Quality of Service for Distributed Simulation Environments

Eberhard K. Kieslich
ARCIC/JAMSD
Fort Knox, KY
ekieslich@trideum.com

Diana Pineda
ARCIC/JAMSD
Orlando, FL
dpineda@trideum.com

ABSTRACT

This paper summarizes discoveries and remedies of simulation protocol data loss across the Battle Lab Collaborative Simulation Environment (BLCSE) Wide Area Network (WAN). The use case for a network supporting large-scale constructive simulation, combined with other traffic, carries special requirements outside the typical boundaries for normal systems administration of a WAN. Understanding the challenges and solutions involved is certainly not mundane. Simulation data losses in excess of 1% can impose compounded causal effects that will easily jeopardize the analytical benefit of an experiment. The possibility of data loss must be vigilantly monitored and vigorously interdicted. One root cause of packet loss is network congestion. Congestion occurs at chokepoints, which exist in nearly all network topologies where a number of hosts on a local network aim to connect to resources at remote destinations via a shared infrastructure. Another, unanticipated cause of loss is technology integration conflict, based on original design assumptions. Finally a more surprising and insidious cause is the burstiness of simulation traffic in which High Level Architecture (HLA) packet loss occurs at utilization levels far below the bandwidth threshold (i. e. without congestion). Over time, the approach of the Cyber Enterprise Support Center (CESC) and Army Capabilities Integration Center (ARCIC)/Joint Modeling and Simulation Divisions' (JAMSD) management of the BLCSE network has evolved from reaction, to monitoring, to deliberate stimulation and most recently to the intentional, governing configuration and application of the Cisco IOS Quality of Service (QoS) technology. The hard-felt experiences of the BLCSE community, as well as the powerful off-the-shelf and custom technologies used will provide a tremendous value to the greater modeling and simulation community, and should be seriously considered for other wide-area simulation environments.

ABOUT THE AUTHORS

Eberhard Kieslich is the chief architect for Trideum Corporation's Simulation Support Operations. Current work includes BLCSE support, where he performs system integration, testing, capability development, planning, and risk analysis. Mr. Kieslich has over 25 years' experience in modeling and distributed simulation. He earned a BSEE-equivalent (state-certified Engineer) diploma from Wurzburg Technical College in Germany. He is a certified modeling and simulation professional (CMSP).

Diana Pineda, CMSP is a software engineer supporting ARCIC's Battle Lab Collaborative Simulation Environment. She holds a Bachelor of Science Degree in Engineering Physics from Embry Riddle Aeronautical University. She has been involved in systems engineering of military satellites and missile system and as a software engineer for modeling and simulation for the past 15 years.

Quality of Service for Distributed Simulation Environments

Eberhard K. Kieslich
ARCIC/JAMSD
Fort Knox, KY
ekieslich@trideum.com

Diana Pineda
ARCIC/JAMSD
Orlando, FL
dpineda@trideum.com

INTRODUCTION

The Battle Lab Collaborative Simulation Environment (BLCSE) is a multi-site, multi-federate simulation environment with the purpose of data collection for Training and Doctrine development together; with technology demonstrations and evaluations that enhance the warfighter's ability on the future battlefield.

In this environment, primarily used for experimentation, the preservation and understanding of data is fundamentally critical for success. BLCSE Simulation events have dramatically increased in volume and user base in recent years, placing ever-changing load conditions (due to data) onto the LAN and WAN architectures. Packet losses of the HLA protocol have been an on-going challenge for the engineers. Non-symmetries in data collection and reporting since 2011 were some initial indications of the problem. Through a myriad of fixes, reconfiguration and upgrades to the simulation, protocol and network, the data packet losses have been significantly reduced.

Nonetheless, packet losses were not entirely eliminated, and class-based[1] Quality of Service (QoS)[2] on the network has been thoroughly examined on the bench and the production network. Several intermediate development efforts were necessary for a successful implementation. Thus far, the quest for a "zero loss" solution of the simulation protocol has produced demonstrably beneficial results in terms of measured traffic. More fine tuning may be necessary, and the incorporation of a Daily Readiness Test into the application to verify the preferential treatment of the critical HLA streams under full load by a single operator are future goals for this capability.

OVERVIEW

The BLCSE consists of 20 Army and Joint Battle Labs connected via the Defense Research and Engineering Network (DREN III) Wide Area Network (WAN) to enable distributed, large-scale war-gaming and experimentation. In support of this mission, BLCSE contains collaboration services, cutting edge web-technologies, and various communication tools. The Army Capabilities Integration Center (ARCIC) Unified Challenge 2015 experiment contained about 64,000 simulated entities, and in the upcoming event of 2017 there is an estimated requirement of 100,000 entities, as shown in Figure 1. The generated real-time simulation load by the various federates is not only based on

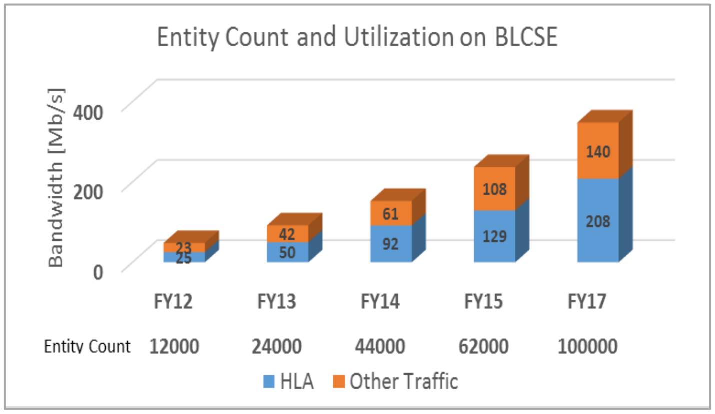


Figure 1. Entity Count on BLCSE for Given Events

[1] A **QoS class** determines the priority and bandwidth for traffic matching a QoS Policy rule.
[2] **QoS** is a set of technologies that work on a network to guarantee its ability to dependably run high-priority applications and traffic under limited network capacity. QoS technologies accomplish this by providing differentiated handling and capacity allocation to specific flows in network traffic.

entity quantities, but is also dependent on the scenario; the more activity (e.g. movement, sensing and target engagements), the greater the load. Simulation load does not equal network load, although a proportional relationship exists. This paper explores the “on-the-wire” network load.

Hub and Spoke Topology

The BLCSE features a classical Hub and Spoke architecture, with the Ft. Gordon Cyber Enterprise Support Center (CESC) as the hub as shown in Figure 2. The CESC provides the core architecture that ensures connectivity and packet distribution between all endpoints. The various 20 labs differ greatly in volume generation of HLA traffic. The federates at each site share a gigabit Ethernet Local Area Network (LAN) that sends and receives several multicast streams to/from the hub over a 50Mb/s or 100 Mb/s pipe, dependent on the requirements of the lab. The access speed at the hub is shared by all pipes and is currently limited to 700 Mb/s. The network uses encryption devices to securely transmit the encrypted real-time data over unencrypted WAN links, utilizing GRE tunnels[3].

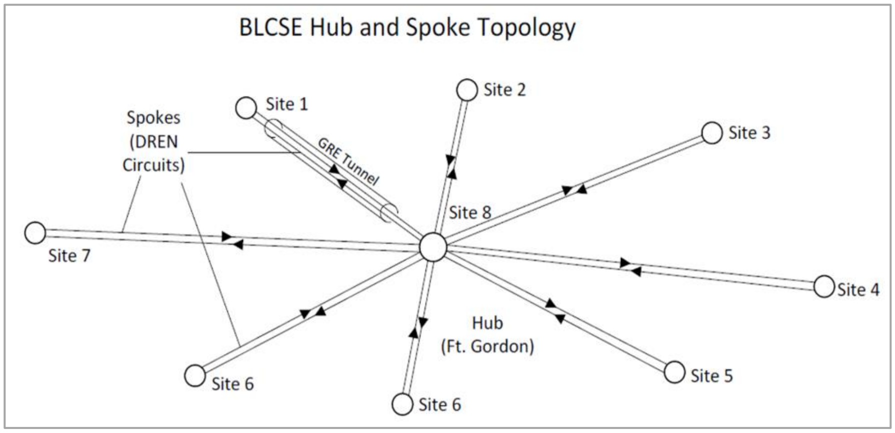


Figure 2. Hub and Spoke Topology

Interestingly, network overload and the associated packet losses can be a result of the mere cost of–bringing systems on line to observe, expand, or enhance the ongoing simulation. Due to the nature of the hub and spoke architecture, any additional site coming on-line places an additional load on the core router as well as on the WAN interface at the hub. The diagram in figure 3 shows the details of an additional load as site E comes on-line, the outbound load at the hub increases by 12 Mb/s.

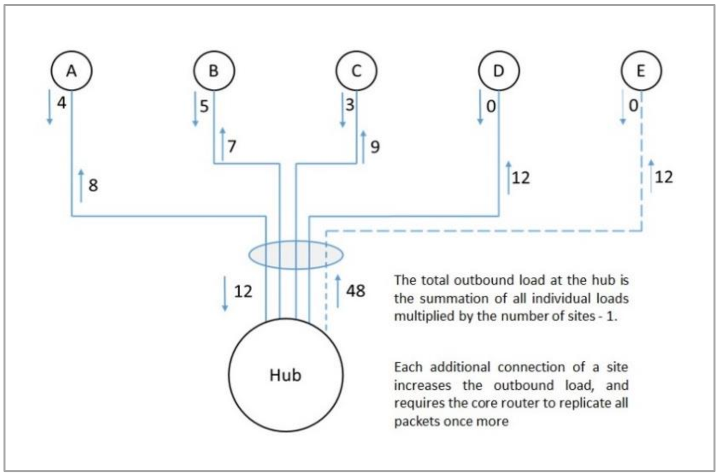
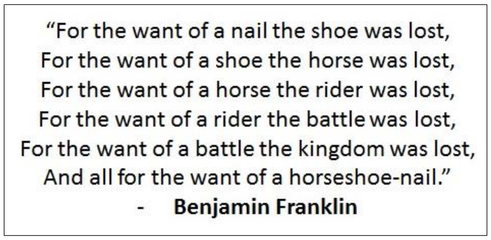

A
B
C
D
E
4
8
5
7
3
9
0
12
0
12
12
48
Hub
The total outbound load at the hub is the summation of all individual loads multiplied by the number of sites - 1.
Each additional connection of a site increases the outbound load, and requires the core router to replicate all packets once more

Figure 3. Example of the Effects of Multicast in a Hub and Spoke Network

[3] Generic Routing Encapsulation (**GRE**) is a tunneling protocol developed by Cisco that allows the encapsulation of a wide variety of network layer protocols inside point-to-point links. A GRE tunnel is used when packets need to be sent from one network to another over the Internet or an insecure network.

Situation

The integrity and context of the collected data is essential in any meaningful science experiment! As in Benjamin Franklin's quote, what appears to be a minor omission, can have far-reaching, and exotic consequences to causality. This is especially true, when the losses are unnoticed. The undetected absence of some data can create avalanches of compounded false positives and negatives creating statistical noise that looks like valid signal. An analysis based on such noise may suggest that a weapon is ineffective against a particular target at a certain range, or that certain countermeasures did not provide the desired protection threshold.



"For the want of a nail the shoe was lost,
For the want of a shoe the horse was lost,
For the want of a horse the rider was lost,
For the want of a rider the battle was lost,
For the want of a battle the kingdom was lost,
And all for the want of a horseshoe-nail."
- **Benjamin Franklin**

Figure 4. Quote from Benjamin Franklin

PREVIOUS DIAGNOSTIC APPROACHES AND FIXES

Connectionless RTI to improve Scalability and Fault Tolerance

During the 2013 BLCSE simulation exercise (SIMEX), which involved about 24,000 simulated entities, reliable messages managed by the Run-time Infrastructure (RTI) Executable experienced substantial failures. Federates would constantly drop off and rejoin to alleviate Central RTI Component (CRC) hang-ups, which became very disruptive to the ongoing human-in-the-loop war-gaming.

The diverse, geographically-dispersed federates intermittent presence on the network, made the federation almost unmanageable. Although LAN and WAN instability were undeniably a major cause, it was impossible to accurately quantify these issues for the network engineers to act upon. The federation utilized the RTI-1.3NGmatrex software and the accompanying MATREX Federation Object Model (FOM). During the months following the event, the architecture was revised to a connectionless RTI hosting heart-beated simulation objects. On the wire, this radical change reduced the previous traffic routing problems of "reliable" and "best effort" traffic in combination, to purely multicast issues, which were anticipated to be more straightforward to track and correct. The immediate benefit of the connectionless RTI configuration was that individual federates could join and resign the federation without the Central Runtime Component (CRC) hang-ups that were observed during the SIMEX. While focusing all our efforts on best-effort traffic, it soon became obvious that this traffic was heavily affected by other coinciding network loads; locally and throughout the network.

In the same SIMEX, a major identifiable inadequacy was the sizable discrepancies in entity counts as seen by simulation federates at each site. The initial reaction and approach was to develop systems that would compare and report on entities received to quantify the losses at a glance. A OneSAF based HTML/Web-browser displayed tool, named Reporter, was developed in early 2014 and has been in service since. It enables staff and participants to quickly detect differences in entity counts and birddogs suspected WAN links for troubleshooting. This innovation was a decisive step in visualizing and benchmarking data losses by tracking entity counts across the federation. Diagnosing and localizing the cause of these losses proved to be a formidable task, much bigger than was apprehended!

Simulation Dashboard to Manage Network and Simulation Loads (Kieslich, 2013)

An initial "passive/reactive" strategy, to mitigate high loads, was to build a dashboard that consisted of three main capabilities/Systems:

(1) Reporter, a web interface of a OneSAF composition that shows entity counts at each site. Because it used a browser it can be viewed anywhere on the network.

(2) Message Reporter, a tool written in C++ that reports on specific multicast groups on the network. The tool is based on tcpdump[4], and is configured to show color coded, continuously updating, stacked line charts that show bandwidth utilization and packet counts by HLA routing space.
(3) Statseeker, a commercial off the shelf (COTS) tool that uses Simple Network Management Protocol (SNMP) to record utilization to of each router and switch interface to a database. An interface to Statseeker was developed to provide a continuously updating line chart showing the actual network total load at each router interface inbound and outbound at a glance.

The advantage of the Simulation Dashboard was the ability to view loads from anywhere on the network, and in case of rising cumulative loads, mitigating steps could be taken to alleviate an outage or severe losses. The disadvantage was that network management was a mostly manual process, and direct intervention by the operations staff to apprehend and mitigate overloads. As the demands and expectations for the simulation exercise grow, and the increased use of cutting-edge web technologies become more prevalent, monitoring alone becomes insufficient and more passive in nature, and speculative or reactive administration of network resources is increasingly ineffective. Additionally, the “monitor and react” strategy did nothing to directly address packet loss, but tried to avoid situations where packet loss is imminent. Another disadvantage was that some packet loss occurred, but was undetectable. Yet, this approach of visually tracking load levels on the various components of the Simulation Dashboard laid the diagnostic foundation for the implementation of QoS.

The Difficulty Diagnosing Distributed Environments

In order to devise a solution that preserves simulation data traversing between all endpoints in real-time, under conceivable conditions, the tendencies for loss must be imagined, simulated, measured and thoroughly studied. The simulation federate software itself is not a useful tool for this purpose for these reasons: (1) the packets on the wire do not directly correspond to the models in the simulation, and (2) it is difficult to generate the proper load for troubleshooting without involving a lot of staff and numerous operators to run the scenario. Additionally, packet losses have been known to occur in inconsistent places.

None of the available Commercial Off of the Shelf (COTS) network stimulation tools: (iPerf, Nuttcp) provided enough directional or “destinational” robustness to generate a representative, distributed load nonstop to allow troubleshooting while the load was present, and none generated statistics about the integrity of each packet within the entire network. It is widely believed that large problems can be dissected into smaller chunks and solved that way – many can, but it is not true in this case. At this point, the goal was to adapt an existing tool, or develop a new one to create a distributed multicast load that could be operated in the following modes of operation: (1) Point-to-point, (2) Point-to-multipoint, and (3) Multipoint-to-multipoint. Under a steady load in each mode the cause of data packet loss could be determined and corrected.

PREREQUISITES FOR A SUCCESSFUL QOS IMPLEMENTATION

For a successful QoS implementation, any lower layer throughput matters have to be fully addressed. The initial fixes listed below can be considered prerequisites for a successful QoS implementation in a distributed environment. Problems at the physical layer are impossible to overcome at a higher layer, especially on the WAN.

Common WAN Connectivity Issues

Following the “bottom up approach” of the OSI[5] model, several issues of data packet loss on the WAN were easily identified and addressed. Throughput limitations caused by buffer overruns at encryption devices were discovered in nearly all locations and increasing the send and receive buffers was the easy fix. Issues with auto-negotiation of a router or switch with an in-line encryptor left the interface in either half-duplex mode, or configured for a lower speed

[4] **tcpdump** is a common packet analyzer that runs under the command line. It allows the user to display TCP/IP and other packets being transmitted or received over a network to which the computer is attached.
[5] The Open Systems Interconnection model (**OSI model**) is a conceptual model that characterizes and standardizes the communication functions of a telecommunication or computing system without regard to their underlying internal structure and technology.

after power outage. Once back online, the throughput limitation is not obvious until that traffic condition is met. Therefore, both interfaces are preconfigured to a fixed speed, eliminating a poorly auto-negotiated and functioning, half-duplex, default state.

Another problem surfaced at peak load, causing some burstiness of the HLA traffic. Low-end media converters merely used to convert from ST fiber connectors to SC were used at many locations where the DREN connection came onto the installation. The performance of the faulty connections was marginal and accounted for some losses at certain loads. This problem was simply corrected by replacing the devices with a SM fiber patch cord equipped with the right connectors. During connectivity and load testing of the WAN circuits, another source of packet loss was discovered; packet loss during very low multicast load or at the startup after a pause. This problem was difficult to identify because of the assumption that if the network supports a full load, a low load will never pose a problem.

Protocol Independent Multicast – Sparse Mode (PIM-SM) (B. Fenner, 2006) is a multicast routing protocol that builds unidirectional shared trees rooted at a Rendezvous Point (RP) per group, optionally creating shortest-path trees per source. Routing trees are established and collapsed by “Join” and “Prune” messages, respectively, involving source router, destination router and the RP. However, if there is no multicast routing activity of a given group, for greater than 180 seconds, the “Expiry Timer” times out and the tree collapses, even without a distinct “Prune” message ordering it to do so. Consequently, there is significant simulation data multicast UDP latency and loss when the trees get repeatedly rebuilt. In certain HLA routing spaces it is very common that no messages are sent for time periods exceeding 180 seconds by far. Each time a routing tree collapsed and needed to be re-established, up to 150 messages were lost and when the routing tree was back up, several duplicate packets were resent. This issue was identified as a by design, characteristic of the PIM-SM routing protocol. The solution to the problem was to keep the routing trees from collapsing by developing a low load stimulation application based on Beacon tool described later. This capability is now incorporated into the MATREX Local Runtime Component (LRC), fully resolving this issue, and making the MATREX RTI WAN-proof. The resolution of this issue was critical, because QoS would have been unable to correct this technology integration conflict, in original design conventions, between the multicast routing algorithm and the simulation protocol.

Diagnosing Packet Loss in Distributed Environments

After all lower layer connectivity issues had been fully addressed, and Iperf point-to-point testing passed in all cases, occasional packet losses could still be confirmed by inaccurate entity counts at some sites; however, the delta in entity counts was not consistent, and neither locations where differences were observed. The available COTS tools were unable to identify the data packet loss. These tools operate on the premise that any problem can be dissected into smaller chunks and localized in that manner. In this case this methodology does not lead to success! These tools merely look at point-to-point packet flow, but are unable to diagnose deficiencies in a full multipoint-to-multipoint packet distribution on the network.

To effectively analyze the problem, a representative load condition was needed that remained steady for the course of troubleshooting. The first attempt to mimic the simulation network load was through a custom tool called Beacon. It generated and subscribed to a continuous stream of 1300-byte multicast packets. Incoming data packets were logged to a file, and their counts were compared with the sent packets. This process was cumbersome, but was effective in finding some losses. The benefit of this strategy was to expose a hardware inadequacy in servicing the data packet distribution over the substantial number of GRE Tunnels. As a result, the Ft. Gordon CESC upgraded its core architecture and in-line encryptors. This upgrade positively corrected data packet loss caused within the core architecture. Additionally, the Beacon tool helped establish performance benchmarks.

THE FINAL SOLUTION TO ANY RESIDUAL DATA LOSSES: QOS

With the immense investment of a BLCSE experiment, as well as the months of preparation, the ARCIC leadership wants certainty that under conceivable load conditions, the simulation is not compromised, and experimental objectives are met. Quality of Service is an industry-wide set of standards and mechanisms for ensuring high-quality performance for critical applications. It allows the prioritization of certain types of network traffic to ensure that crucial applications or protocols always receive the required priority and amount of bandwidth.

Goals of QoS

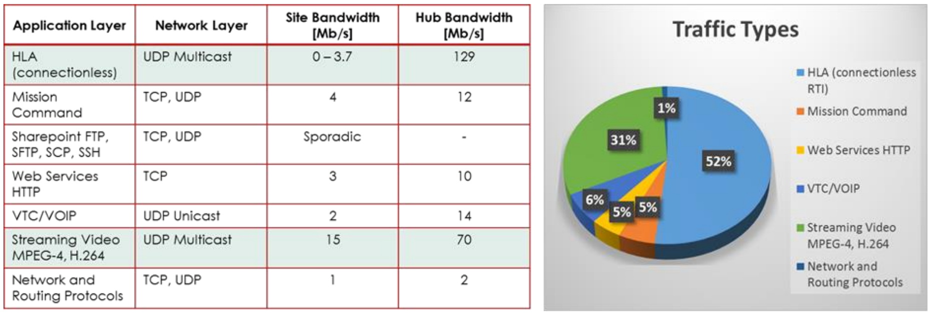
ARCIC has established the primary network objective as eliminating HLA packet loss under any circumstance. Latency and jitter are often mentioned in conjunction with WAN performance, but HLA is quite resilient to latency. On BLCSE, the Round trip time (RTT) to the farthest location is 95 milliseconds. Latency and Jitter have not had adverse effects on the BLCSE simulation. Queuing and buffering, although they add some latency, are negligible. Therefore, the focus of the QoS strategy is on the preferential treatment of HLA traffic. Avoidance of network congestion, is a significant concern. On BLCSE the LAN speed is 1 Gb/s, and the WAN speed is merely 50 Mb/s or 100 Mb/s, which is only 5% or 10%, respectively. WAN access has to be carefully managed to avoid over-extension of the available bandwidth. The cost is dropped packets on the DREN side and thus is beyond the control or correction of the team. Queuing or shaping of traffic may be necessary to overcome WAN access bottlenecks. The impact on other applications or services on the same network must be minor. Additionally, the need to manually manage network load should no longer be required, even at a growing environment.

Class-Based QoS

At the router, incoming packets are checked, and are assigned to traffic classes based on matching criteria. Traffic classes are categories of traffic (packets) that are grouped on the basis of similarity. Traffic that fails to match any defined traffic class is assigned to a default class of traffic. Such groups of traffic are called class maps. In order to prioritize traffic, a series of QoS policies have to be created and applied to one or more class maps. Finally the QoS policy is applied to either the inbound or outbound side of a router interface. On BLCSE the load pattern is constantly changing. The scenario of the next SIMEX, in 2017, is expected to encompass nearly 100,000 simulated entities. It will include more federates than ever before and additional sites will come on line. New web-based tools and services will be added, more video and analytical collaboration will be required. This expansive welter of progress will further contribute to the complexity and determination of network integrity. All data traffic loads have to be carefully considered when designing the policies.

Conflict with Other Traffic

Figure 5 provides examples from a previous exercise. A file download at a site from the SharePoint caused a significant load increase on the site's WAN connection, resulting in major data packet losses of the simulation traffic. During the file transfer, the site's Reporter page indicated over 4000 missing entities. This occurred during a Spiral exercise that involved only about 12,000 entities total. TCP/IP uses the sliding window protocol[6], which aggressively increases bandwidth utilization during a file transfer. Simultaneous presence of a file transfer and Multicast stream leads to significant losses of the best-effort multicast packets. The subscription to dozens of full motion video (FMV) multicast streams generated by UAV platforms at several sites, caused a much more widespread network overload, including at the network's core. This type of unanticipated activity by the sites, almost jeopardized the analytical worth of the entire experiment.



Application Layer	Network Layer	Site Bandwidth [Mb/s]	Hub Bandwidth [Mb/s]
HLA (connectionless)	UDP Multicast	0 – 3.7	129
Mission Command	TCP, UDP	4	12
Sharepoint FTP, SFTP, SCP, SSH	TCP, UDP	Sporadic	-
Web Services HTTP	TCP	3	10
VTC/VOIP	UDP Unicast	2	14
Streaming Video MPEG-4, H.264	UDP Multicast	15	70
Network and Routing Protocols	TCP, UDP	1	2

Traffic Types
HLA (connectionless RTI)
Mission Command
Web Services HTTP
VTC/VOIP
Streaming Video MPEG-4, H.264
Network and Routing Protocols
52%
1%
31%
6%
5%
5%

Figure 5. Other traffic on BLCSE

[6] A **sliding window protocol** is a feature of packet-based data transmission protocols. Sliding window protocols are used where reliable in-order delivery of packets is required, such as in the Data Link Layer (OSI model) as well as in the Transmission Control Protocol (TCP).

Prioritization of Traffic

Per direction from LTC Barry, Branch Chief, Technical Integration & Research at ARCIC, JAMSD, the focus of the military leadership is on the fidelity and dependability of the simulation. Capability development of the simulation at key battle labs is a major investment of TRADOC/ARCIC. GAMEX's and the TIEs which directly feed into the development toward a SIMEX. Figure 6 shows the significance of the SIMEX at the top of the pyramid.

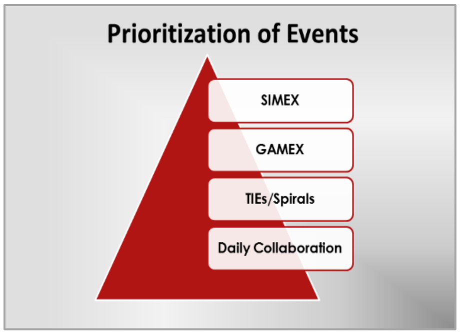


Figure 6. Priority of SIMEX

While communication and collaboration services are heavily utilized in the preparation of an experiment, they rank lower in criticality compared to the simulation exercise. Therefore, the highest priority of traffic on BLCSE is that of the HLA protocol, because it provides the medium of inter-communication between simulation federates and the experimental objectives that directly depend on it. Other, non-simulation, services are expected to work reasonably well all the time. However, the simulation must work "perfectly well" at any given time, and hold up to any analytical scrutiny. Thus, if an unexpected overload condition happens to occur, other services must be provisionally affected, rather than the simulation.

Effects on other Applications

In the case of inadvertent file uploads/downloads, because it is TCP/IP, a reliable mechanism, the transfer is expected to slow down in favor of the HLA traffic, but will not cease to work. Streaming video uses multicast as a transport, which can rapidly overload the total outbound bandwidth at the CESC, because of the required replication of packets in the hub and spoke architecture. Depending on the number of streams and the total load, the video may exhibit degradation or malfunction as a result. The load caused by streaming video has to be carefully planned to avoid congestion by means of too many streams to too many endpoints. VTC and VoIP have been thoroughly evaluated. The available bandwidth is sufficient for simultaneous use of VoIP and VTC to the limit of the servers while the simulation is running. In the future these real-time tools should fall into a QoS class of their own, that gets priority over the rest of the traffic, but only second to HLA. Web services use a combination of TCP/IP and UDP. During peak loads users may experience some sluggishness in performance, but can always wait or intervene by reloading a page.

Packet Marking for End-to-End QoS

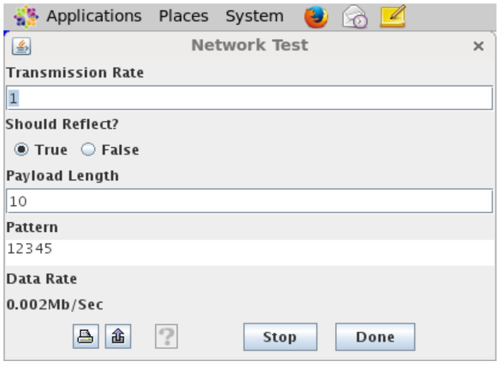
Marking is similar conceptually to "service class" designation on an airplane ticket: first, business, or economy. This value reflects the level (quality) of service the passenger should receive. Similarly, the application (or the router) marks a value in the packet to indicate the service class (henceforth termed service-class) for that packet as it traverses the network. By looking at the marked value, network elements can decide how to treat the marked packet. People in business-class may have used a variety of means to achieve that designation. They may have paid extra, used air-miles, or been lucky and booked at the normal rate when no other seat was available. Elsewhere, someone performed the complex task of classification - determining eligibility for a particular service-class then marked the ticket with a mere designation: first-class, business-class, or economy-class. The flight-attendant is unconcerned with how eligibility was determined; he or she simply looks at the class marked on the ticket and provides that level of service. This dynamic plays out in the networking world. One device may perform complex classification on the data in a flow, determining an appropriate service-class for that flow. Other network elements "trust" the value marked in packets they receive and provide service appropriate for that designation (Cisco, 2008) The QoS policies define the priority through a Differentiated Services Code Point (DSCP) value. The DSCP applies a value (0–63) within the Type of Service (TOS) field in an IPv4 packet's header. This DSCP value provides classification at the Internet Protocol (IP) level, which routers can use to decide queuing behavior. Packet marking of HLA messages is achieved by setting the DSCP value to "EF" (Expedited Forwarding) in the HLA RID file. Marking the service class in the application rather than in the switch or router, is less CPU-intensive than classification elsewhere in the network. As the marked packet traverses the network, each network device can act on its marking and provide the proper treatment. Unmarked packets will be forwarded on a best-effort basis, or discarded if the bandwidth threshold gets exceeded.

QoS Policy Development

BLCSE is an active environment year-round. Testing of QoS was deemed too much of a burden on ongoing integration efforts. Therefore, the goal was to develop and demonstrate the capability on a bench network and document each step for subsequent implementation on the production network. As the simulation federate is unable to generate the required load levels to develop the QoS policies on the test bench, a representative network load generator was needed. Ideally, this tool uses the HLA Local Runtime Component (LRC), just like a simulation federate, but is able to generate a constant stream with an adjustable transmission rate. This decision proved to be a vital step. Having the exact traffic transmission characteristics as the packets generated by a federate is closer to the actual simulation load than any of the COTS tools could possibly deliver. The LRC affords bundling, buffering and other traffic shaping methods, which are already optimized for the BLCSE federation. The idiosyncrasies of the HLA traffic at various loads are key to devising the proper QoS solution. In addition to the packet generation, another custom tool is needed and was developed, that is capable of measuring and accounting for every packet as it traverses each network link.

NetTest and NetTest Reporter

This new tool is a OneSAF composition named Network Tester, created by Paul Monday of JAMSD. The tool employs "Network-Test" interactions as a means to generate a continuous multicast stream, routed to all subscribers on the network. Packet size and transmission rate can be specified in the GUI. The tool offers a Mirror mode, in which a single stream suffices, because it is reflected by other sites/federates as show in Figure 7. Having been reflected once, a stream is not reflected again. With this simple set up, a network load can be generated that replicates the network characteristics of a fully loaded HLA simulation (100k entity scenario) by providing a steady load at a specified bit rate. An additional benefit is that the Network Tester involves the majority BLCSE federate: OneSAF, as well as the MATREX LRC. The toolset faithfully simulates the network traffic of an actual simulation.



Network Test

Transmission Rate

1

Should Reflect?

● True ○ False

Payload Length

10

Pattern

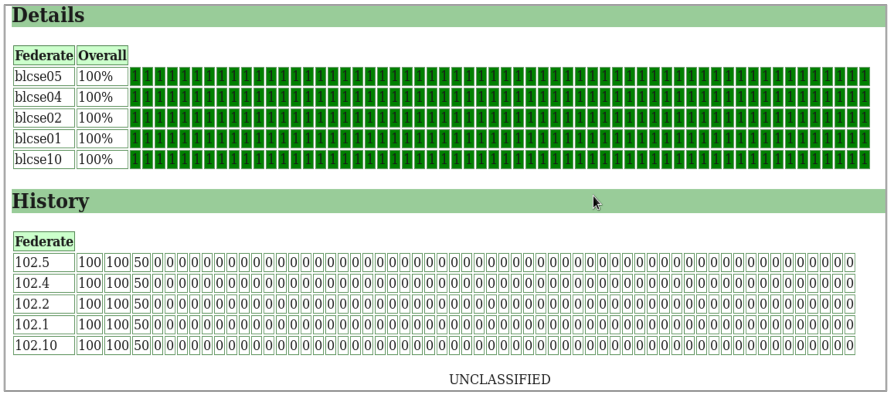
12345

Data Rate

0.002Mb/Sec

Figure 7. NetTest GUI

The NetTest Reporter, another OneSAF composition, provides statistics on how many NetTest interactions are received of a specific count per second. The NetTest interaction contains the transmission rate as part of the payload of the data packets sent within a second. The NetTest Reporter page as shown in Figure 8 is a web-based interface with the purpose of visualizing the integrity of packets sent vs. received, displaying how many packets are received within a second. The NetTest Reporter uses a color coding scheme to clearly illustrate the success rate: green means 100%, red means significant losses (>5%) and yellow means marginal losses (3% to 5%). Multiple browser windows can be opened on the same screen to compare the counts of various endpoints. The tool records transmission success rate in a table every 30 seconds, thus making it possible to observe the streams over time.



Details

Federate	Overall																																																												
blcse05	100%	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
blcse04	100%	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
blcse02	100%	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
blcse01	100%	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
blcse10	100%	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1

History

Federate																																																												
102.5	100	100	50	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
102.4	100	100	50	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
102.2	100	100	50	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
102.1	100	100	50	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
102.10	100	100	50	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

UNCLASSIFIED

Figure 8. NetTest Reporter Page

Test Bench Setup

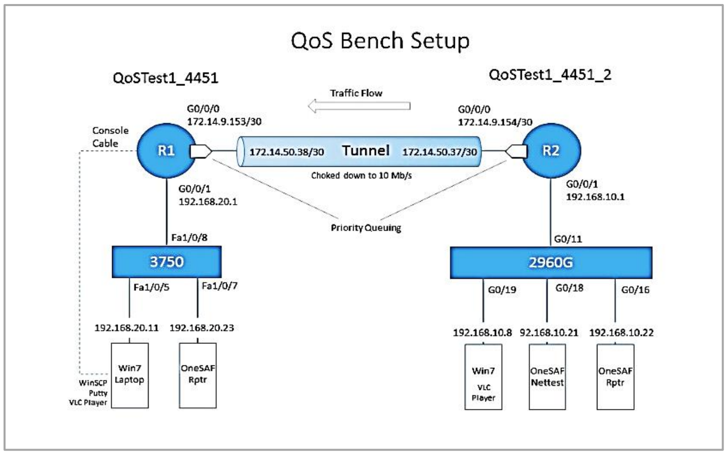


Figure 9. QoS Test Bench

To demonstrate this capability on the test-bench, a WAN link is emulated, using two routers, two switches, and several workstations. To emulate the reduced bandwidth of the WAN, the router interfaces are choked down to 10 Mb/s. Two 4451 CISCO routers connected via a GRE tunnel represented the common WAN connection of a site with Ft. Gordon. Each router is connected to a LAN switch and several workstations are connected to each switch. The bottleneck in each direction is at router interface G0/0/0 where its streams exceeding 10 Mb/s results in dropped data packets. A NetTest Reporter is placed on each LAN to show packets dropped locally.

Unanticipated Loss of NetTest Interactions

The bench test is configured to support up to 10 Mb/s of any traffic. However, when transmitting a message count of a mere 700 messages/second or 2.68 Mb/s, the NetTest Reporter demonstrated an unexpected packet loss. The burstiness of the LRC scheduled output to the network causes the bandwidth threshold to be momentarily (but regularly) exceeded every 200 ms.

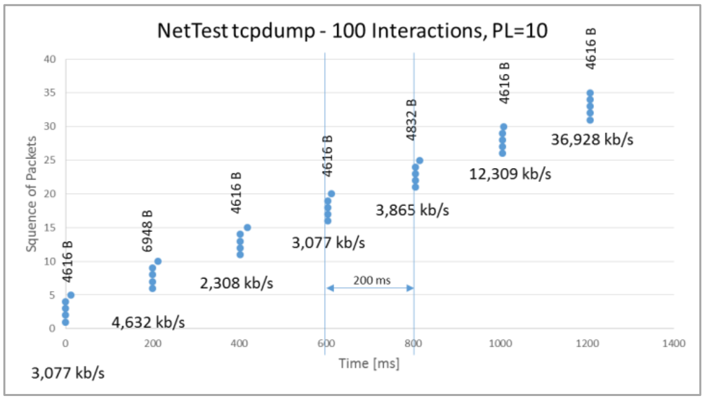


Figure 10. LRC transmitting 100 Interactions

Analysis of Lost Packets

Using tcpdump, the HLA packets are written to a file and analyzed, showing that "bundling" of packets in the LRC is problematic for the routers. At only 100 interactions per second, during a 30 second test, a distinct pattern, as shown in Figure 10, is seen. About five packets (4,616 bytes) are transmitted every 200 milliseconds. At 3,000 interactions (7.19 Mb/s) about 147 packets, totaling (150.3 KB) are sent every 200 milliseconds at the speed of the network interface. See Figure 11.

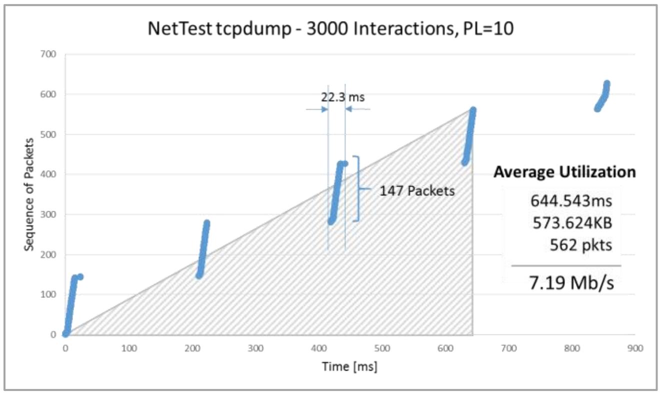
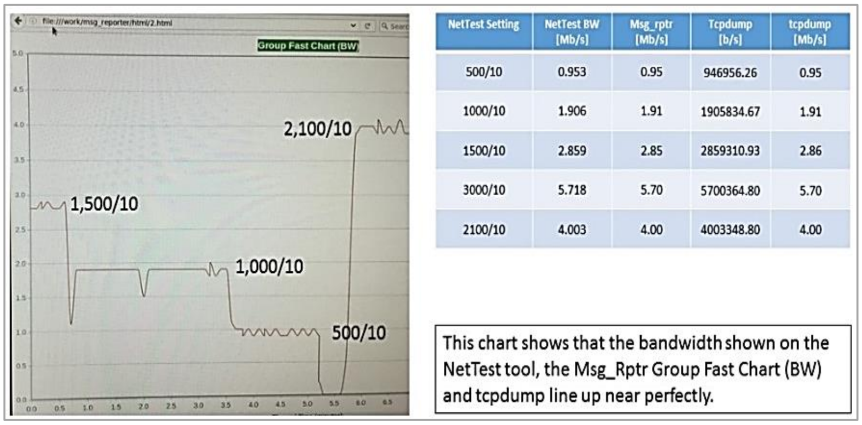


Figure 11. LRC transmitting 3000 Interactions

Table 1. Transmission Pattern

Time [ms]	Size [B]
0	1300
0.039	1300
0.053	1300
0.066	196
1.859	1300
1.885	1300
1.899	1300
1.911	196
3.107	1300
3.113	1300

The maximum packet size is 1,300 bytes. It is verified by running a OneSAF scenario, that the traffic pattern at a load, generated



NetTest Setting	NetTest BW [Mb/s]	Msg_rptr [Mb/s]	Tcpdump [b/s]	tcpdump [Mb/s]
500/10	0.953	0.95	946956.26	0.95
1000/10	1.906	1.91	1905834.67	1.91
1500/10	2.859	2.85	2859310.93	2.86
3000/10	5.718	5.70	5700364.80	5.70
2100/10	4.003	4.00	4003348.80	4.00

This chart shows that the bandwidth shown on the NetTest tool, the Msg_Rptr Group Fast Chart (BW) and tcpdump line up near perfectly.

Figure 12. Validation of NetTest bandwidth Readout

by 20,000 moving entities, produces a very similar pattern as shown in Table 1. The greater the load, the more “full” packets are sent; usually more than 70%. Another custom tool, called the Message Reporter, provides a continuously updating line chart of Multicast packets received on the network by group, or in the case of HLA, per routing space. Stacked charts are available to show either bandwidth or packets (messages). This tool is used to validate the bandwidth settings on the NetTest tool. The chart and table show that output from NetTest, tcpdump and Message Reporter correspond closely. Repetitive spikes, every 200 milliseconds, place a peculiarly perverse load on the network, deliberately causing interface buffers to fill up quickly and packet drops to occur.

Priority Queuing[7] (PQ)

Smoothing Benefit Observed

Considering the use cases and the significance of the HLA traffic, CISCO Priority Queuing on the outbound router interface is the most sensible QoS solution. With Priority Queuing enabled, the router allocates memory to hold traffic in the queue for subsequent transfers. This feature reduces high traffic spikes and provides smoothing. On the test bench, HLA traffic in volume up to the connection’s bandwidth threshold, traversed reliably over the emulated WAN link. Both NetTest Reporters displayed a 99 % success rate of NetTest Interactions received. Satisfyingly, with PQ in place, the router was routing HLA packets up to almost 10 Mb/s. During a concurrent file transfer of a very large file that lasted about 15 minutes, no HLA packets were lost. Video streams sent over the network using VLC Player® at the same time resulted in no packet loss, of either application, as long as the total available bandwidth was not exceeded. When it was exceeded HLA clearly prevailed without losses.

HLA Traffic vs. Video Stream

Table 2 shows the test results of sending a video stream over the same connection concurrent with HLA traffic. The NetTest Reporter showed near 100 % success during the entire test, while the video stream lost packets as HLA started consuming more bandwidth.

Table 2. HLA Traffic vs. Video Stream

HLA BW [Mb/s]	Video Stream [Mb/s]	Total BW [Mb/s]	RPTR Msg Loss [%]	Video Quality
2.68	4.75	7.39	0	good
3.83	4.75	8.38	0	good
7.74	1.61	9.35	0	bad
	Video packets dropped to preserve HLA stream			

HLA Traffic vs. TCP/IP

While sending HLA traffic at a bandwidth of 9.58 Mb/s a file transfer was started. The transfer speed was at 176 kb/s (very slow). Both Reporters showed zero loss. When the OneSAF NetTest was stopped from sending HLA traffic, the transfer speed jumped to 9.6 Mb/s, demonstrating the QoS constrained speed of the tunnel.

QoS Implementation on the Production Environment

The implementation, of the bench tested QoS solution, on BLCSE went smoothly. The same router commands were applied and testing on the bench was performed, albeit at a larger scale. The results were analogous to those seen on the test bench from the sites to the network core at the CESC. In the other directions multiple GRE tunnels share the

[7] Priority Queuing (**PQ**) ensures that important traffic gets the fastest handling at each point where it is used. It was designed to give strict priority to important traffic.

same in line encryption devices and switch ports. This part of the configuration could not be evaluated on the test bench. PQ was applied to the physical interface in outbound direction. Testing has been very successful in terms of recorded metrics.

DAILY READINESS TEST (DRT)

The CESC/JAMSD team developed a Daily Readiness Test that exercises the network using NetTest in conjunction with a simple OneSAF scenario. By applying a simple formula using the expected load, available bandwidth and site reflector count, the NetTest settings can be determined at the outset. The test can be run before a simulation event or whenever a problem is suspected. By applying concurrent loads, such as streaming video or file transfers; the network, the federation, and the QoS configuration are mutually evaluated as part of the DRT.

WAY AHEAD

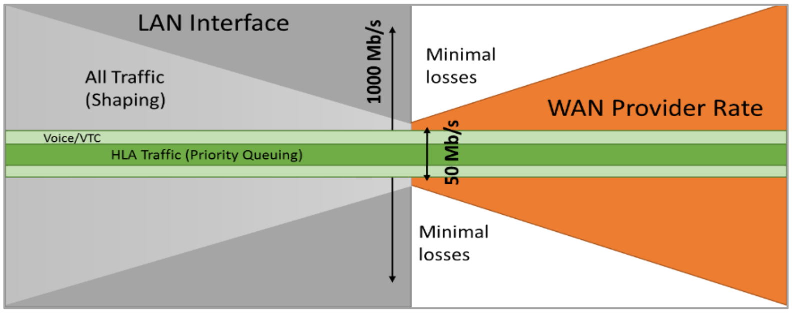


Figure 13. Future Enhancements of QoS: Shaping

QoS Priority Queuing has ensured that known data losses in the experiment are eliminated in the current load state. However, if the load of competing assets increases more significantly, additional policies may be needed to ensure communications such as simulated Mission Command systems, as well as VoIP and VTC receive additional levels of protection. At present, WAN components at the sites are configured for 100Mb/s. However, in some cases only 50 Mb/s are provided. If the DREN implements a policing policy on the edge device interface, or starts charging for overuse, Traffic Shaping may become necessary as shown in Figure 13. A future innovation will be to incorporate the Daily Readiness Test into OneSAF for optional use during experiments. NetTest interactions are ignored by the simulation, but are counted by the Reporter. Similar to the proverbial canary in a coal mine, an intact stream of NetTest interaction is construed to be a confirmation by extension of an intact flow of other HLA traffic at the same time while other loads are present.

CONCLUSION

In the BLCSE environment, where the preservation and collection of data is as critical as the simulation itself, continuous efforts to increase data reliability have each brought considerable success. As the overall load on the network continues to proliferate, a multitude of other applications directly compete with mission-critical data for bandwidth. Several preparatory steps, including the development of new tools were essential to implementing and complementing the class-based QoS on the BLCSE. The results are gratifying in two ways. The throughput of the bundled HLA messages became more efficient, while the prioritization of HLA traffic results in no losses from end-to-end. The effects on the secondary applications are not excessively disruptive. Packet losses of audio and video streams at full load have not shown a human-detectable degradation. A slowing of web-services and file transfers is very minimal (< 5%) at full load. Fine-tuning of the QoS policies may be necessary as load levels increase further in the future. Further development and enhancement of the DRT into the applications and schedule, so that a single administrator can quickly verify the preferential treatment of the simulation protocol under full load are likely the next steps.

ACKNOWLEDGEMENTS

We would like to acknowledge the following individuals for their support. Paul Monday's advice and work was significant in the overall approach. Brian Perry was instrumental in the integration of QoS on BLCSE. Thanks to Don Meinshausen for his help in preparing the data for this paper. And finally, to LTC John Barry, Warren Jones, Favio Lopez, Roy Zinser, Andy Gross, Brian Hobson, Tiffany Kinsella, Donnell Walker and the entire BLCSE Community of Practice for their patient provision during testing and implementation of this capability.

REERENCES

Kieslich & Zinser (IITSEC 2013). Simulation Dashboard to Manage Multi-Federate Simulation Events
On-line Resource (2008). Implementing Quality of Service Policies with DSCP, CISCO Document ID: 10103
B. Fenner (2006). CISCO RFC: 4601 Protocol Independent Multicast – Sparse Mode (PIM-SM)