Copy No.________

Defence Research and Development Canada
Recherche et développement pour la défense Canada

DEFENCE DÉFENSE

Final Report on the Investigation of Asynchronous Events in Non-Time-Managed Federations

Dan Bleichman
CogSim Technologies

CogSim Technologies
83 Ossington Avenue, Suite 1
Ottawa, Ontario
K1S 3B5

Contract Number: W7707-042793/001/HAL

Contract Scientific Authority: Mark Hazen, 902-426-3100 x176

Defence R&D Canada – Atlantic

Contract Report
DRDC Atlantic CR 2005-182
September 2005

Canada

This page intentionally left blank.

Final Report on the Investigation of Asynchronous Events in Non-Time-Managed Federations

Dan Bleichman
CogSim Technologies

CogSim Technologies
83 Ossington Avenue, Suite 1
Ottawa, Ontario
K1S 3B5

Contract number: W7707-042793/001/HAL

Contract Scientific Authority: Mark Hazen, 902-426-3100 x176

Defence R&D Canada – Atlantic

Contract Report

DRDC Atlantic CR 2005 – 182

September 2005

Author

Dan Bleichman

Approved by

Mark G. Hazen

Contract Scientific Authority

Approved for release by

Kirk Foster

DRP Chair

Abstract

A study was developed and conducted at DRDC Atlantic to examine the relationship between (1) Wide Area Network (WAN) characteristics (i.e., latency and latency variability) and (2) High Level Architecture (HLA) federation characteristics (e.g., frequency of the occurrence of events as indexed by the frequency of the exchange of messages, event handling time and federation size) to the probability of an asynchronous event occurrence in non-time-managed federations.

The study was conducted using federates sending or receiving wall-clock time-stamped HLA messages over an emulated WAN while all hardware clocks of the participating machines were synchronised. The time-stamp of each received message was examined and used in order to determine out-of-sequence arrival.

Results show that in a non-time-managed federation executing over a WAN with realistic network delays, the probability of asynchronous events occurring is significant and depends on both network characteristics and message rates.

Résumé

Une étude a été élaborée et menée à RDDC Atlantique dans le but d'examiner les relations entre (1) les caractéristiques d'un réseau étendu (WAN) (temps d'attente et variabilité du temps d'attente) et (2) les caractéristiques de regroupement d'une architecture de haut niveau (HLA) (fréquence d'occurrence des événements indiquée par la fréquence d'échange des messages, temps de traitement des événements et taille des regroupements) par rapport à la probabilité d'occurrence d'événements asynchrones dans les regroupements qui ne sont pas gérés en fonction du temps.

L'étude a été effectuée à partir d'éléments regroupés qui émettaient ou recevaient des messages HLA horodatés sur un WAN émulé alors que toutes les horloges matérielles des appareils participants étaient synchronisées. L'horodatage de chaque message reçu a été examiné et utilisé pour déterminer l'arrivée hors séquence.

Les résultats révèlent que, dans un regroupement non géré en fonction du temps et fonctionnant sur un WAN à temps de traitement réaliste, la probabilité que se produisent des événements asynchrones est élevée et dépend à la fois des caractéristiques du réseau et de la fréquence des messages.

This page intentionally left blank.

Executive summary

Introduction

The Virtual Combat Systems (VCS) Group at Defence Research & Development Canada – Atlantic have run High Level Architecture (HLA) conservatively time-managed federations (simulations) over a wide area network (WAN), and experienced a significant slow-down compared to running on a local area network (LAN.) Measurements showed low CPU utilization and network traffic compared to the resources available. It was hypothesized that latency in the network leading to serialization of the time-management infrastructure was the main cause of the performance decrease. The alternative of running non-time-managed (receive-ordered) federations, however, would allow errors in the order delivery of messages. This study was conducted, under contract by COGSIM, to determine the impact of these errors including the number of errors in message order and their magnitude.

Results

Two sets of experiments were run on WANE, a WAN Emulator constructed by the VCS group which can emulate latency, bandwidth and transmission errors effects; one set for time-managed federations and the other for receive-ordered federations. The experiments used a variety of network conditions including constant and normally distributed latencies across the WAN. The results show that the probability of asynchronous events occurring is significant for non-time managed federations running with realistic WAN delays. The DMSO RTI 1.3NGv6 could not handle high message rates (>200Hz) but at lower rates (<5Hz) careful filtering allowed the use of 80% of received messages. The effects of network latency can be minimized by designing for uniformity in latency across the links. In time-managed federations performance was proportional to $\log_2(n)$*latency, where n is the number of federates.

Significance

This work has quantified tradeoffs between time and non-time-managed federations in the presence of network latency, which is a major federation design issue. The work has also clarified the impact of latency on HLA performance and proved the worth of the WANE system in distributed simulation experimentation.

Future Work

This contract was the first in a series of activities to investigate the use of HLA federations on WANs with significant latency. Future work will include investigations of serialization in conservative time-managed federations, the effect of run-time-infrastructure algorithms on federation performance, and other processes for mitigating the effects of latency on federation performance.

"Bleichman, D" 2005 *Investigation of Asynchronous events in Non-Time-Managed Federations*. DRDC Atlantic CR 2005-182. Defence R&D Canada – Atlantic.

Sommaire

Introduction

Le Groupe des systèmes de combat virtuel (SCV) de Recherche et développement pour la défense Canada – Atlantique a fait fonctionner sur un réseau étendu (WAN) des regroupements (simulés) à architecture de haut niveau (HLA) gérés prudemment en fonction du temps, et il a décelé un ralentissement significatif par rapport au fonctionnement sur un réseau local (LAN). Les mesures ont révélé une faible utilisation de la CPU et la présence sur le réseau de trafic peu intense comparativement aux ressources disponibles. On a supposé que la baisse du rendement résultait principalement du fait que le temps d'attente du réseau menait à la sérialisation de l'infrastructure gérée en fonction du temps. La possibilité de faire fonctionner des regroupements non gérés en fonction du temps (ordonnés selon la réception) permettrait toutefois des erreurs dans la livraison ordonnée des messages. Cette étude, menée dans le cadre d'un contrat avec COGSIM, visait à déterminer l'impact de ces erreurs, y compris le nombre des erreurs dans l'ordre des messages et leur grandeur.

Résultats

Deux ensembles d'expériences ont été effectués avec WANE, émulateur WAN constitué par le groupe des SCV et capable d'émuler les effets du temps d'attente, de la largeur de bande et des erreurs de transmission; un ensemble a porté sur les regroupements gérés en fonction du temps et l'autre, sur les regroupements ordonnés selon la réception. Les expériences se fondaient sur diverses conditions de réseau, notamment les temps d'attente constants et à répartition normale sur le WAN. Les résultats révèlent que la probabilité d'événements asynchrones est élevée dans les regroupements non gérés en fonction du temps sur un WAN à retard réaliste. Le logiciel DMSO RTI 1.3NG, version 6, n'a pas pu traiter les fréquences de messages élevées (> 200 Hz) mais, aux fréquences inférieures (< 5 Hz), un filtrage minutieux a permis l'utilisation de 80 % des messages reçus. Il est possible de réduire les effets du temps d'attente du réseau en établissant une conception axée sur l'uniformité des temps d'attente d'une liaison à l'autre. Dans les regroupements gérés en fonction du temps, le rendement était proportionnel à log2(n)*temps d'attente, où n désigne le nombre d'éléments regroupés.

Portée

Ces expériences ont permis de quantifier les compromis entre les regroupements gérés en fonction du temps et non gérés en fonction du temps lorsqu'il existe des temps d'attente sur le réseau, ce qui constitue un aspect important de la conception des regroupements. Elles ont aussi procuré des éclaircissements quant à l'impact des

temps d’attente sur le rendement de la HLA et démontré la valeur du système WANE dans les expériences de simulation par répartition.

Recherches futures

Ce contrat était le premier d’une série visant à étudier l’utilisation des regroupements HLA sur des WAN à temps d’attente considérable. Des recherches futures auront trait à la sérialisation dans les regroupements gérés prudemment en fonction du temps, aux effets des algorithmes d’infrastructure d’exécution sur le rendement des regroupements et à d’autres procédés permettant de réduire les effets du temps d’attente sur le rendement des regroupements.

"Bleichman, D" 2005 *Investigation of Asynchronous events in Non-Time-Managed Federations*. DRDC Atlantic CR 2005-182. Defence R&D Canada – Atlantic.

This page intentionally left blank.

Table of contents

Abstract i

Executive summary iii

Sommaire iv

Table of contents vii

List of Figures viii

1. Introduction 1

2. Test Procedures 5

    2.1 Probability of asynchronous events in a non time managed federation 5

        2.1.1 WAN setup 7

        2.1.2 Test Runs 8

        2.1.3 Results 9

    2.2 Message rate in a time managed federation 20

        2.2.1 Results 21

3. WAN Emulator 23

4. Conclusions 26

5. Future Work 27

6. References 28

List of symbols/abbreviations/acronyms/initialisms 29

Distribution List 30

List of Figures

Figure 2-1 A Non-Time-Managed Federation ... 6

Figure 2-2 Experimental Network Setup ... 7

Figure 2-3 Message Arrival on WAN: Graph 1 ... 10

Figure 2-4 Message Arrival on WAN: Graph 2 ... 11

Figure 2-5 Message Arrival on WAN: Graph 3 ... 12

Figure 2-6 Message Arrival on WAN: Graph 4 ... 13

Figure 2-7 Message Arrival on WAN: Graph 5 ... 14

Figure 2-8 Message Arrival on WAN: Graph 6 ... 15

Figure 2-9 Message Arrival on WAN: Graph 7 ... 16

Figure 2-10 Message Arrival on WAN: Graph 8 ... 17

Figure 2-11 Probability at 1 Hz ... 18

Figure 2-12 Probability at 5 Hz ... 18

Figure 2-13 Message Arrival on WAN: Graph 9 ... 19

Figure 2-14 Time-Managed Federation ... 20

Figure 2-15 Experimental Network Setup ... 21

Figure 2-16 Time Advance Grant Throughput ... 22

Figure 3-1 WAN Emulator ... 23

Figure 3-2 Actual federation deployment ... 24

1. Introduction

The Canadian Forces are making use of the High Level Architecture (**HLA**) to support a multitude of programs, ranging from training systems and distributed mission operations exercises, to component model & simulation development for research. Inherent to the existence of HLA federations is the role played by the Run-Time Infrastructure (**RTI**) and the management services that it provides. In the past, federations deemed to be “real-time” (i.e., human-in-the-loop simulations) have typically declined the use of RTI time management services due to the performance overhead imposed by invoking these services. Non real-time federations have been implemented with and without the use of time management services in a somewhat ad-hoc fashion, in many cases without a strong understanding of the implications of doing so, or without an effective means of quantifying these implications.

As a result, a need has arisen to develop an understanding of the issues involved in determining the suitability of using time-management services for a given federation. This question has impact on the type of RTI used and the complexity of the federation management required. DRDC Atlantic is developing a wide area network (**WAN**) emulator for conducting federation/RTI testing to investigate questions of this nature.

HLA time management is concerned with the mechanism for controlling the advancement of each federate along the federation time axis. A federate that becomes time-regulating may associate some of its activities (such as updating instance attribute values and sending interactions) with points on the federation time axis. A federate that is time-constrained requires that notifications of relevant updates be received (such as reflecting instance attribute values and receiving interactions) in a federation-wide, time-stamped order. Use of time management services allows this type of coordination among time-regulating and time-constrained federates in an execution. Time management in HLA guarantees the following:

- A federate receives messages in a time stamped order; and

- A federate won’t receive a message in its past (relative to it’s local simulation logical time).

However, a number of issues can arise with time management:

- At any instance during federation execution different federates can be at a different logical time;

- The duration of a single logical time unit can vary in terms of wall-clock time during the simulation; and

- The federation advances in time according to the **slowest performing** time-regulating federate.

Wall-clock time can be used as an alternative to logical time management (from the HLA prospect this is a non-time-managed federation). Use of wall-clock time usually requires time synchronization between all federates. It is possible to achieve an acceptable level of time synchronization with wall-clock time when executing on a Local Area Network (**LAN**). However, time synchronization on a WAN requires the use of Global Positioning System (**GPS**) receivers (one per site).

The use of the Network Time Protocol (**NTP**) may provide adequate means for achieving synchronization of wall-clock time across a federation, for some cases. Each outgoing message can be time stamped at the sender's site with an accurate local time. The receiver can then compare its local time with the originating time stamp and apply extrapolation (dead reckoning) on the received data or can simply decide to ignore it if data is stale or does not meet other timing criteria. This method **cannot** guarantee that messages are received in their correct order (asynchronous event occurrence) since variance in network delay between different federates may cause later generated messages to arrive at the receiving federate prior to messages that were generated at an earlier time. For example, consider a 3 federate federation with network delay of tndBA between federates B and A and network delay tndCA between federates C and A. Assume:

$$\text{tndCA} = \text{tndBA} + 0.5\text{sec}$$

and federate C sends an interaction to A at time t while federate B sends an interaction to A at time t+0.2 seconds. Federate A will receive the message from B first and respond to it and only later on will read the message from C even though the message from C represents an earlier event. Furthermore, asynchronous events can occur between two federates that are federated over a complex and slow WAN, resulting in a message being delayed to the extent that it arrives later than a subsequently sent message. Network delays between federates can be divided into two major categories: (1) The delay in the transmission time of messages between two federates produces a normal or Gaussian distribution with a constant mean. (2) In addition the delay in the transmission time of messages between different federates produces normal or Gaussian distributions but with different mean values. The problem described above can be effectively resolved in some cases by applying incoming message queuing and time-stamp based message reordering at the receiving federate.

The main obstacle when using a time-managed federation is the reduced performance due to costly time management overhead. In a time-managed federation, at each point on the logical time scale, certain simulation steps can take place. In order for the federation to advance to the next activity in simulation, the federation's time must be advanced. This is done through a request initiated by the federate and a grant given by the RTI. The process of computing the next granted time requires the calculation of the Lower Bound on Time Stamp (**LBTS**). The following considerations should be made with respect to the LBTS:

- LBTS value computed for a federate is the lower bound on the time stamps of messages that may be received and that are destined for that federate later in the execution;
- For a federate, the RTI must ensure that:
    - Time Stamp Ordered (**TSO**) messages are delivered to the federate in time-stamped order; and
    - No message is delivered to the federate with a time stamp that is smaller than its logical time.
- Once LBTS for a given federate is computed:
    - The RTI can deliver to the federate all TSO messages containing a time stamp less than LBTS; and
    - If the RTI prevents the federate from advancing its logical time beyond LBTS, it can guarantee that the federate will not receive any messages in its past (relative to local simulation logical time).
- To compute LBTS, the RTI must consider:
    - The smallest time stamp of any TSO message any federate might generate in the future (the current logical time of a federate is one bound since no federate can generate a TSO message in its past); and
    - The time stamps of messages within the RTI and the interconnection network (transient messages).

A single LBTS computation that is initiated each time that a federate issues a time advance request or a next event request requires the order of $N\log_2 N$ (Fujimoto & Hoare, 1998) messages to be exchanged between the federates (where N is the sum of time regulating and time constrained federates in the federation). All RTIs that implement time management services use an LBTS calculation algorithm with complexity of this order, but due to different means of dealing with transient messages the actual LBTS calculation may require between $2N\log_2 N$ messages and possibly double that figure. Some of these messages can be passed and processed concurrently but a significant degree of serialization will always be present (the amount of serialization is proportional to the depth of a binary tree with a federate at each of the tree's nodes). For example, consider a federation consisting of 8 federates (a binary tree of 8 nodes has a depth of 3), all time regulating and time constrained. When interconnected over a WAN with an average network latency of 0.1 seconds, in the worse case situation the minimal delay required so a federate could advance to the next point in time may take over 1 second. Such a delay would not be acceptable for a human-in-the-loop simulation.

The approaches to time management described above have benefits and drawbacks. A non-time-managed federation may allow faster execution, but at a cost of asynchronous event occurrence. A time-managed federation will eliminate this problem but may not meet the performance requirements of the simulation.

2. Test Procedures

Two major tests in an emulated WAN environment were completed. The first test measured the probability of asynchronous events and maximum, as well as average time, of the expected asynchronicities in a non-time-managed federation. The second test measured maximum and average message throughput in a time-managed federation of an identical deployment (identical emulated WAN environment and identical size of federation).

2.1 Probability of asynchronous events in a non time managed federation

The test federation in Figure 2-1 was deployed at DRDC Atlantic over the experimental network setup as shown in Figure 2-2 and consisted of a single receive federate (subscriber) and 3 senders on 3 machines (one federate per LAN) or 6 senders on 3 machines (two federates per LAN). Each node (PC) synchronized its internal hardware clock by connecting to an NTP server that resided on one of the sender machines. The network delays between the 3 sender machines were set to the minimum possible value, which is about 4 ms, one way. This allowed for sufficient degree of clock synchronization. High quality of time synchronization with receiver node was not required as only the time stamps obtained from the sending federates were used for detection of out of order messages. All sender nodes were connected to the receiver node through the WAN emulator (please refer to section 2.2.3 for a detailed description). The receiver node collected and logged the data about arrival of out of time stamp order interactions. After passing the first synchronization point, all senders started sending time stamped interactions to the receive federate simultaneously. The interactions were sent at a predefined frequency that varied between runs. The message frequencies were: 100 Hz, 50 Hz, 20 Hz, 10 Hz, 5 Hz, 2 Hz, 1 Hz and 0.5 Hz. The actual delay between messages was computed using a random component in order to better simulate “real world” conditions and to ensure that the data sending sequences were typically out of phase such that they did not fall into an “alignment” on the time scale. The random component was as high as 200% for very high (100 Hz, 50 Hz) message frequencies and was reduced to about 100% for lower message frequencies. For all experiments, messages that were out of time stamp order by less than 10 ms were not categorized. Consequently, these values can be attributed to suboptimal time synchronization between senders.

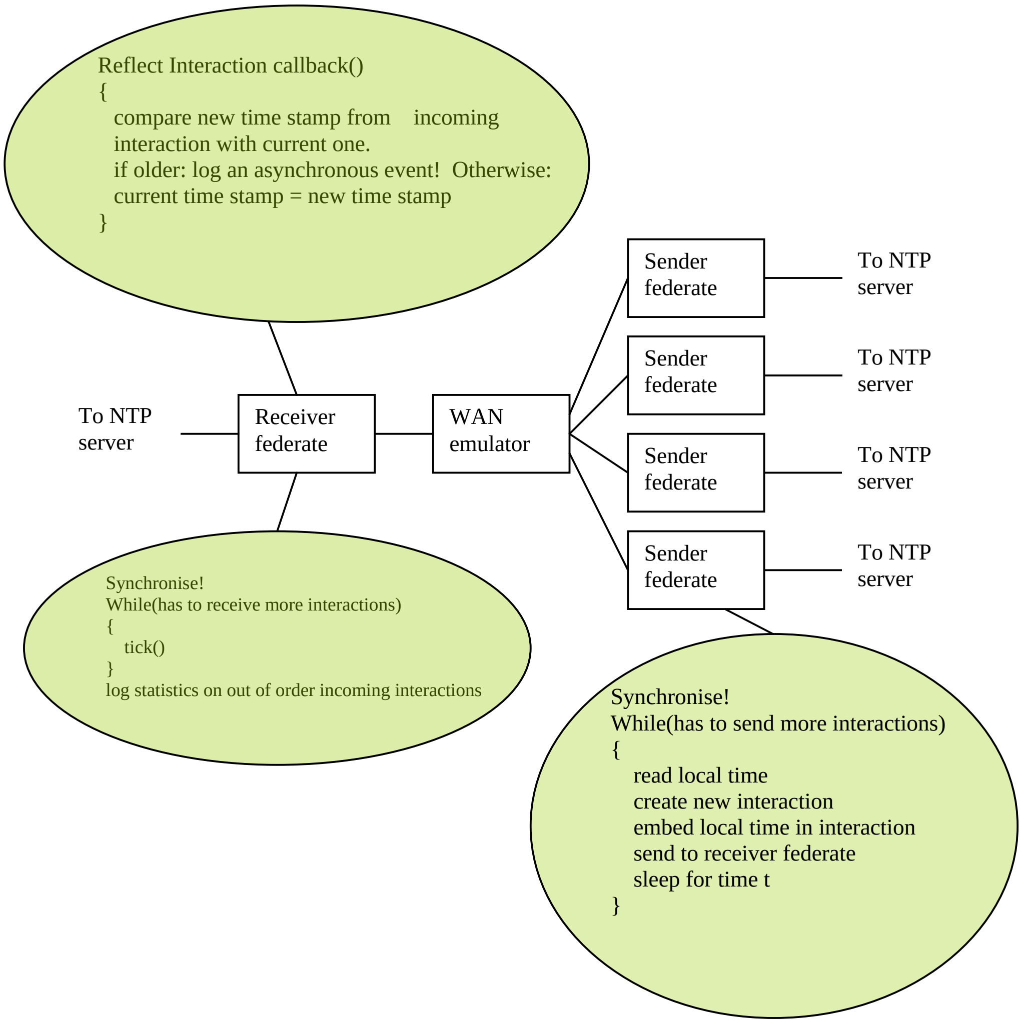

Figure 2-1 A Non-Time-Managed Federation

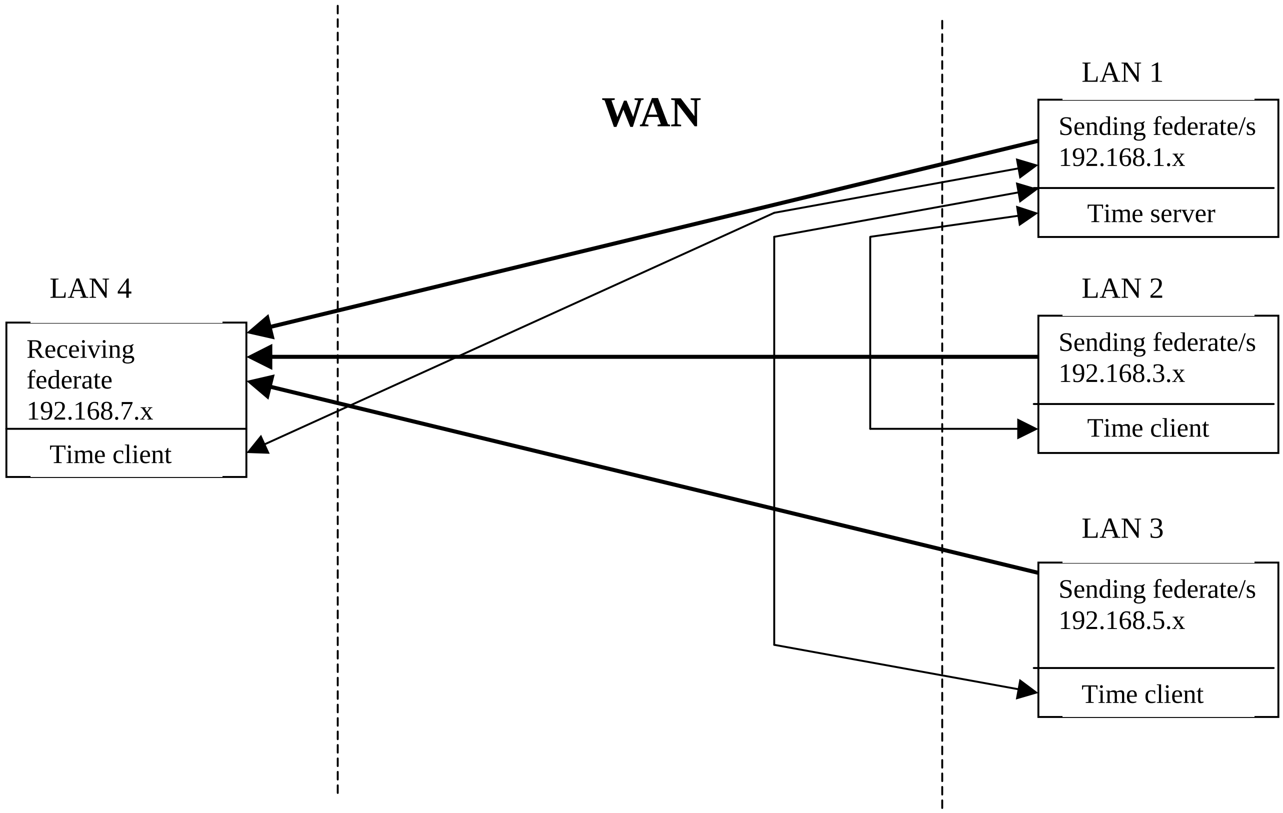

Figure 2-2 Experimental Network Setup

2.1.1 WAN setup

Setup #1:

Mean network delays were set at 100 ms and were normally distributed. Although network delays typically vary between different nodes, this parameter was fixed here to determine the baseline probability of obtaining asynchronous event deliveries. We had chosen a standard deviation value of 20 ms, which is a reasonable estimate of the variability of network delays.

Setup #2:

The mean network delay between LAN1 and LAN4 was set at 50 ms with standard deviation of 10 ms.

The mean network delay between LAN2 and LAN4 was set at 100 ms with standard deviation of 20 ms.

The mean network delay between LAN3 and LAN4 was set at 200 ms with standard deviation of 40 ms.

These values represent the delays associated with the actual WAN used by DRDC Atlantic and partners in the UK (50 ms), Australia (100 ms), and N.Z (200 ms). Again the values chosen for standard deviation represent an estimate of the variability of network delays.

2.1.2 Test Runs

The following test runs were performed:

Three (3) sending federates, one (1) receiver federate, network delays with a mean of 100 ms between all senders (WAN setup #1) and receiver federate. Outgoing message frequencies were: 100 Hz, 50 Hz, 20 Hz and 10 Hz (total of 4 runs).

Six (6) sending federates, one (1) receiver federate, network delays with a mean of 100ms between all senders (WAN setup #1) and receiver federate. Outgoing message frequencies were: 50Hz, 20Hz, 10Hz, 5Hz, 2Hz, 1Hz and 0.5Hz (total of 7 runs).

Six (6) sending federates, one (1) receiver federate, network delays of 50ms, 100ms and 200ms between sending and receiving federates (WAN setup #2). Outgoing message frequencies were: 50Hz, 20Hz, 10Hz, 5Hz, 2Hz, 1Hz and 0.5Hz (total of 7 runs).

Each run included 10 cycles where after each cycle all federates resign and the federation is destroyed. The cycles were invoked using a shell script. During each cycle the following steps were performed:

- Join the federation.
- Publish or subscribe to a single interaction class.
- Synchronise at synchronization point #1.
- Each sending federate sends 500 interactions in a loop at the given message frequency.
- The federate calls sleep() between sends where the sleep period is defined by the message frequency but has a random portion to ensure that messages will not be sent with a constant time gap, this is done in order to avoid a situation where messages arrive at a constant timing alignment at the receiving federate.

- After all 500 interactions have been sent the federation synchronises on synchronization point #2, the receiving federate finalizes the logging and all federates resign.

- Federation is destroyed.

2.1.3 Results

Graph 1 presents the results from test runs with 4 federates (1 receiver and 3 senders) where all federates (LANs) were interconnected with an identical network delay of 100ms.

When using message frequencies of 10Hz-100Hz 10% or less of the messages arrived out of time stamp order while less than 2% had their time stamp in the past by more than 40ms. An increase in message frequency causes more messages to arrive out of time stamp order and by a larger delay. An attempt to increase message frequency to 200 Hz resulted in a dramatic increase in both the probability of asynchronous events (almost 100%) and amount of “time into the past” that the message represents (around 1000 ms). The network was set to introduce network delays in the range of 73ms-127ms (the far edges of the normal distribution curve were not represented) so the network itself can only be responsible to out of order delays of up to 54ms. Higher out of order delays are due to RTI and IP layer characteristics such as internal message queuing by IP layer buffers and RTI message queues, interrupt disable period by network hardware irq (interrupt request) and implementation of RTI:Tick(). Those software components are especially sensitive to traffic load and will be the dominant reason for the out of order delay at higher message frequencies.

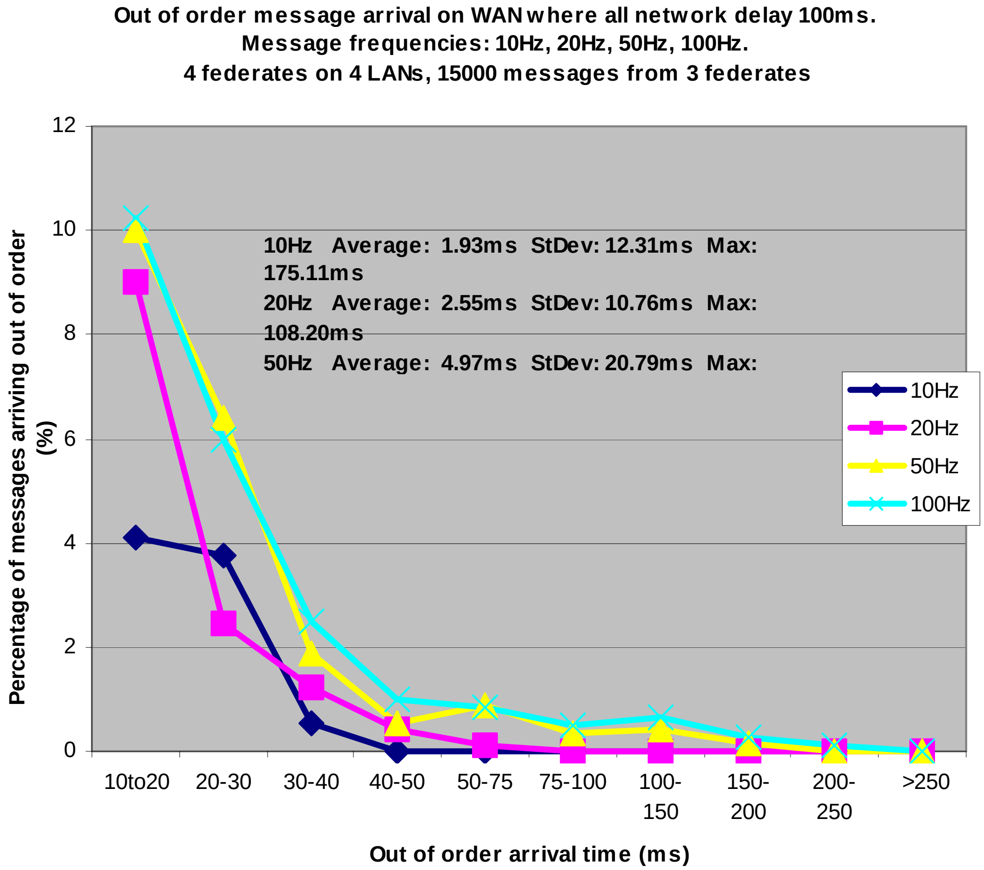

Figure 2-3 Message Arrival on WAN: Graph 1

Graphs 2 and 3 present the results from test runs with 7 federates (1 receiver and 6 senders) where all federates (LANs) are interconnected with an identical network delay of 100ms. Two sending federates were residing on each LAN that was simulated by a single machine. Each federate was running in its own shell as a process. Default scheduling policy was used resulting in an equal time slice of 10 milliseconds for each federate unless the federate suspended itself by invoking a system call (as had been used during the simulation when federates were calling sleep() between calls to sendInteraction()). Since the process suspension period was greater than the operating systems time slice, and the operation of sending an interaction (including all RTI and IP layer overhead) requires much less than 10 milliseconds (less than 2 milliseconds as benchmark shows) the processor utilization generated by each sending federates was lower than 10% (as indicated by the operating systems performance monitor). It is safe to assume that under those conditions (low processor utilization) execution of two federates on a single machine produces the same behaviour as when each federates is being executed on its own machine.

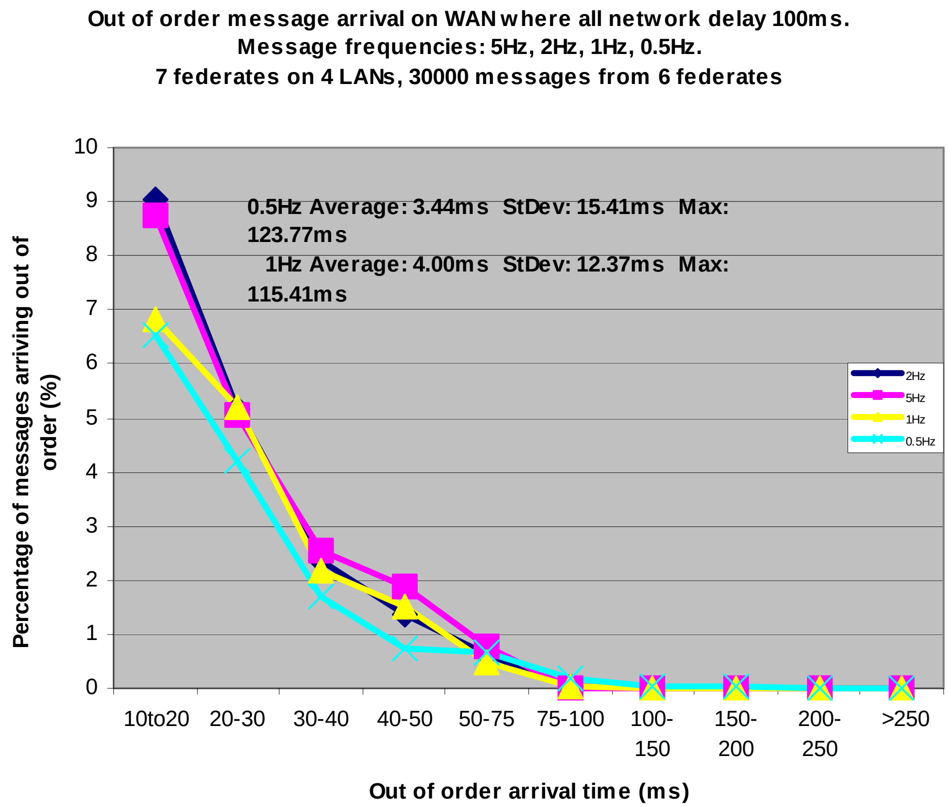

Figure 2-4 Message Arrival on WAN: Graph 2

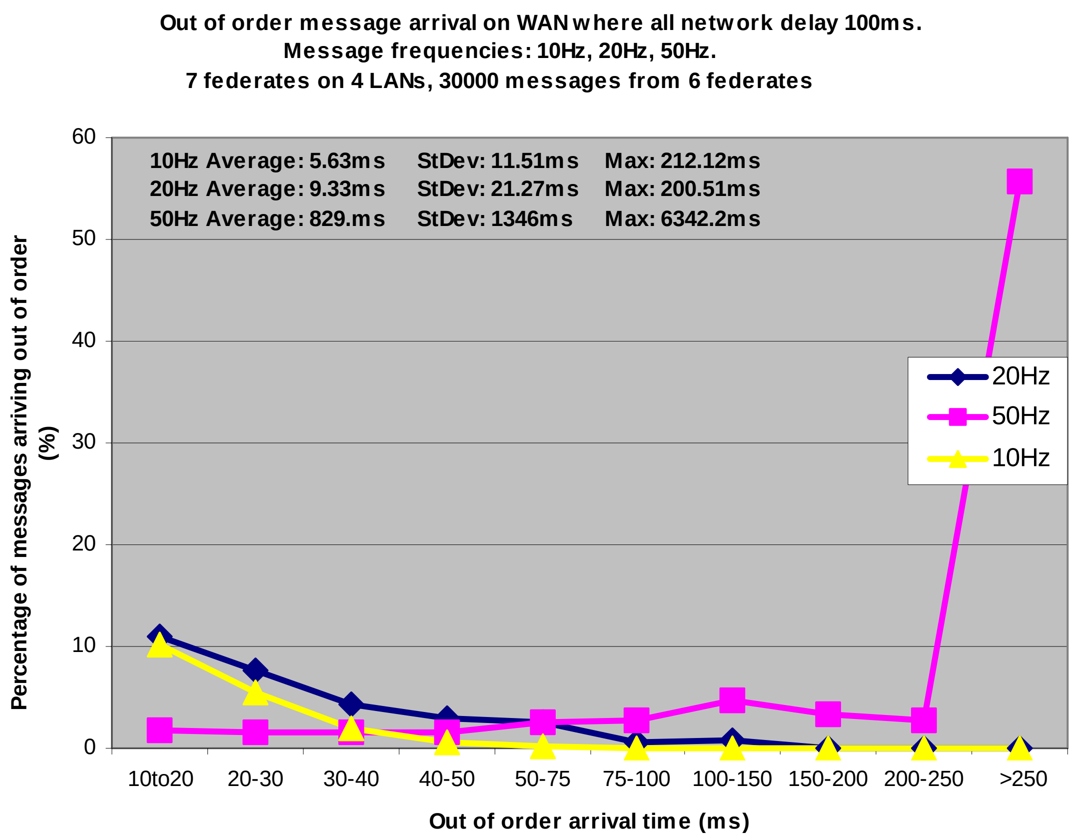

Figure 2-5 Message Arrival on WAN: Graph 3

The huge jump in out of order delay (most messages arrived with out of time stamp order delay around 1000ms) when message frequency was 50Hz (Graph – 3) is due to queuing delays when the RTI on the receiving node was unable to keep with the flow of incoming HLA traffic. An identical behaviour was observed when using 4 federates and message frequency slightly greater than 100Hz. For lower message frequencies (Graph 2) less than 2% of messages had their time stamp at the past (relatively to the latest received time stamp) by more than 40ms while at message frequency of 20Hz 3%-4% of messages arrived with an out of order time stamp of up to 75ms.

Graphs 4 through 8 present the probability of asynchronous message arrival at a federation that is executing over a typical WAN with network delays of 50 ms, 100 ms and 200 ms between federates. Each graph presents results for a single message frequency:

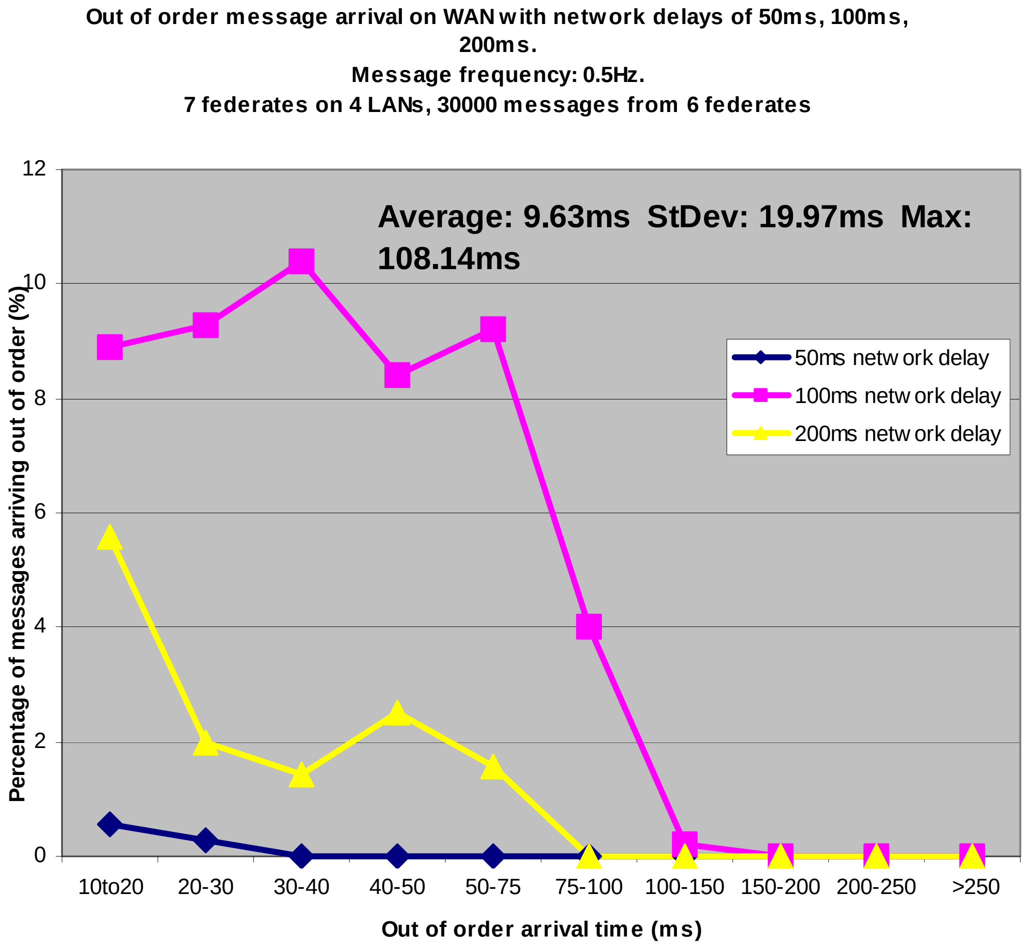

Figure 2-6 Message Arrival on WAN: Graph 4

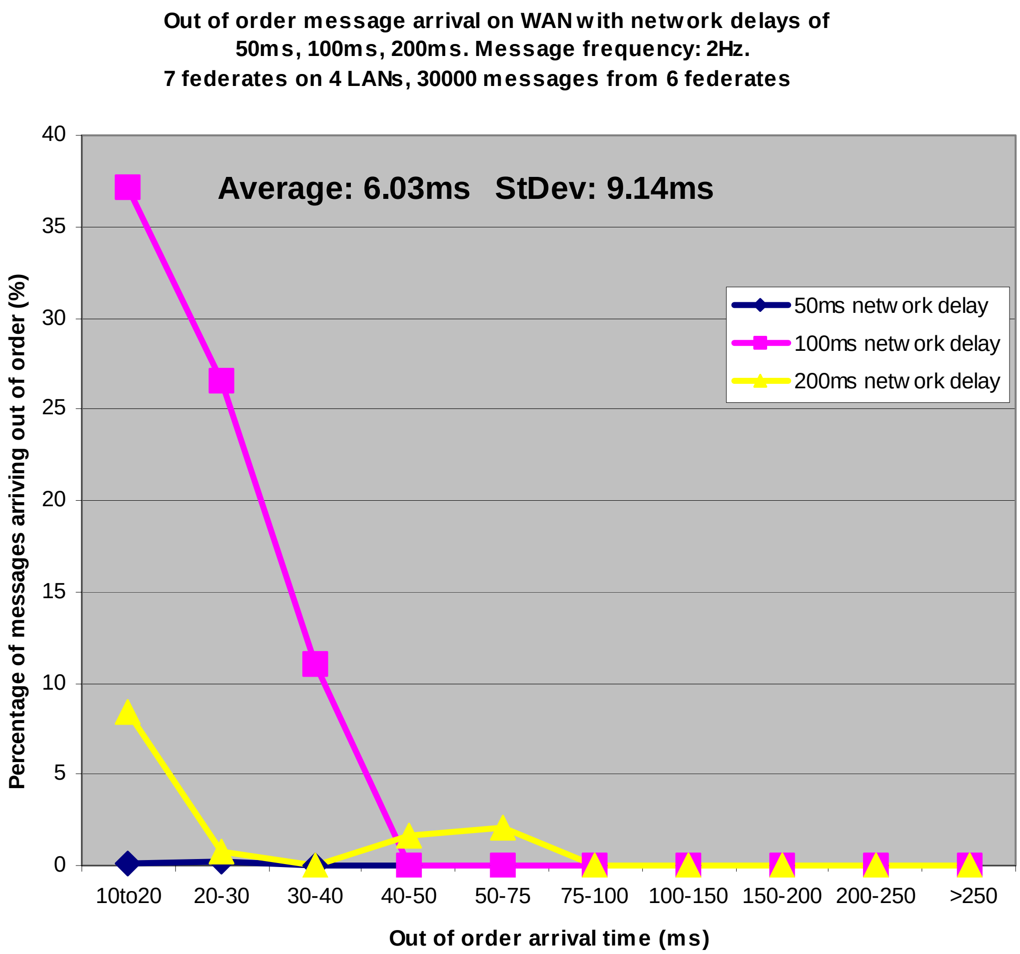

Figure 2-7 Message Arrival on WAN: Graph 5

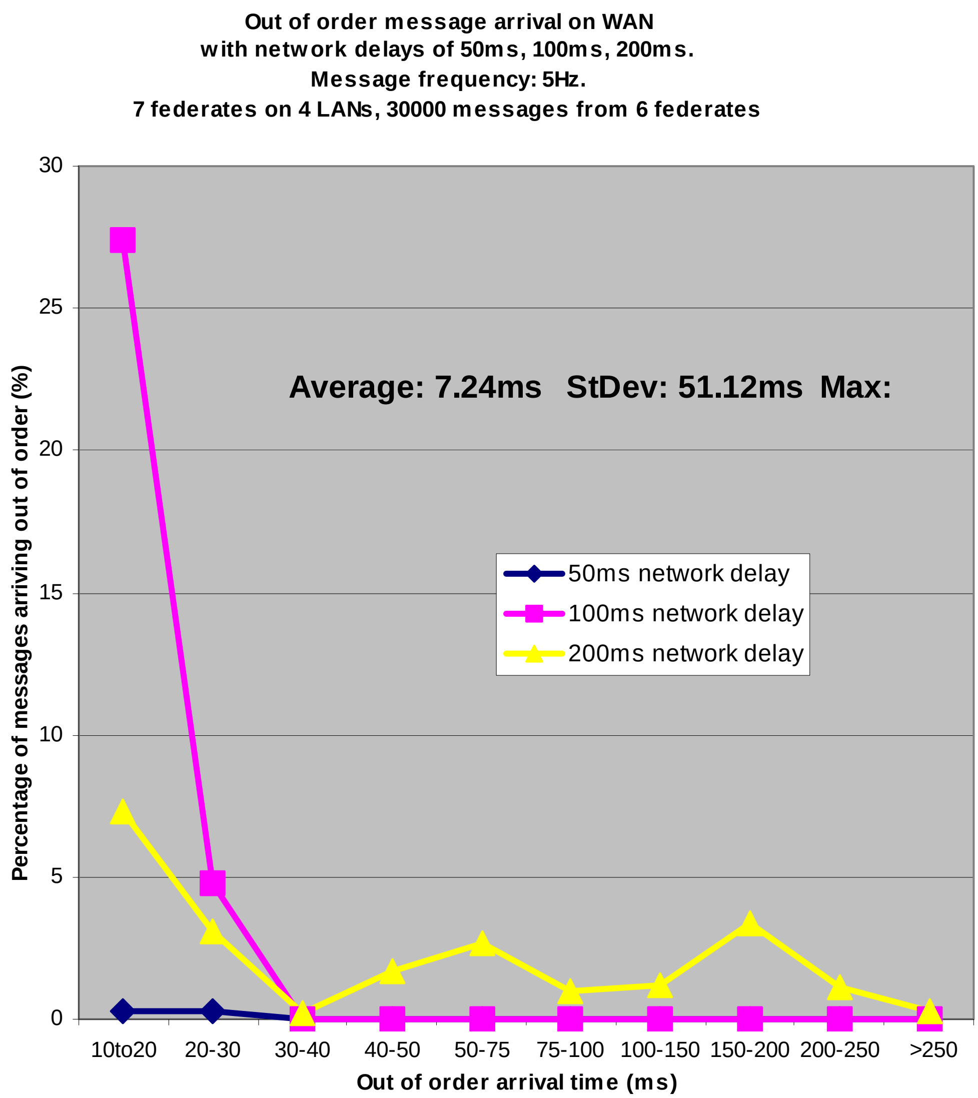

Figure 2-8 Message Arrival on WAN: Graph 6

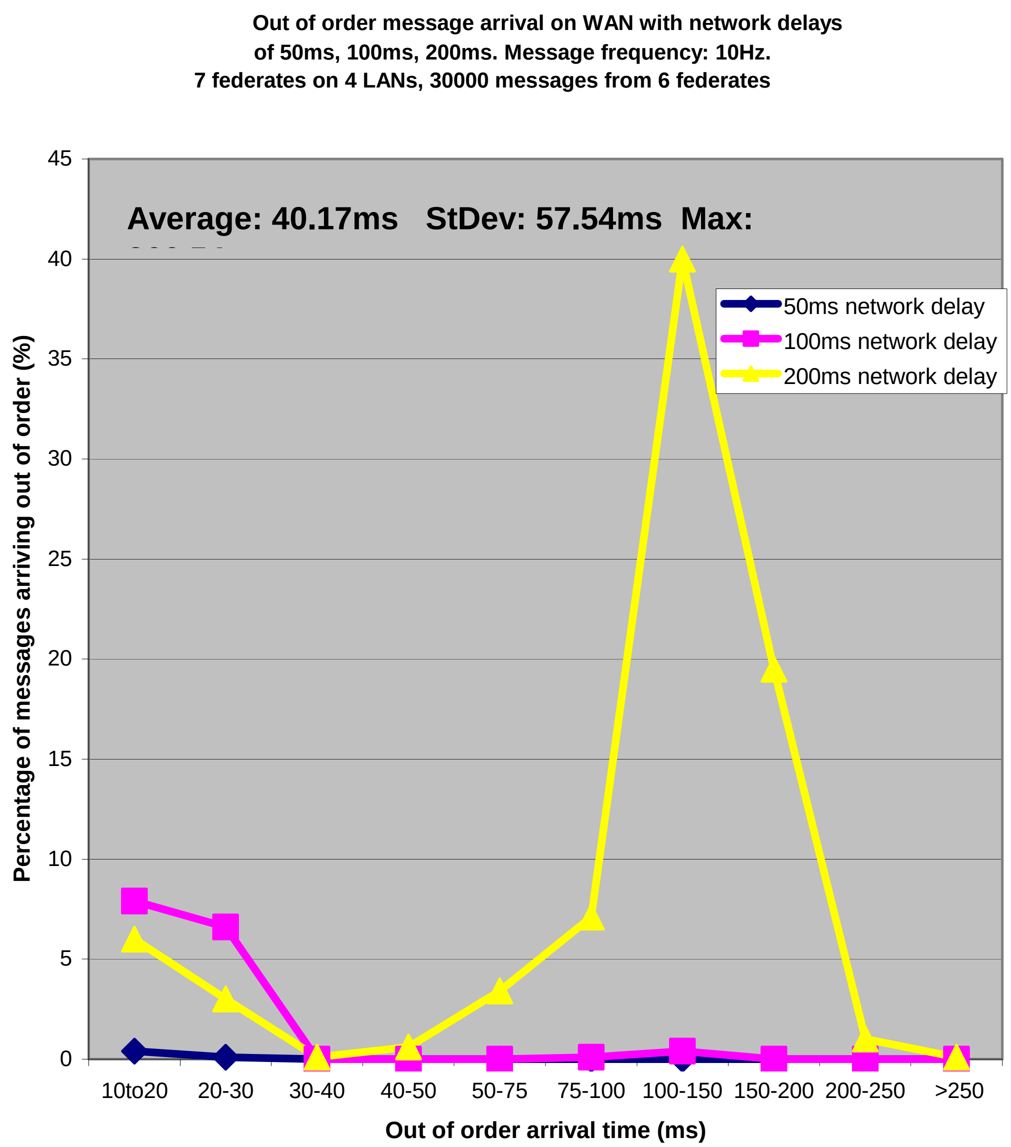

Figure 2-9 Message Arrival on WAN: Graph 7

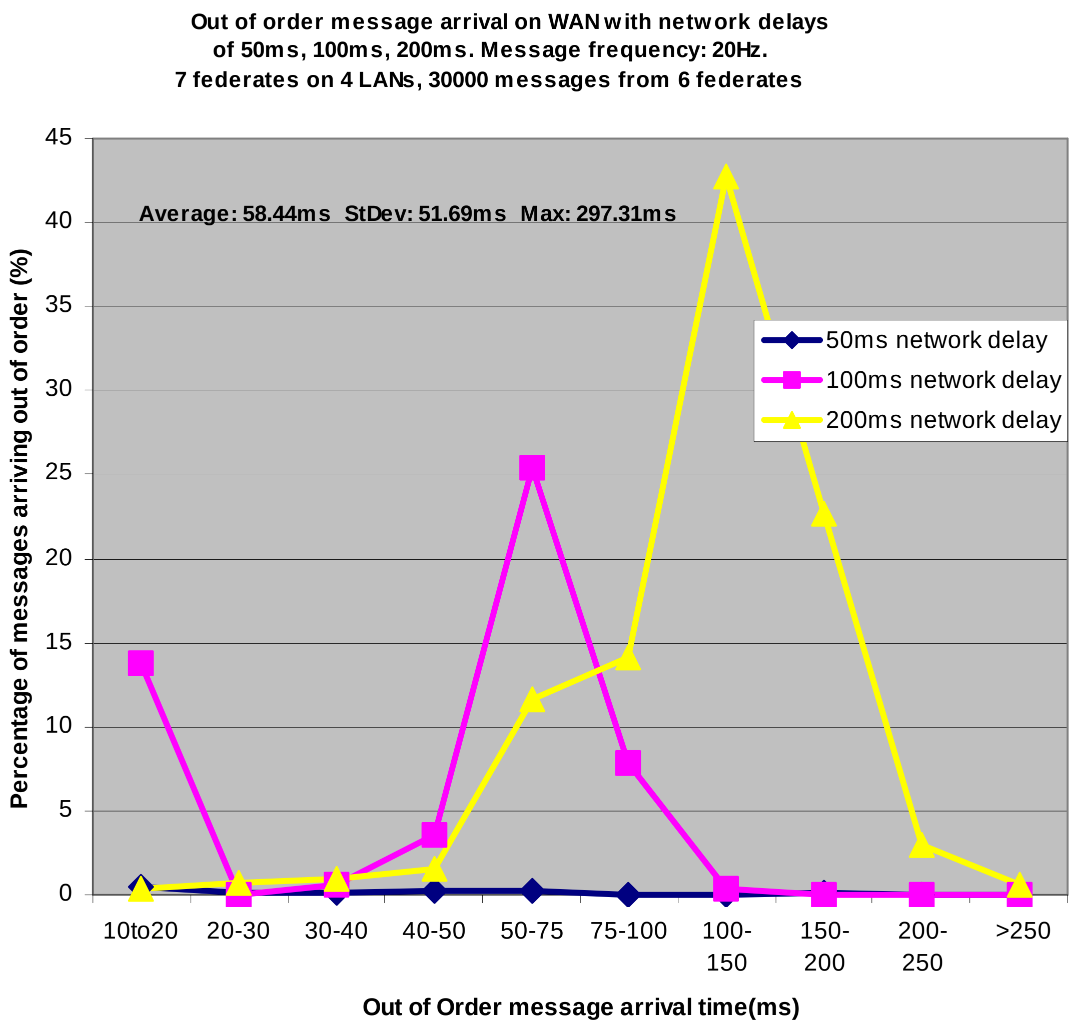

Figure 2-10 Message Arrival on WAN: Graph 8

Since the minimal network delay for a federate on a 50ms delay segment of the WAN can be as low as 40ms while the maximum network delay for a federate on a 200ms delay segment of the WAN can go as high as 252ms, the WAN itself can cause out of order message arrival times as long as 212ms. On average depending on data rates, the out of order message arrival times were between 7ms for very low data rates and almost 60ms for high data rates. The percentage of messages arriving out of time order sequence increases with the increase in data rate. This can be explained by the fact that as soon as time gap between consecutive messages originating at the same federate is not much grater than the difference in network delays, the probability that timing of message sending will compensate for the difference in network delays, decreases. This is demonstrated in the following example: assume a federation with federates A, B and C. A and B are sending messages to C where the network delay from A to C is fixed at 50 ms while the network delay from B to C is fixed at 200 ms. In order for the messages to arrive at C at an out of order sequence, A and B have to send their message during the same 150ms time period and B has to send prior to A. The probability for out of order message arrival is:

$$P_{\text{outOfOrderArrival}} = P_{\text{B sendsPriorTo A}} * P_{\text{A and B sendDuringSame150msPeriod}}$$

Where $P_{\text{A and B sendDuringSame150msPeriod}}$ = 150ms / time gap between msgs

And $P_{\text{B sendsPriorTo A}} = 0.5$

Figure 2-11 illustrates the condition for out of order arrival at a message frequency of 1Hz

Figure 2-11 Probability at 1 Hz

Figure 2-12 illustrates the condition for out of order arrival at a message frequency of 5Hz

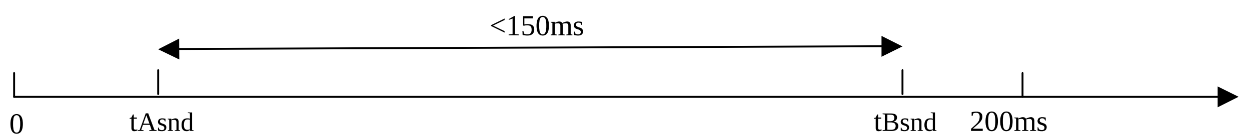

Figure 2-12 Probability at 5 Hz

Hence, the probability for an out of order message arrival when message frequency is 1Hz is: 0.5*150ms/1000ms=7.5% while at a message frequency of 5Hz the probability is: 0.5*150ms/200ms=32.5%.

Since the average out of order arrival time is calculated over all the messages that were received at the receiving federate, a higher probability of out of order message arrival also increases the average value of the out of order arrival time; this is better expressed in Graph 9. As expected, messages originated at the sending federates on LANs closer to the receiving federate (in terms of network delay) were less likely to arrive out of order. Messages from the 50ms network delay LAN still arrived occasionally out of their original send order because there were two individual senders on this LAN and the network delay was normally distributed around 50ms delay.

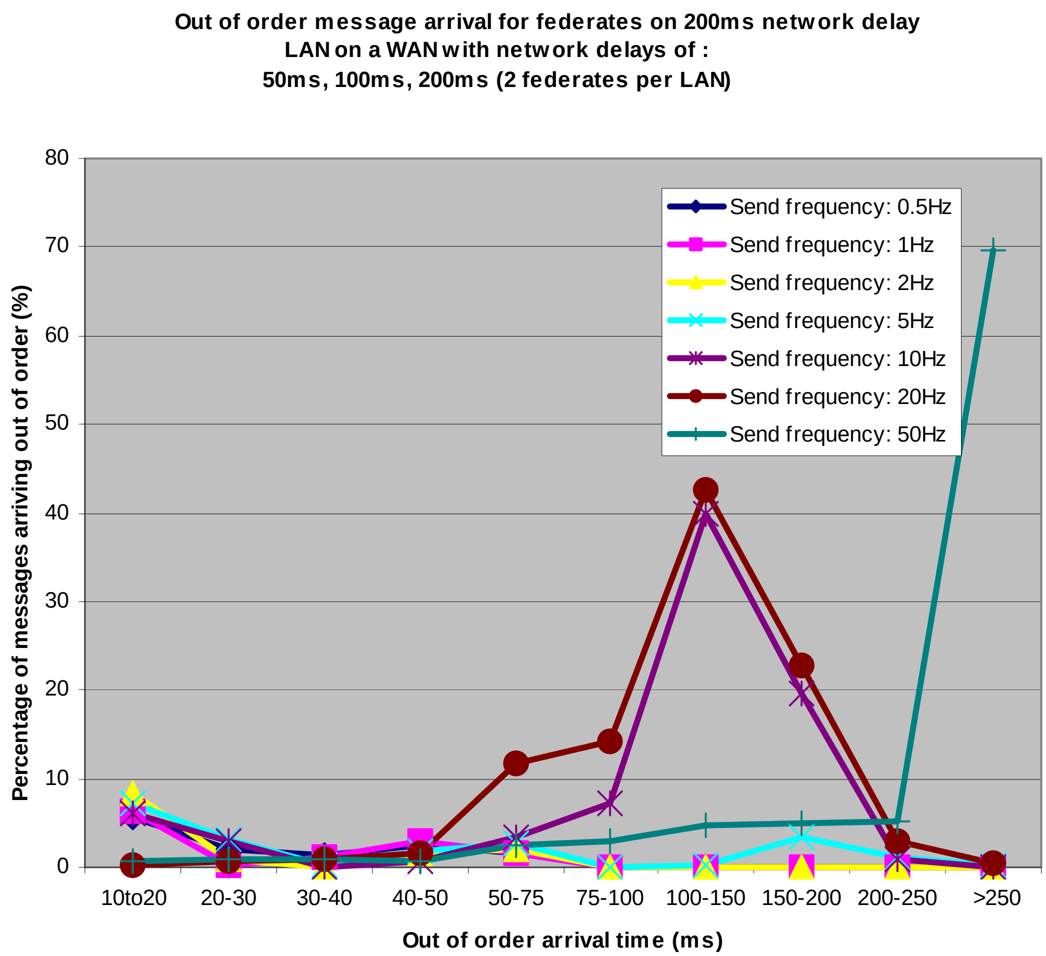

Figure 2-13 Message Arrival on WAN: Graph 9

The results for the 2 federates on the 200 ms network delay segment when message frequency was 50 Hz, suggests again that at this data rate the receiving federate failed to keep with the flow of incoming data and messages started to queue for as long as

1000 ms. Average out of order arrival time in this case was 709.3ms with a maximum value of 6000ms (6 seconds).

2.2 Message rate in a time managed federation

The time-managed federation in 2-14 can provide the worst-case rate of which the federation will advance in time. This rate provides a good indication about the amount of new events that can be generated by the federation during a given period of wall clock time.

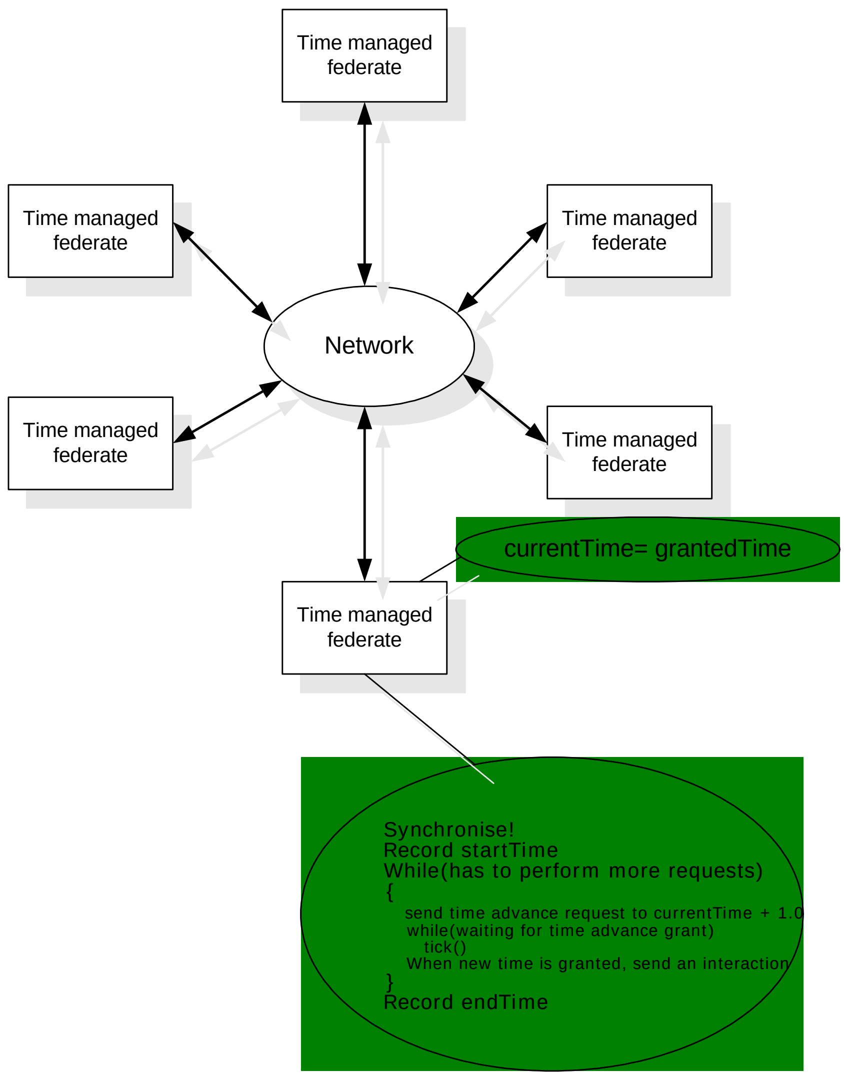

Figure 2-14 Time-Managed Federation

The experimental setup shown in Figure 2–15 was used with 2, 4, 6 and 8 federates over network delays of 50ms, 100ms and 200ms between LANs. Five runs were performed for each network delay (network delays were set for a fixed delay without a distribution function) and a given number of federates, with 500 time advance requests by each federate per run. After each run all federates resigned the federation and the federation was destroyed. A look ahead value of 1.0 (identical to the federations time step) had been used.

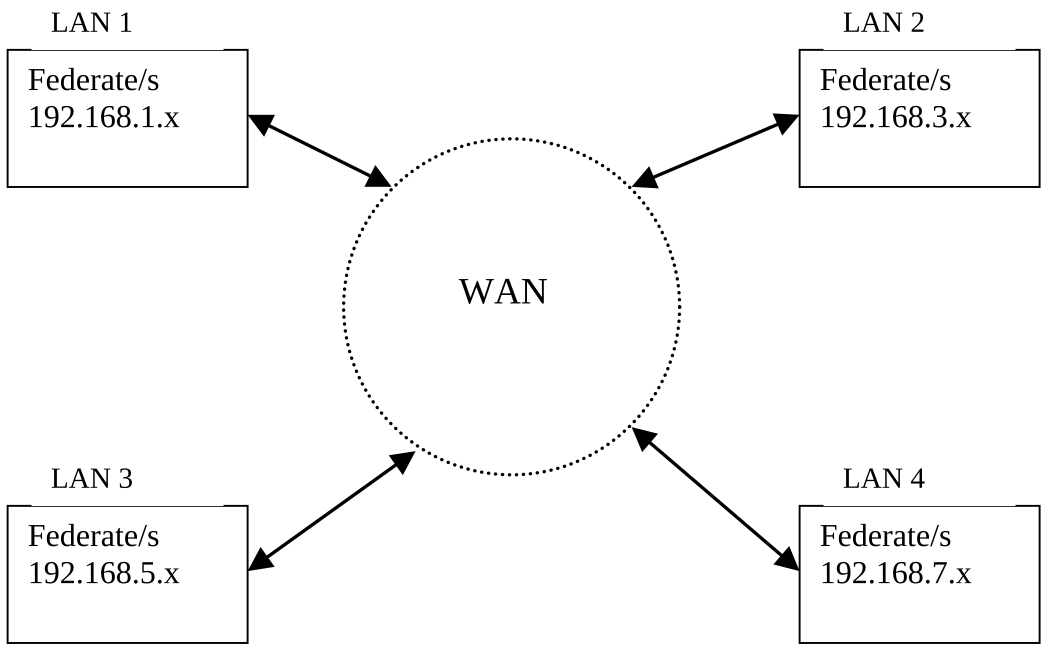

Figure 2-15 Experimental Network Setup

2.2.1 Results

Results for the various conditions listed above are shown in Figure 2-16. For larger size federations (6-8 federates) where all federates are time regulating and time constrained and the network delay is 100ms or longer, less than two time advance requests can be processed each second. Earlier experience with a different RTI (from Georgia Tech version 4.0) provided a 50% improvement over the DMSO NG 4 RTI (DMSO NG 6 was not reported to have a performance advantage over DMSO NG 4 in this respect) when used over LAN so some improvement can be expected here as well. Figure 2-16 shows that the LBTS computation time is proportional to $\log_2 N$ where N is the number of federates. This is consistent with (Fujimoto & Hoare, 1998) who state that LBTS calculation requires $\log_2 N$ serialized steps.

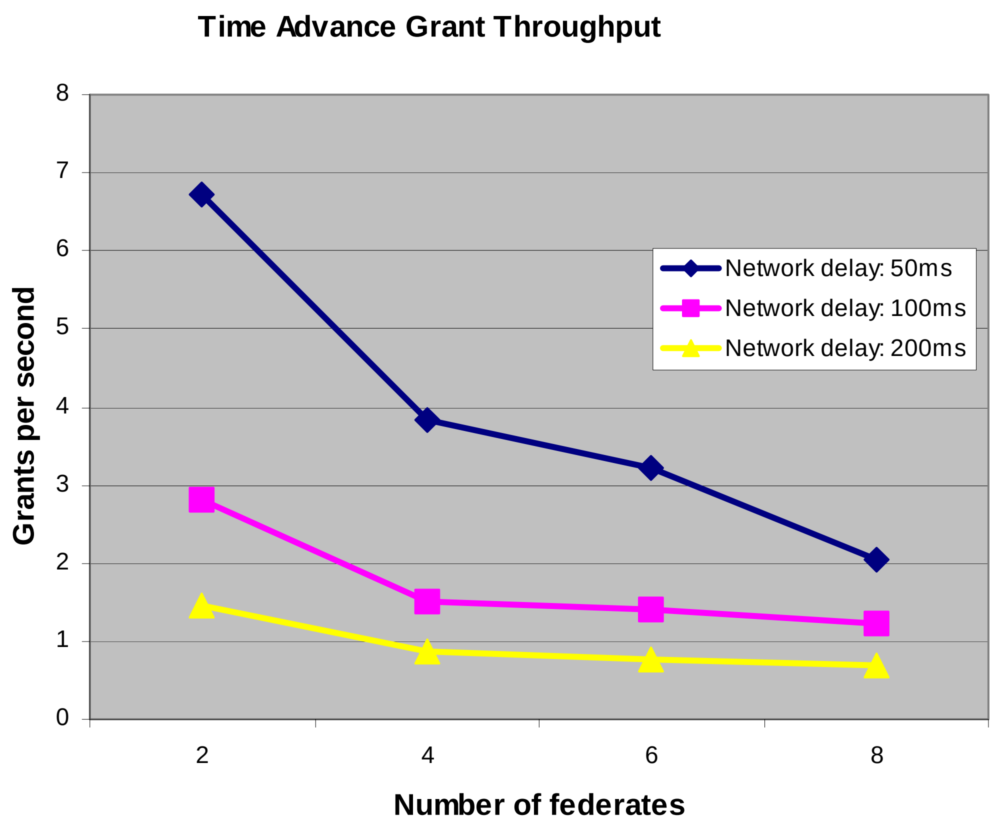

Figure 2-16 Time Advance Grant Throughput

3. WAN Emulator

The WAN emulator shown in Figure 3 - 1 consists of 4 LANs interconnected in a "star" deployment to resemble a WAN. Each LAN represents a participant country as shown in Figure 3 - 2. In the "real world" each one of the four participating countries consists of about 4 or 5 nodes (machines) on a single LAN. The four participating countries are deployed in a "star" type WAN with the stars center in Virginia USA. During experiments each LAN (country) was represented by a single node (a Linux RedHat 9.0 PC).

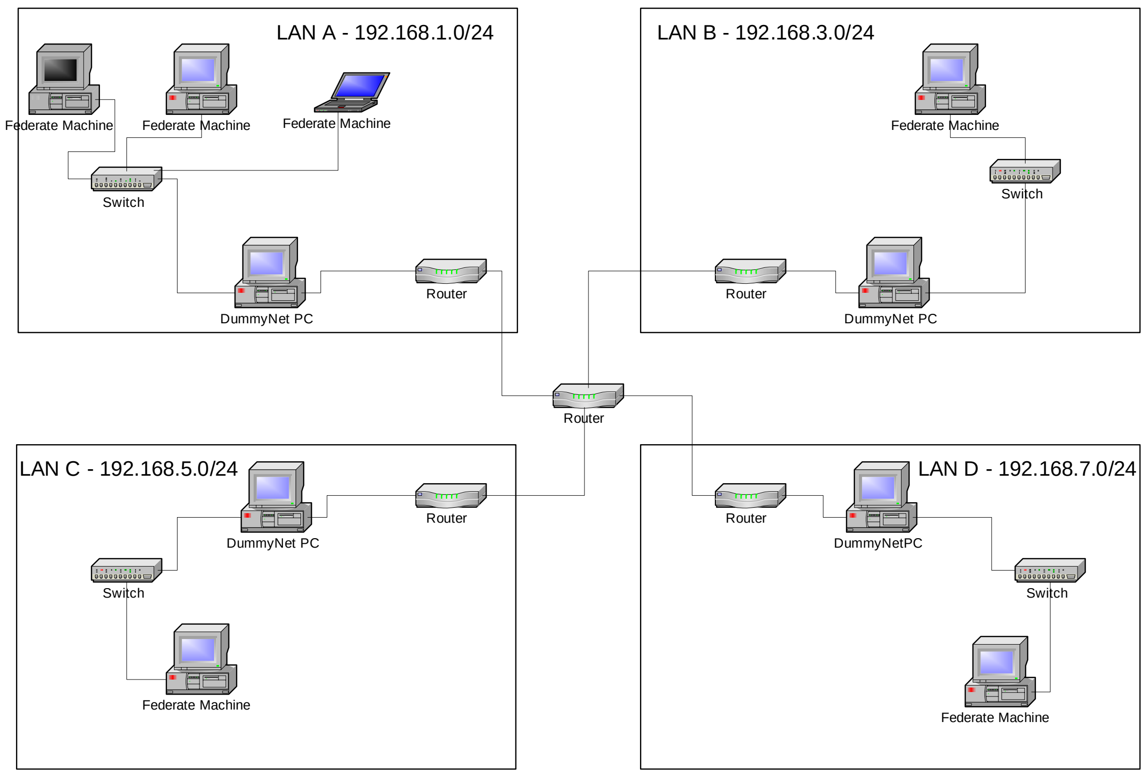

Figure 3-1 WAN Emulator

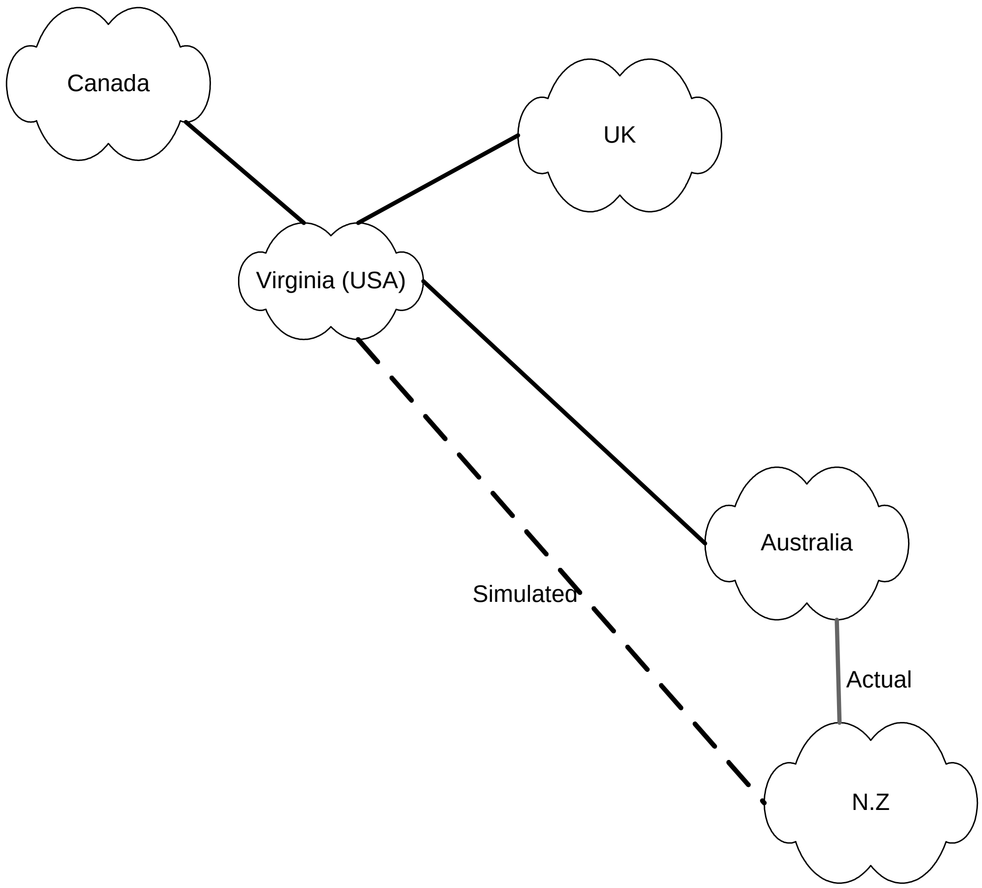

Figure 3-2 Actual federation deployment

The core of the emulator is based on the DummyNet machines running BSD 5.3. Each DummyNet machine is equipped with two NICs (Network Interface Cards). Each packet that is passed through the DummyNet machine enters it through one of the NICs end exits through the other NIC. DummyNet is capable to control network traffic between the two NICs by controlling parameters such as network capacity between the NICs or introduction of delays for each packet that is passed between the NICs. Furthermore, the delay can vary based on a user defined probability function.

In addition it is possible to set network delay based on source and destination IP so that a specific DummyNet machine can apply specific delay to each packet according to its source and destination addresses.

DummyNet also provides functionality to control the amount of lost (dropped) IP packages.

It is possible to shape UDP and TCP traffic characteristics through the emulated WAN independently.

By using the WAN emulator as described above it is possible and easy to accurately set the network delay and the delay characteristics between any two LANs (countries) as well as other critical parameters such as bandwidth between any given two countries, and reliability of packet delivery.

The WAN emulator’s components are responsible for about a 3ms delay end to end.

4. Conclusions

As expected, in a non-time managed federation executing over WAN with realistic network delays, the probability of asynchronous events occurring is significant and depends on both network characteristics and message rates.

The DMSO RTI (version NG-6) is unable to handle efficiently higher rates of incoming messages (200 messages per second or more). This results in message queuing and dramatically increases both the probability of asynchronous events and amount of "time into the past" that the message represents.

On the other hand, with conservative message rates (5 Hz or less) and careful filtering of out of sync messages at the federate ambassador code, and in addition by allowing messages that are only 20 ms – 40 ms out of order, over 80% of the received messages can still be used.

Even when network delays across the federation are uniform, as long as there is variability in network delays there is a significant probability that messages will be received out of their original send order.

The non-time managed federation experiment represents the worse case condition for a federation of the discussed sizes (4 or 7 federates) since all the sending federates were sending concurrently to a single receiver. Depending on actual simulation constraints and characteristics, results may be somewhat more favorable (smaller percentage of event notification arrive out of their send order).

It is possible to achieve a significant improvement by making network delays across the federation more uniform. This can be done by introduction of additional artificial network delay to the faster segments of the WAN.

A completely time managed federation across a slow WAN is impractical in most cases.

The WAN emulator at DRDC Atlantic is a valuable research tool for investigation of HLA federations over WAN.

The WAN emulator is flexible robust and easy to configure for emulation of a wide range of WAN behaviours.

DMSO RTI version NG-6 performed in a robust way. It didn't crash a single time during days of consecutive runs. Federates never failed to join or resign cleanly despite using a relatively large number of federates and long and uneven network delays.

5. Future Work

In order to prove the validity of the findings and concepts from previous sections it would be helpful to construct a practical simulation using a non-time managed federation over the emulated WAN at DRDC Atlantic. This simulation should include/implement the following:

1. A GPS receiver to provide accurate time to all nodes (a setup that includes all hardware and software is less than $1K).

2. A conservative message rate (2 Hz-5 Hz).

3. Implementation of incoming message queuing and time-stamp based message reordering at the receiving federate ambassador code.

4. And/or filtering out messages that arrived with a time stamp that is more that TBD seconds (around 20 milliseconds) older than the latest received time stamp.

Investigation into more efficient time management algorithms that can be implemented and tested on an open source RTI.

6. References

1. Hoare, P. & Fujimoto, R.(1998). *HLA RTI performance in high speed LAN environments*. In Proceedings of the 1998 Fall Simulation Interoperability Workshop. (Orlando, FL, September). 501-510.

List of symbols/abbreviations/acronyms/initialisms

DND	Department of National Defence
DMSO	Defense Modeling & Simulation Office
DRDC	Defence Research & Development Canada
GPS	Global Positioning System
HLA	High Level Architecture
IP	Intellectual Property / Internet Protocol
Irq	Interrupt request
LAN	Local Area Network
LBTS	Lower Bound Time Stamp
NIC	Network Interface Card
NTP	Network Time Protocol
PC	Personal Computer
RTI	Run-Time Infrastructure
TSO	Time Stamp Ordered
WAN	Wide Area Network

Distribution List

DRDC Atlantic CR 2005-182

Internal Distribution: (DRDC Atlantic Responsibility)

6 DRDC Atlantic Library File Copies

1 M.G. Hazen, Scientific Authority

1 B. Dillman

1 J. Murphy

1 D. Bleichman, Contractor

1 A. Gillis

External Distribution by DRDKIM

1 NDHQ/DRDKIM 3

2 Robert Elliot
Synthetic Environment Coordination Office (SECO)
Canadian Forces Experimentation Centre
3701 Carling Avenue
Ottawa, Ontario, K1A 0Z4

1 A. Vallerand, H/FFSE
Defence R&D Canada – Ottawa
3701 Carling Avenue
Ottawa, Ontario, K1A 0Z4

15 Total Copies required

DOCUMENT CONTROL DATA

(Security classification of title, body of abstract and indexing annotation must be entered when the overall document is classified)

1\. ORIGINATOR (the name and address of the organization preparing the document. Organizations for whom the document was prepared, e.g. Centre sponsoring a contractor's report, or tasking agency, are entered in section 8.)

Defence Research & Development Canada – Atlantic

2\. SECURITY CLASSIFICATION (overall security classification of the document including special warning terms if applicable).

UNCLASSIFIED

3\. TITLE (the complete document title as indicated on the title page. Its classification should be indicated by the appropriate abbreviation (S,C,R or U) in parentheses after the title).

Investigation of Asynchronous Events in Non-Time-Managed Federations (U)

4\. AUTHORS (Last name, first name, middle initial. If military, show rank, e.g. Doe, Maj. John E.)

Dan Bleichman

5\. DATE OF PUBLICATION (month and year of publication of document)

September 2005

6a. NO. OF PAGES (total containing information Include Annexes, Appendices, etc).

38 (approx.)

6b. NO. OF REFS (total cited in document)

1

7\. DESCRIPTIVE NOTES (the category of the document, e.g. technical report, technical note or memorandum. If appropriate, enter the type of report, e.g. interim, progress, summary, annual or final. Give the inclusive dates when a specific reporting period is covered).

CONTRACT REPORT

8\. SPONSORING ACTIVITY (the name of the department project office or laboratory sponsoring the research and development. Include address).

Defence R&D Canada – Atlantic
PO Box 1012
Dartmouth, NS, Canada B2Y 3Z7

9a. PROJECT OR GRANT NO. (if appropriate, the applicable research and development project or grant number under which the document was written. Please specify whether project or grant).
CF SECO / 11BK

9b. CONTRACT NO. (if appropriate, the applicable number under which the document was written).

W7707-042793/001/HAL

10a ORIGINATOR'S DOCUMENT NUMBER (the official document number by which the document is identified by the originating activity. This number must be unique to this document.)

DRDC Atlantic CR 2005-182

10b OTHER DOCUMENT NOs. (Any other numbers which may be assigned this document either by the originator or by the sponsor.)

11\. DOCUMENT AVAILABILITY (any limitations on further dissemination of the document, other than those imposed by security classification)

(X) Unlimited distribution
(   ) Defence departments and defence contractors; further distribution only as approved
(   ) Defence departments and Canadian defence contractors; further distribution only as approved
(   ) Government departments and agencies; further distribution only as approved
(   ) Defence departments; further distribution only as approved
(   ) Other (please specify):

12\. DOCUMENT ANNOUNCEMENT (any limitation to the bibliographic announcement of this document. This will normally correspond to the Document Availability (11). However, where further distribution (beyond the audience specified in (11) is possible, a wider announcement audience may be selected).

13. ABSTRACT (a brief and factual summary of the document. It may also appear elsewhere in the body of the document itself. It is highly desirable that the abstract of classified documents be unclassified. Each paragraph of the abstract shall begin with an indication of the security classification of the information in the paragraph (unless the document itself is unclassified) represented as (S), (C), (R), or (U). It is not necessary to include here abstracts in both official languages unless the text is bilingual).

A study was developed and conducted at DRDC Atlantic to examine the relationship between (1) Wide Area Network (WAN) characteristics (i.e., latency and latency variability) and (2) federation characteristics (e.g., frequency of the occurrence of events as indexed by the frequency of the exchange of messages, event handling time and federation size) to the probability of an asynchronous event occurrence in non-time-managed federations.

The study was conducted using federates sending or receiving wall-clock time-stamped HLA messages over an emulated WAN while all hardware clocks of the participating machines were synchronised. The time-stamp of each received message was examined and used in order to determine out-of-sync arrival.

Results show that in a non-time-managed federation executing over a WAN with realistic network delays, the probability of asynchronous events occurring is significant and depends on both network characteristics and message rates.

14. KEYWORDS, DESCRIPTORS or IDENTIFIERS (technically meaningful terms or short phrases that characterize a document and could be helpful in cataloguing the document. They should be selected so that no security classification is required. Identifiers, such as equipment model designation, trade name, military project code name, geographic location may also be included. If possible keywords should be selected from a published thesaurus. e.g. Thesaurus of Engineering and Scientific Terms (TEST) and that thesaurus-identified. If it not possible to select indexing terms which are Unclassified, the classification of each should be indicated as with the title).

HLA
Wide Area Network (WAN)
Latency
Federation
Distributed simulation

This page intentionally left blank.

Defence R&D Canada

Canada's leader in defence
and National Security
Science and Technology

R & D pour la défense Canada

Chef de file au Canada en matière
de science et de technologie pour
la défense et la sécurité nationale

DEFENCE DÉFENSE

www.drdc-rddc.gc.ca