federation objects (entities with attributes and simulation events) is handled by publish and subscribe mechanisms, which typically involve filtering interest by determining overlapping publish and subscribe regions that exist within some set of routing spaces.

The JLVC Federation makes use of one, large region (the "hyperspace") providing federate-level flexibility via the use of application spaces, or app spaces (Technical Development and Innovation Branch (TDIB), 13 January 2010). Application spaces are federate-specific subspaces of the routing space in conjunction with some set of object classes such that all related classes are bound to the routing space. The resulting routing space is the minimal subset of all unique attribute dimensions in all app spaces, allowing federates to specialize the application space for their purpose while still providing an embedding into a shared, though more complex, routine space. The JLVC Federation then subdivides app spaces for particular object classes.

Filtering of data occurs primarily at the IP-level by assigning these partitions to multicast addresses rather than using exact match, region-based region filtering via LRCs. For the JLVC, subscription regions correspond to these application subspaces, and update regions are treated as single points. That is, the RTI ensures that simulation programs that produce shared

data send that information out on multicast group addresses associated with the app space regions in which the program has update activity, and the simulation programs that need to consume data in some application region subscribe to the associated multicast address. In this way, there can be (in the worst case) a separate multicast group for every application space partition, though not all exercises need to use all of these. This means there is an intrinsic tradeoff between resolution of the regions and the quantity of multicast groups.

We focused our research on investigating how to mitigate resource challenges imposed by the JLVC's heavy dependency on multicast groups; however, there is a broader issue concerning how to handle interest management within the JLVC Federation more efficiently because, as noted above, there is a potentially close relationship between DDM and multicast as it is currently implemented. While the methods by which the federation does or does not use hardware level multicast certainly address aspects of the interest management question at the level of how federates are connected and how information between them is routed, such a focus omits earlier components in the interest management pipeline—including, and most notably, how federates communicate their interests in the first place. We believe the core scaling challenges within the JLVC Federation stem from the mechanisms by which it chooses to declare federate interests rather than whether or how multicast is used, as explained below.

Region Based DDM
The DDM is a flexible publish/subscribe paradigm for managing the data interests of federates within simulation federations. Federates typically express interest in data consumption by subscribing to regions of a larger space (the routing space) for given entity types, and they typically describe their ability to produce data for the federation by publishing about types (again, typically within regions). There are many methods used to accomplish effective interest management and data filtering within this paradigm. Indeed, conceptually one can separate its basic responsibilities as follows (Petty & Morse, July 2004):

- **Declaration**: The process whereby federates modify their update and subscription regions as their data interests change and then inform other federates of these changes;
- **Matching**: The process by which the RTI determines which update and subscription regions from which federates overlap;

- **Connectivity**: The process that uses the match results to establish network connectivity (e.g., joins to multicast groups, etc.);
- **Routing**: The process whereby the RTI transports data from one federate to another using the established connectivity.

Use of hardware-based multicast groups refers simply to the last two pieces. When connectivity and routing are implemented using entirely software-based methods, DDM itself becomes a type of application level multicasting, such as those described in the previous subsection regarding the communication methods used by modern MMOGs. But there are still a variety of ways to handle declaration and matching, each of which may or may not make use of multicast groups. To some extent, the scaling problems regarding the number of multicast groups is an artifact of scaling problems created by implicit choices in these first two pieces, declaration services in particular.

Our view is that addressing the core problems existing within the JLVC Federation will require a long-term change that focuses more on declaration services. We believe it is possible to address these scalability problems by focusing future efforts on investigating and developing DDM-oriented services to help federates make better use of the DDM.

Determining Subscription and Update Regions: The True Challenge
While there's some focus in the literature about how to improve performance using sender-side filtering (e.g., (Van Hook, Rak, & Calvin, 1996)), our position is that the true challenge is in establishing appropriate update and subscription regions in the first place. Indeed, one advantage to grid-based approaches (and, by extension, the app spaces approach that the JLVC Federation uses) is that it simplifies the question of how to effectively design these regions by reducing a host of complex engineering decisions to one central issue: cell resolution. Unfortunately, this reduction also removes the possibility of some more sophisticated and more efficient system level solutions.

In region-based approaches, when the update and subscription regions change very quickly, declaration services itself can become a network bottleneck. When one thinks of the routing space as simply a simulated physical geography and encourages federates to establish simple physical bounding regions for entities without consideration to movement and behavior, this problem crops up naturally in simulations with fast-paced, highly dynamic entities. But the dynamism of simulation objects can be built into the DDM by taking

care to consider how federates treat the routing space and how they manage changes to update and subscription regions (Morse, Bic, Dillencourt, & Tsai, 1999). Moreover, there are known methods for incorporating dynamic coordinates for entities that allow for physically correct filtering[3] (Morse & Steinman, 1997). These methods attempt to deal with such situations by bounding how much a routing space coordinate value can change over some clear unit of time and minimizing the sampling rate of both update and subscription regions by taking into account velocity and range-of-motion aspects of entities (among other things). In that sense, such approaches require that federates are implicitly involved in data distribution management in a peripheral way, in the sense that they must calculate regions and generate region updates in a way that permits efficient interest management within the federation. It also requires a somewhat more sophisticated use of routing space than is traditionally done in the JLVC (for example), though federation agreement regarding routing space is always an issue that affects federates, regardless of how DDM is approached.

Unfortunately, many simulation applications that serve as federates within the JLVC Federation are already well-established software projects, and some are no longer under active development. Moreover, most of these programs operate in a number of joint simulation environments, including HLA federations other than the JLVC. This means that agreements about routing spaces requires simulations take careful consideration in being flexible about how to deal with update and subscription regions within a variety of routing spaces. For some, integration of the program into a large-scale HLA federation was not a primary design goal at the outset of development. While the HLA and DDM services are quite flexible, because of these issues it is challenging to leverage this to attain highly efficient interest management.

DDM Service-Oriented Components for Federates
There is an essential dilemma within the challenges of designing effective interest management schemes within the DDM: We want the RTI to be a general tool for which individual simulations are as unaware of as possible, but we need simulations to incorporate at least some sim-specific information into declaration services, at least in terms of how regions are managed. This

dilemma affects federation operation in terms of computations performed by the simulation for more sophisticated declaration management, as well as interoperability issues generated by incorporating more sim-specific information into the DDM. To wit, solutions to this problem need to be, above all, flexible and modular.

Participants in our Simulation group (the others being Network and RTI) of the mid-term workshop for this project introduced an idea for a long-term, modular solution. They suggested that there be a way to provide services to federate simulations that help them flexibly incorporate the use of things like dynamic coordinates into their construction and maintenance of subscription and update regions. This idea would be akin to the current proposed use of encoder and decoder helpers within HLA Evolved 2010. The concept behind encoder and decoder functions is to give federate developers modular functions that they can use inside their simulations to encode and decode shared data items in order to alleviate the burden of data translation agreements. Of course, tying the simulation to an HLA and by extension to a particular RTI has a cost and utility impact on the simulation.

Like encoder and decoder helpers, these components would not be a generalized operational component of the RTI, but would rather be a collection of interoperability tools and services (perhaps provided with some RTI distribution) to be operationalized within the federate itself. These would make it possible for simulation federates to capitalize on service-oriented approaches, such as those used with technologies such as Common Object Request Broker Architecture (CORBA) and Distributed Component Object Model (DCOM)[4]. Among the things that such tools might provide for federates, are means of obtaining services to convert velocity, position, and time in terms of a provided routing space, in addition to a means of identifying sufficiently infrequent rates for region modification requests, etc. If the tools developed are useful and appropriately modular, then as the community constructs new tools, their availability (and consequently, their utility) would increase.

---

[3] By "physically correct filtering", we mean that distributed objects are discovered in time to be responded to correctly by a federate at the correct time in the simulation.

[4] CORBA and DCOM are distributed, shared object technologies meant to help developers create general services to be used by a variety of applications. Because interoperability is crucial with such technologies, developers make use of a variety of ways of packaging and un-packaging data in platform independent ways.

Future research into this idea may capitalize on lessons-learned from the as-yet future application of HLA Evolved encoder and decoder helpers. Additionally, it may be possible to base design and implementation ideas on existing service-oriented approaches to distributed interest management (Steinman, Valinski, Lammers, & Steinman, 2010) (Pan, Turner, Cai, & Li, 2009).

RECOMMENDED SOLUTIONS

We created four recommended solutions, ranging from short-term fixes that emphasize mitigation of the immediate troubles over resolving underlying scaling problems to long-term fixes that seek to address the core issues that we believe are leading to the current scaling problems. Our recommend solutions are: use multicast in sparse mode, perform thorough optimization of the RID, make use of software routing facilities within the RTI, and investigate a thorough restructuring of interest management within the JLVC Federation. We also emphasize the need to create an independent test lab as a means of facilitating all solutions. This test lab is useful for performance analysis and testing of all proposed solutions.

We have investigated the possibility of adopting IPv6 as an alternative to IPv4 and it is evident that there are some advantages associated with transitioning from IPv4 to IPv6 with no real disadvantages. However, there is currently no real improvement on the performance of data delivery, though there may be a theoretical benefit to reduced latency times associated with the, as yet untested, use of jumbo-grams. Any possible return on investment for the Department of Defense (DoD) will be likely associated with the increased address space and heightened level of security provided by address obscurity and the protocol requirement of Internet Protocol Security (IPSec). We do not recommend the use of SSM or IGMPv3 with large numbers of multicast groups and suggest an alternative approach, SM, to resolving the challenges associated with large numbers of multicast groups being faced by current and future implementations of a large federation.

We have additionally determined from our investigation and confirmed our conclusions from the workshop that applications of IPv6 vs. IPv4 to the resolution of the DM associated challenges being faced by the current implementation of the JTEN will not resolve these challenges. Though IPv6 provides the ability of including embedded rendezvous points into multicast addresses, this still does not resolve the requirement for a large volume of multicast groups associated with large federations in the JTEN. As a result we encourage the transition from IPv4 to IPv6 due to DoD mandate and the associated advantages it presents in terms of security and added functionality.

Solution #1 – Sparse Mode

The USJFCOM Mid-Term workshop participants agreed with IST recommendation that a transition to Sparse Mode (SM) multicast communication is the quickest and most cost effective way of reducing the dependence on intermediary floods and prunes associated with Dense Mode (DM). The transition to SM will require performance analysis and testing of SM on the JLVC Federation testbed.

Deployment of SM must be carried out all at once throughout the whole network. This is imperative since all federates communication via multicast and these will not be able to communicate with one another if the network is divided in DM and SM groups. The router configuration changes that would be necessary require that; for IPv4, routers use MSDP to coordinate multiple RPs and for IPv6, routers use embedded RPs to coordinate multiple RPs. Disadvantages associated with transitioning to SM include the fact that it cannot be approached through a phased plan of change over, does not decrease the number of multicast groups, does not reduce the configuration challenge, continues to carry the risk of routers confusing or recognizing packet amplification effects of multicast traffic as a Denial of Service DoS attack, and may lose short duration events. Many of these risks exist primarily when requiring that the performance of the network scale well with the number of multicast groups. For the current scale and planned implementation of this solution presents a good investment with low risk in the short term.

Solution #2 – Optimized RID

The USJFCOM Mid-Term workshop participants recommend that the JLVC Federation RTI Initialization Data (RID) should be optimized for usage of multicast groups. The RID file defines the initialization data for the RTI which includes the number of subspace partitions the RTI will create. The number of multicast groups is defined in the RID. The RID we examined had several groups of application spaces that are no longer used, so these definitions could be removed. In addition, there were many remaining applications spaces that were undefined. Each of these undefined application spaces will be assigned to one multicast

group. Once these two groups of application spaces are removed, then the maximum number of subspaces can be reduced which will reduce the number of multicast groups created.

We also recommend an investigation into the feasibility of developing a tool(s) that would optimize the RID in terms of the number of multicast groups and traffic. This would reduce the time spent determining and updating the number of partition spaces. In addition, each exercise would have an optimized RID in place of the general one currently used for all exercises.

Solution #3 – RTI

RTIs that allow each communication channel to be assigned as being Transmission Control Protocol (TCP), UDP, or Message Passing Interface (MPI) provide flexibility that is beneficial to large federations such as the JLVC. For long distance Wide Area Network (WAN) communications, some RTIs allow UDP that eliminates the acknowledgement messages. Time critical WAN communications could utilize TCP. A hybrid approach allows both connectionless and connection-oriented communications within the WAN or Local Area Network (LAN). We recommend migrating to a hybrid approach since channels can be configured for the best communications method.

Some of RTIs support alternate communication modes like UDP broadcast, semi-reliable, reliable multicast, state consistent, minimum rate, and Simulation-Aware Routing (SAR). These RTI alternate communication modes offer a way to reduce the JLVC dependence on UDP multicast.

Some of the alternate communication mode features offered by newer products are:

- Eliminates UDP multicast over WAN by using TCP;
- Allows flexible of UDP Multicast or TCP over LAN or WAN;
- Supports using TCP, UDP, and MPI assigned on a per communication channel;
- Quality of Service options such as rate limit, data bundling, and compression;
- Additional transport types of semi-reliable, minimum rate, state consistent besides TCP & UDP.

Solution #4 – Interest Management Restructuring

In addition to short- and medium-term solutions for mitigating resource challenges imposed by the JLVC's heavy dependency on multicast groups, we believe it is in JLVC Federation's best interest to consider more foundational improvements to interest management within the JLVC Federation—particularly how federation objects are defined and organized, as well as how federates declare and communicate their interests. Expenditures now will result in savings in the future as the number and complexity of planned exercises increases. We suggest that the community should investigate long-term changes in the FOM structure and DDM that address such issues.

Specifically, we suggest JLVC Federation investigate how to eliminate the use of hardware-level multicast in favor of dynamic, physically correct, region-based exact filtering at the RTI / LRC level. This would involve modifying DDM to make use of both update and subscription regions, as well as modifying federation agreements regarding the routing space to incorporate a dynamic, region-based approach. The idea is to reduce the overall traffic by addressing the fundamental problem of inefficient interest management.

In addition, any long-term restructuring should consider migration to HLA Evolved (IEEE 1516-2010). In this case, restructuring the FOM suggests investigation into the use of modular FOMs, a new feature of HLA evolved that should allow more flexible management and reuse of FOM components. Additionally, the JLVC Federation would benefit by the use of new efficiency features within HLA Evolved, such as smart update rates, improved fault tolerance, etc. On a related note, steps should be taken to ensure future RTI's are backward compatible and interoperable between vendors at some minimal level.

These suggestions entail a longer-term and more costly effort than the previous recommendations. Most notably, more sophisticated DDM that allows for physically correct, dynamic region-based filtering will necessitate changes to underlying federates. Still, there may be ways to reduce the cost faced by requirements on federate developers to make changes to accommodate more sophisticated declaration management. We suggest that the community direct future research toward investigating how to provide DDM services to federate simulations that help them flexibly and modularly incorporate the use of things like dynamic coordinates into their construction and maintenance of subscription and update regions. We also suggest developing user-oriented, visual tools for helping manage DDM, RID, and FOM changes

between exercises. Such tools can perform graph-based optimizations to try to reduce interactions and generate exercise-specific configuration files (such as the RID file) from more general templates. These will likely be quite useful, even if the DDM is not restructured.

In addition to costs and time associated with potential changes to federate simulation applications, it may be necessary to work through certain challenges with an RTI developer. Depending on how the above changes are approached, they may require changes to the configuration of communication within the federation. Because of efficiency and fault tolerance issues, it may be worthwhile investigating hybrid configurations using a mixture of best-effort, connectionless communication for high frequency and broadly disseminated data with connection-oriented communications for more directed data delivery. Further, we recommend taking this opportunity to begin the transition to IPv6, which will require operating the federation in a dual stack configuration for some period of time, a configuration that is not possible for some RTI implementations.

CONCLUSIONS

This paper summarizes results of IST's investigation into the scaling and configuration challenges associated with the JLVC Federation dependence on UDP multicast. Our view is that the JLVC Federation's use of large numbers of multicast groups during exercises is a symptom of a broader problem rather than the root cause. Simply eliminating UDP multicast by shifting to broadcast and receiver-filter methods would serve to reduce hardware dependencies related explicitly to multicast, but at the cost of severe latency and throughput challenges. The core issue is the scalability of the federation in general, which touches on a number of factors.

First, we discussed some simple, cost effective, short term solutions that alleviate difficulties associated with the use of UDP multicast in the near future. These include network configuration changes to permit the use of multicast in sparse mode and thorough optimization of the RID file used by federation exercises. The first will not reduce the number of multicast groups, but may alleviate some of the traffic challenges and thus reduce some of the hardware burden in the near-term. The second may make modest decreases in the number of multicast groups, but is unlikely to have significant impact on future scaling.

Second, we presented longer-termed avenues that address the scaling challenges more foundationally. In the first case, we recommend alleviating complexities of traffic management on the WAN by making use of RTI software routing technologies. More critically, we believe that the long-term scalability of the JLVC Federation hinges on its ability to improve the effectiveness of data distribution management by incorporating more sophisticated mechanisms for declaration management—namely, incorporating entity dynamics into subscription and publication region updates and taking a deeper look at how to optimize federation and DDM configurations. While the return on investment for such a longer-termed view may seem to deter such action, our view is that sustainability and scalability of JLVC exercises are not addressed by the near-term options described above.

ACKNOWLEDGEMENTS

REFERENCES

Bruck, J., Ho, C.-T., Kipnis, S., & Weathersby, D. (1994). Efficient algorithms for all-to-all communications in multi-port message-passing systems. *Proceedings of the sixth annual ACM Symposium on Parallel Algorithms and Architectures (SPAA)* (pp. 298-309). Cape May, New Jersey: ACM.

Chan, E., Heimlich, M., Purkayastha, A., & Van de Geijn, R. A. (2004). On Optimizing Collective Communication. *Proceedings of the IEEE Conference on Cluster Computing* (pp. 145-155). San Diego, CA: IEEE.

Chen, K.-T., Huang, P., & Lei, C.-L. (2005). Game traffic Analysis: An MMORPG perspective. *Proceedings of the International Workshop on Network and Operating Systems Support for Digital Audio and Video* (pp. 19–24). Skamania, Washington: ACM.

Internet Engineering Task Force. (2010). *Request for Comments (RFC)*. Retrieved November 1, 2010, from Internet Engineering Task Force (IETF): http://www.ietf.org/rfc.html.

Institute for Simulation and Training at UCF. (2010). *Comprehensive Roadmap and Final Report*. Orlando, FL: University of Central Florida.

Möller, B., Morse, K. L., Lightner, M., Little, R., & Lutz, R. (2008). HLA Evolved – A Summary of Major Technical Improvements (08F-SIW-064). *2008 Fall Simulation Interoperability Workshop*. Orlando, FL: SISO.

Morse, K. L., & Steinman, J. (1997). Data distribution management in the HLA: Multidimensional regions and physically correct filtering. *Proceedings of the Spring 2009 Simulation Interoperability Workshop* (pp. 97S-SIW-052). San Diego, CA: SISO.

Morse, K. L., Bic, L., Dillencourt, M., & Tsai, K. (1999). Multicast grouping for dynamic data distribution management. *Proceedings of the 31st Society for Computer Simulation Conference*.

Pan, K., Turner, S. J., Cai, W., & Li, Z. (2009). Implementation of Data Distribution Management Services in a Service Oriented HLA RTI. *Proceedings of the 2009 Winter Simulation Conference.* Austin, TX : WSC.

Petty, M. D., & Morse, K. L. (July 2004). The computational complexity of the high level architecture data distribution management matching and connecting processes. *Modeling Practice and Theory*, 12 (3-4), 217-237.

Steinman, J. S., Valinski, M., Lammers, C., & Steinman, W. (2010). Scalable Publish and Subscribe Data Distribution. *Proceedings from the 2010 Simulation and Interoperability Workshop* (pp. 10F-SIW-027). Orlando, FL: SISO.

Technical Development and Innovation Branch (TDIB). (13 January 2010). *Joint Live Virtual and Constructive (JLVC) Federation Integration Guide.* Suffolk, VA: United States Joint Forces Command (USJFCOM).

Thakur, R., Rabenseifner, R., & Gropp, W. (2005). Optimization of Collective Communication Operations in MPICH. *International Journal of High Performance Computing Applications*, 19 (1), 49-66.

Van Hook, D. J., Rak, S. J., & Calvin, J. O. (1996). Approaches to RTI Implementations of HLA Data Distribution Management Services. *Proceedings of the 15th DIS Workshop,* (pp. 535-544). Orlando, FL.

Yeo, C. K., Lee, B. S., & Er, M. H. (2004). A survey of application level multicast techniques. *Computer Communications, 15*, 1547-1568.