Industry
Canada
CRC

A STUDY OF ISO 10589 PROTOCOL IN A DYNAMIC QoS-BASED ROUTING ENVIRONMENT

by

Claude Bilodeau

CRC REPORT NO. 96-001

April 1996
Ottawa

Industry Canada Industrie Canada

The work was developed under the NATO CSNI project which was sponsored in Canada by the Department of National Defence.

Canada

A Study of ISO 10589 Protocol in a Dynamic QoS-Based Routing Environment

by

Claude Bilodeau
Network Technologies Research

The work was developed under the NATO CSNI project which was sponsored in Canada by the Department of National Defence.

COMMUNICATION RESEARCH CENTRE, INDUSTRY CANADA
CRC REPORT NO. 96-001

Canada

April 1996
Ottawa

TRADEMARKS

OPNET is a registered trademark of MIL3, Inc.
Retix is a registered trademark of Retix
SPARC-II is a registered trademark of Sun Microsystems, Inc.

Table of Contents

PART-I: Stability and Promptness

PART-II: Efficiency

PART-III: Wrap Up

FOREWORD

The work described in this report was performed under the collaborative Communications Systems Network Interoperability (CSNI) R&D project.

Very briefly, CSNI is a five-year project initiated at the end of 1991 and includes six participating nations (Canada, France, Germany, The Netherlands, The United Kingdom and The United States) with the Shape Technical Centre also as a participant.

The principal CSNI objective is to develop, test and demonstrate multiservice (voice, data, messages) communications across mixed media transmission networks (HF, VHF, UHF, SHF) employing open systems principles and Commercial Off-The-Shelf (COTS) products to the greatest extent possible. R&D results from the project are made available to the international standards community for consideration in the promulgation of emerging standards.

The CSNI project organization, R&D schedule, demonstration testbed and overall major project accomplishments are summarized in [3].

A Study of ISO 10589 Protocol in a Dynamic QoS-Based Routing Environment

by

Claude Bilodeau

ABSTRACT

A computer simulation model of the CSNI network topology was developed to assess the effectiveness of QoS-driven routing in heterogeneous radio subnetworks. The simulated topology contains 11 routers, 33 subnetwork access controllers and as many end systems forming a single IS-IS routing domain of 13 subnets. The study investigates how to set the ISO 10589 protocol timers and QoS measurement period to best control the rate of link state advertisements within the network and to minimize the routing data overhead over the links. Simulation results show that the ISO 10589 : 1992 (E) protocol imposes several limitations when the routing information bases are dynamically modified to adapt to QoS changes of the subnets.

RÉSUMÉ

Un modèle du réseau de CSNI a été développé pour évaluer par simulation, et lorsqu' assujetti à la qualité de service disponible, l'efficacité de routage. La topologie du réseau inclus 11 routeurs, 33 controlleurs d'accès aux sous-réseaux, et autant de systèmes d'extrémité, le tout formant un domaine unique de routage de systèmes intermédiaires à systèmes intermédiaires de 13 sous-réseaux radio hétérogènes. Afin de bien controller le taux de diffusion des paquets de routage sur l'ensemble du réseau et de minimiser la charge de chacun des liens, l'étude détermine les valeurs de réglage optimum associées aux divers compteurs du protocol ISO 10589 et à la fréquence d'échantillonnage mesurant la qualité de service des sous-réseaux. Les résultats de simulation indiquent que le protocol ISO 10589 : 1992 (E) impose plusieurs contraintes dûes aux changements de l'information de routage lors de la variation de la qualité de service des sous-réseaux.

EXECUTIVE SUMMARY

The CSNI project adopted open systems interconnect (OSI) principles in developing an open architecture that could accommodate diverse users and a broad array of deployed systems and subnetworks. Central to this approach, the ISO 10589 intermediate system to intermediate system (IS-IS) protocol provides the intra-domain routing information exchange between the CSNI nodes. This report summarizes the results of a simulation study on the performance of the ISO 10589 standard. The simulation is tailor-made for the CSNI demonstrator but results are directly applicable to other non-homogeneous networks where multiple low bandwidth subnets and/or dynamic routing based on the users' requested Quality of Service (QoS) are the prime system characteristics.

The study identifies several limitations imposed by the ISO 10589 protocol, in particular with regard to the stability of routing information bases (RIBs). For instance, it determines the amount of time the network requires to adapt to each QoS change initiated by the subnetwork access controllers (SNACs) and reach the state where all RIBs are synchronized. For a network of the size of the CSNI demonstrator, it is possible to maintain the network routing stability over a loading range of 0 to 100% and a range of QoS metric generation rates of 3 to 15 minutes. For much larger networks, it is likely that longer and more frequent transient periods may put these networks in a constant state of sub-optimal routing. The report also estimates the amount of routing data overhead that can be expected over the CSNI links. It investigates how to set the ISO 10589 protocol timers to best control the rate of link state advertisements within the various subnetworks and concurrently minimize the routing data overhead over the links.

Overall, this simulation study shows that the April 1992 edition of the ISO 10589 protocol is not well suited to a dynamic routing and low-bandwidth broadcast subnetwork environment. Dynamic routing as conceived by CSNI or otherwise, was not within the design scope of the ISO 10589 protocol. The protocol was designed for static or quasi-static (QoS) routing. When used over broadcast subnetworks, the protocol is intended to operate on ISO 8802 LANs or other similar high bandwidth subnets and is not well suited to the low bandwidth broadcast subnets used in CSNI. In several aspects, the protocol is being asked to perform outside the functional environment that it is designed for or that it can currently support. It is doubtful that the QoS-driven routing concept could successfully be applied to heterogeneous radio networks exceeding several times the size of the CSNI demonstrator without making changes to the protocol and to the QoS-driven routing concept itself.

Several measures for substantially reducing the routing stability and overhead problems observed in this study are proposed herein.

A Study of ISO 10589 Protocol in a Dynamic QoS-Based Routing Environment

1.0 Introduction

This document reports on a simulation study of the ISO 10589 Intermediate System to Intermediate System (IS-IS) intra-domain routing protocol [1]. Despite being tailor-made for the CSNI network, the study provides results that are demonstrative of the protocol's behavior when used over multiple low bandwidth subnets and in a dynamic QoS-based routing environment. The main issues covered in this study include:

(a) *LSP Timers, IIH Timers, etc.*
It investigates how to set the IS-IS (ISO 10589) routing protocol timers to best control the rate of link state advertisements within the CSNI network;

(b) *CSNI Network stability, routing function adaptation speed, etc.*
It determines the amount of time the network requires to adapt to each QoS change and reach the state where all routing information bases (RIBs) are synchronised;

(c) *CSNI Network performance*
It estimates the amount of routing data overhead that can be expected over the CSNI links.

Sections 2.0 to 4.0 deal with some generalities about the simulation while the core of the report is divided into three parts.

Part-I (Sections 5.0 and 6.0) investigates problems (a) and (b) mentioned above. The main objective in this first part is to estimate the limitations imposed by the IS-IS protocol when the RIBs are modified to adapt to QoS changes of the subnets.

Part-II (Sections 7.0 and 8.0) builds on the results obtained in Part-I and introduces one more constraint: that of minimizing the routing data overhead as mentioned in (c) above.

Part-III (Sections 9.0 to 12.0) summarizes the findings of the study. Recommendations, conclusions and references are also provided in this last part.

The IS-IS protocol model developed for this simulation includes the Level 2 routing functions but does not support the Level 1 routing functions. Simplified ES and ES-IS protocol models are included to allow traffic generation and loading of the network. These limitations are non-restrictive since the study is mainly concerned with characterizing the system performance within the CSNI Transit Domain (inter-area routing).

2.0 Simulated Network Topology

This study was conducted using a model of the full CSNI network topology rather than a simple generic configuration of subnetworks. The simulated CSNI Network topology is shown in Figure 1. It contains 11 routers, 33 subnetwork access controllers (SNACs) and as many ESs (not shown) forming an ensemble of 13 subnets, as listed in Table 1.

The SNAC is a standard interface between the SNDCP hosted on the ISs and the diverse SNAcPs of the subnetworks comprosing the CSNI testbed. The SNAC, as developed by the CSNI team, also provides support for congestion avoidance, QoS feedback for dynamic routing metric generation and local management of the subnetwork [3].

For simulation purposes, the link connecting a SNAC to a local IS is bandwidth limited at 10 Mbps. In Figure 1, a point-to-point link is represented by a bidirectional arrow interconnecting two SNACs whereas two unidirectional arrows are used to represent a broadcast link.

TABLE 1. List of subnets included in simulation topology of CSNI Transit domain (inter-area)

	SUBNET	ISO 10589 Circuit Type	Transit Domain Neighbours	Nominal Link Capacity[1] (bps)
1	HF North American (HFNA)	Point-to-Point	CA, US1	2400
2	HF North American (HFNA)	Point-to-Point	CA, US3	2400
3	HF North American (HFNA)	Point-to-Point	US1, US3	2400
4	HF Trans Atlantic (HFTA)	Point-to-Point	CA, US3	2400
5	HF Trans Atlantic (HFTA)	Point-to-Point	UK1, US3	2400
6	HF European (HFEM)	Point-to-Point	FR1, NL	1200
7	HF European (HFEM)	Point-to-Point	FR2, UK2	1200
8	UHF Satcom (FLTSAT1)	Broadcast	CA, STC, UK1, UK2, UK3, US3	4800/6= 800
9	UHF Satcom (FLTSAT7)	Broadcast	CA, US1, US3	4800/3= 1600
10	SHF Satcom (DSCS III)	Broadcast	CA, STC, UK1, US1	48000/4= 12000
11	SHF Satcom (NATO IV)	Point-to-Point	GE, UK1	64000
12	SHF Satcom (NATO IV)	Point-to-Point	GE, STC	64000
13	SHF Satcom (NATO IV)	Point-to-Point	NL, STC	64000

1. Nominal Subnet Capacity divided equally among its members

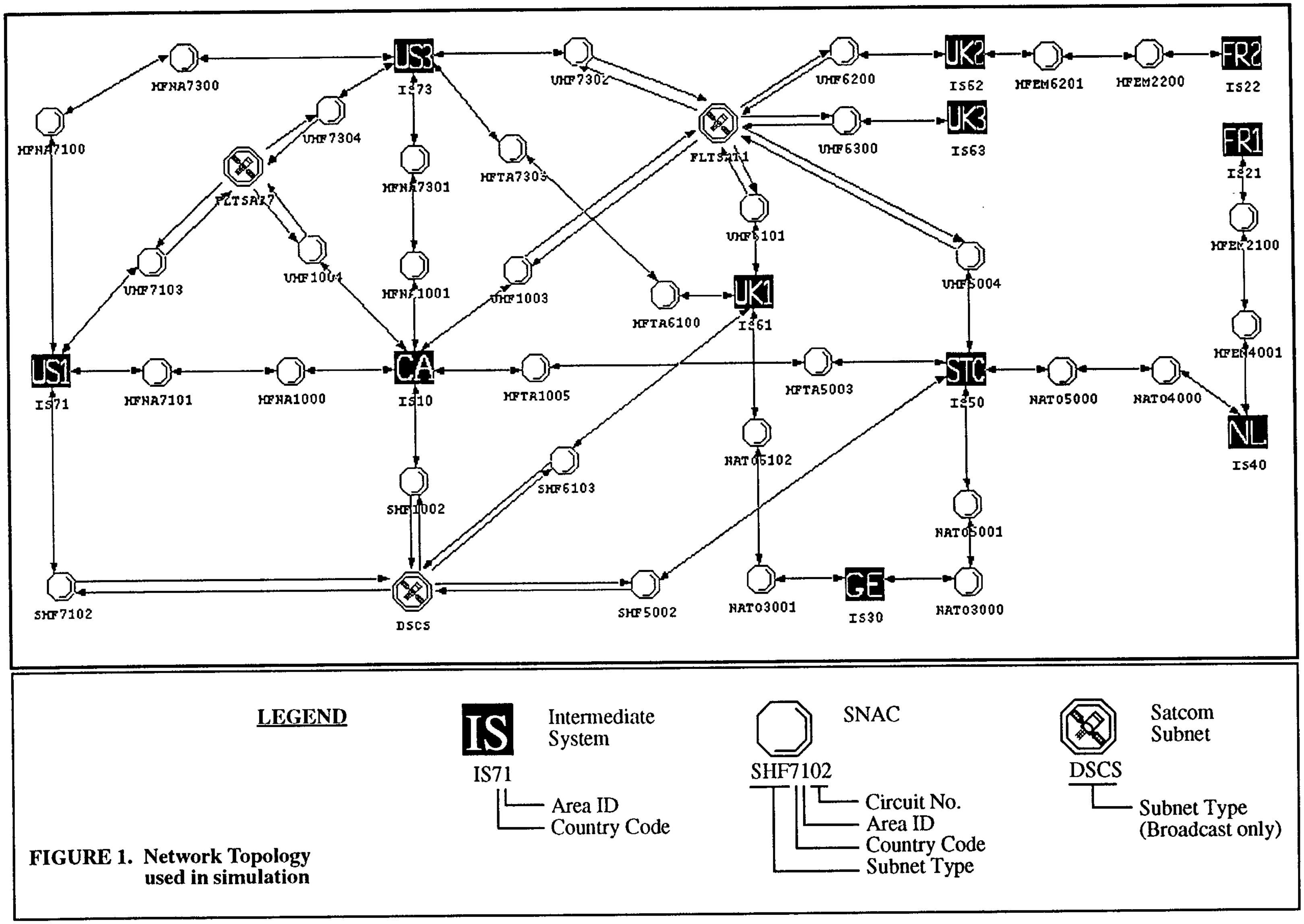


FIGURE 1. Network Topology used in simulation

2.1 Subnetwork Types

In the ISO 10589 standard, subnetworks are classified according to two types:

1. broadcast subnetworks; and,
2. general topology subnetworks.

2.1.1 Broadcast subnetworks

The standard states that the protocol is intended to operate on any broadcast subnetwork which meets the general requirements that it lists in Section 6.7, for example:

> Events like routing PDU non-sequentiality or routing PDU loss due to detected corruption should be rare i.e. on the order of once per 1000 PDUs.

If an average SnSDU size of 512 octets is assumed this figure corresponds to a BER in the order of 10^{-7}. This simulation assumes that these requirements are met by the CSNI SHF and UHF subnets although preliminary field activity reports suggest that this may not be readily achievable. The standard (Section 6.5.1) also recognizes that there are broadcast subnetworks that may not be adequately covered at this time; the low bandwidth broadcast subnets used in CSNI probably fall into this category. Nevertheless, the CSNI requirement for COTS products could be best served by the ISO 8802 LANs (broadcast) subnetwork independent functions which are specifically addressed by ISO 10589. The simulation model includes several of these functions.

2.1.2 General Topology Subnetworks

This group includes:

1. Point-to-point links, i.e. links which connect exactly two systems and stay connected at all times (unless turned off by system management). The IS-IS routing functions provided for point-to-point links in the standard are used in CSNI over the NATO IV links and most HF links.
2. Dynamically assigned (DA) links, i.e. links that are established upon receipt of traffic, and brought down on timer expiration when idle. No IS-IS routing PDUs are exchanged between ISs on a DA circuit.

It is our understanding that although the CSNI European Mesh is essentially a DA type of subnetwork, the routers will not establish switched virtual circuits (SVC) on this subnet. Instead, a number of point-to-point circuits will be activated and artificially initialized and maintained in the "UP" state by the actions of the French SNAC emulating the routing protocol functions. Since this unique approach is beyond the IS-IS protocol design scope and since no documentation of the French SNAC design was available, the HF European subnet was modeled as two static, half-duplex, low-bandwidth (2400 bps) point-to-point links.

2.2 End Systems and Module Characteristics

For clarity of presentation, End Systems are intentionally not depicted in Figure 1. Each area contains in fact as many ESs as the number of circuits supported by the router configured for that area. For example, the US1 IS in area 71 (see Legend in Figure 1) has 4 circuits (2 HF North American, 1 UHF Satcom and 1 SHF Satcom) and therefore 4 ESs are directly attached to this router.

Figure 2 shows how the ESs within an area make use of a single router by sharing the same CLNS Entity through Network Service Access Points (NSAPs). This simplified ES-IS communication model does not support the ISO 10589 Level 1 routing functions but nevertheless allows traffic generation for simulating and testing the Level 2 routing functions.

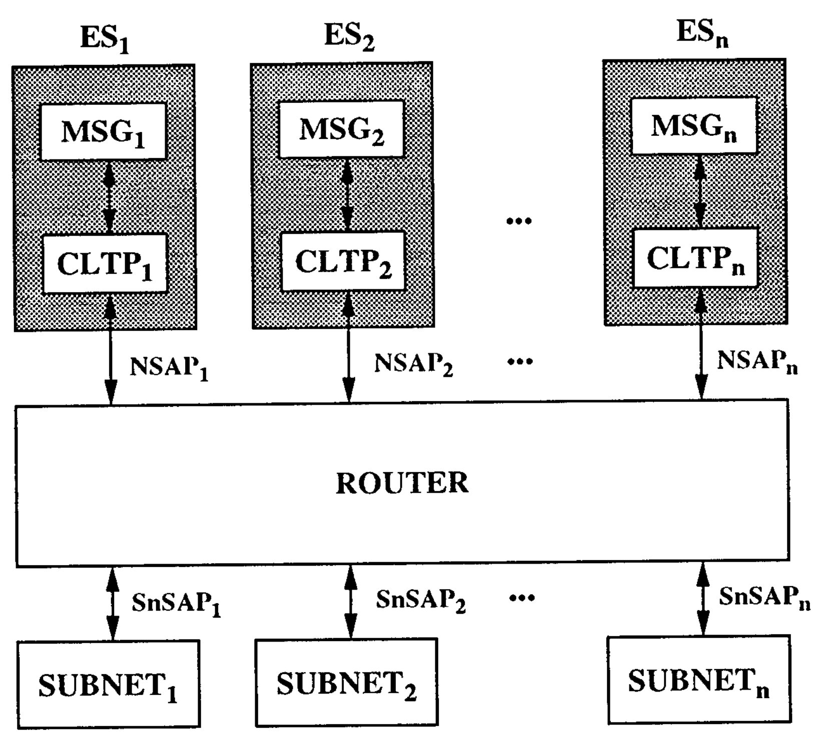


FIGURE 2. ES-IS communication path

The main characteristics of the simulation modules will now be described.

MSG — Packets are generated by a message generator at a rate given by *MSGTransmissionRate* = *nominalLinkCapacity* x *nominalLoad*, where *nominalLinkCapacity* is the nominal capacity (ref. Table 1) of the link the MSG is sending to and, *nominalLoad* is a global parameter controlling the load (0-100%) imposed on the network links during a simulation run.

The size of the packets is dependent on a random process having a uniform distribution within the range [1-1024] bytes. The random generator is set with the same seed after each simulation run. A fixed destination address along with other information such as packet creation time, sequence number, etc., is inserted into each APDU.

CLTP — In the absence of any flow control signal from the router, the CLTP module accepts packets from the MSG and forwards them directly at a rate of 10 Mbps to the router through an NSAP. When a flow control signal is received from the router, the CLTP layer flow controls the MSG and no packets are forwarded.

ROUTER — The packets received from the CLTP layers are forwarded to one of the SNACs based on the current path cost calculated from the local RIB. In this study, all packets are routed using the Expense metric only. This is done so as to maintain controlled and uniform loads throughout the network. This version of the router model includes most of the ISO 10589 Level 2 protocol functions (IIH-PDU generation, LSP/CSNP/PSNP transmissions, RIB updates, etc.)

The IS-IS model does not support dynamic Designated-IS election. Instead, an IS is elected at the start of every new simulation and remains elected for the duration of the simulation. The model creates one LSP per IS and one LSP per Designated IS. For the topology shown in Figure 1, each RIB will contain 11+3=14 LSPs i.e. one LSP for each router (11) and one LSP for each broadcast subnet (3).

SUBNETS — Every SNAC module includes a prioritized queue to buffer the incoming NPDUs while the SNAC transmitter is busy. Included in the SNAC model are the functions to gather queue statistics, generate Statistics (QoS) Reports and dynamically generate new routing metrics. For convenience, the latter functionality is implemented in CSNI by a routing metric generator (RMG) on a host external to the SNAC.

3.0 Performance Parameters

A few terms will now be defined to help understand the results presented in the sections to follow.

In QoS-driven routing, any change in a network link cost causes a LSP to be generated and propagated to all the ISs. The routing function will need some finite amount of time to adapt to each change, i.e. to reach the state where all RIBs are synchronized. A parameter, which we shall call *Divergency*, will be used to monitor whether the RIBs throughout the network contain the most up-to-date metric values for all known paths.

$$\text{Divergency} \equiv \begin{cases} 1 & \text{One RIB (or more) is not in-sync} \\ 0 & \text{All RIBs of the network are in-sync} \end{cases} \tag{1}$$

This parameter is of little use "as is" since it only provides a point indication of the RIBs status. The *Time-Averaged Divergency* will be used to measure the fraction of time the network undergoes transient periods of suboptimal routing due to QoS changes at the subnetwork level.

$$\text{Time-Averaged Divergency} \equiv \frac{1}{T}\int_{0}^{T} \text{Divergency}\,(t)\,dt \tag{2}$$

In Part-I and Part-II of this report, the time average was calculated over the duration T of each simulation run.

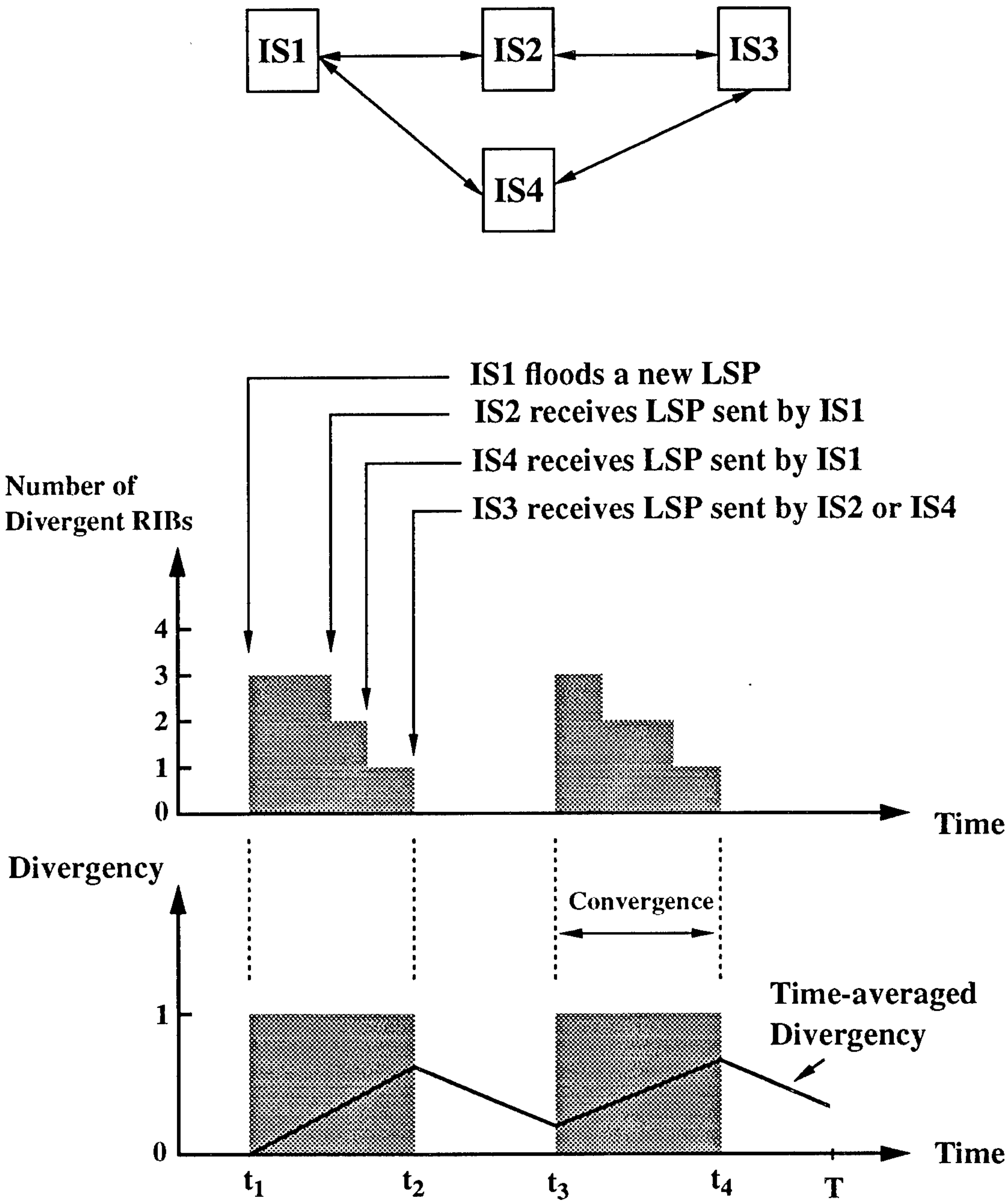
Finally, the amount of time to adapt to each change will be referred to as the *Convergence* time.

$$\text{Convergence} \equiv \begin{cases} \text{Time taken by the Network to re-establish RIB synchronization} \\ \text{(Duration of the Divergency)} \end{cases} \tag{3}$$

Figure 3 illustrates by an example the definitions given above.

4.0 Simulation Parameters

This section provides a brief definition of the main simulation parameters and terms used in this study. The reader should refer to Annex A for an overview of how the CSNI QoS-driven dynamic metric generation mechanisms relate to the routing information database (RIB) update mechanisms provided by the ISO 10589 standard.



$$\text{Time-Averaged Divergency (T)} = \frac{(t_2 - t_1) + (t_4 - t_3)}{T}$$

FIGURE 3. Performance parameters illustrated by an example

4.1 ISO 10589 Parameters

Maximum LSP Age, *MaxAge*

The maximum lifetime of a LSP. This is a protocol constant and its value is 1200 seconds (20 minutes).

Maximum LSP Generation Interval, *maximumLSPGenerationInterval*

The maximum amount of time allowed to elapse between generation of LSPs by a source IS. This timer should be less than the maximum lifetime of a LSP i.e. a source IS should re-generate each LSP periodically at intervals of at most

maximumLSPGenerationInterval < *MaxAge*

Minimum LSP Generation Interval, *minimumLSPGenerationInterval*

The minimum amount of time a source IS should wait before re-generating one of its own LSPs. Setting this timer value too large causes a delay in reporting new link state information. Setting this timer value too small causes too much routing data overhead.

Minimum LSP Transmission Interval[1], *minimumLSPTransmissionInterval*

The minimum amount of time an IS should wait between re-transmissions of a LSP. Setting *minimumLSPTransmissionInterval* greater than *minimumLSPGenerationInterval* makes no sense because the source would be allowed to generate LSPs more quickly than they would be allowed to be sent.

Min. Broadcast LSP Transmission Intvl, *minimumBroadcastLSPTransmissionInterval*

The minimum amount of time between PDU arrivals which can be processed by the slowest IS on a LAN. An IS is permitted to transmit a small number of LSPs/CSNPs (no more than 10) "back to back", provided that no more than

1000/minimumBroadcastLSPTransmissionInterval

LSPs/CSNPs are transmitted in any one second period.

Complete SNP Transmission Interval, *CompleteSNPInterval*

The amount of time between periodic transmissions of a complete set of Sequence Number PDUs by the Designated IS on a broadcast link. CSNPs are only sent during link initialization over point-to-point circuits.

1. The IS model developed for this study follows the CSNI implementation which is based on the OSI Router Software developed by Retix®. Retix interpreted the standard as if this timer was applicable to point-to-point circuits only [2].

Partial SNP Transmission Interval, *PartialSNPInterval*

The amount of time between sending partial Sequence Number PDUs.

IS-IS Hello Transmission Interval, *iSISHelloTimer*

The interval between the generation of IIH PDUs.

DR IS-IS Hello Transmission Interval, *dRISISHelloTimer*

The interval between the generation of IIH PDUs by the Designated IS on a LAN.

4.2 RMG (Routing Metric Generator) Parameters

Minimum Metric Update Interval, *minimumMetricUpdateInterval*

The minimum time interval elapsed between two metric updates by the RMG. This is one of two conditions for the RMG initiating a metric update.

Minimum Metric Delta, *minimumMetricDelta*

The minimum metric change, either absolute (expressed in metric unit, mu) or relative (expressed in percent, %), needed to enable a metric update by the RMG. This is one of two conditions for the RMG initiating a metric update.

4.3 SNAC Parameters

Statistics Report Interval, *StatisticsReportInterval*

The time interval elapsed between two Statistics (QoS) Reports issued by a SNAC to the RMG.

Routing PDU Time-To-Live, *TTL*

The maximum life time assigned to the routing PDUs that the router forwards on a circuit.

4.4 Other Parameters

Nominal Link Capacity, *nominalLinkCapacity*

The *subnetCapacity* divided equally among its members as listed in Table 1.

Subnet Capacity, *subnetCapacity*

The nominal transmission rate of a subnet as per the following table:

SUBNET	*subnetCapacity* (bps)
HFNA	2400
HFTA	2400
HFEM	2400
UHF Satcom	4800
SHF Satcom (DSCS)	48 000
SHF Satcom (NATO IV)	64 000

Nominal Load, *nominalLoad*

A global value between 0 and 100% controlling the transmission rate of all the MSG modules of the network. This parameter is used to control the load (0-100%) imposed on the network links during a simulation run. Note that the Network Load exceeds the *nominalLoad* by an amount related to the amount of CLTP, CLNP and Subnetwork Protocol overheads.

MSG Transmission Rate, *MSGTransmissionRate*

The transmission rate (expressed in bits per second) at which a MSG module sends its messages:

MSGTransmissionRate = *nominalLinkCapacity* x *nominalLoad*

Routing Traffic Percentage, *routingTrafficPercentage*

The fraction of SnSDU bytes generated by the IS-IS protocol that a subnet has accepted and successfully transmitted. The sum of the *routingTrafficPercentage* and *userTrafficPercentage* equals 100%.

User Traffic Percentage, *userTrafficPercentage*

The fraction (expressed in percent) of SnSDU bytes indirectly generated by a MSG that a subnet has accepted and successfully transmitted. The sum of the *routingTrafficPercentage* and *userTrafficPercentage* equals 100%.

Link/Routing Data Overhead, *link/routingDataOverhead*

The link capacity (expressed in bits per second) devoted by a SNAC to the transmission of Level 2 routing data such as LSPs, IIH PDUs, Sequence Number PDUs, etc., including link protocol overheads.

PART - I: STABILITY AND PROMPTNESS

The analysis of the problems listed in the *Introduction* section is non-trivial due to the large number of parameters and subnetwork types involved. Setting some new value for a timer will usually cause:

1. a delay change in propagation of routing information and stabilisation of the routing algorithm; and,
2. a bandwidth change in the resources required to propagate the routing information.

These changes are usually not disjoint and move in directions that do not allow concurrent optimums. For this reason, Part-I of this study mainly explores the setting of the IS-IS/RMG/SNAC timers without a fair consideration of the bandwidth constraint. The routing data overhead requirements will be duly considered in Part-II.

Part I is divided into two sections. The first (Section 5.0) does not allow the metric values to change dynamically as the traffic load is increased. The second (Section 6.0) lets them change according to the rules set for updating the RIBs by the RMGs.

Caution: In both, Part-I and Part-II, most of the simulation runs or averaged statistics represent 5000 seconds (about 1 hr 20 min) of real time network operation and necessitated, depending on the simulation scenario, anywhere from 30 minutes to 8 hours of simulator computation time (SPARC-II workstation). The statistical confidence level of most results presented in this study is less than 85%. Higher levels could have been achieved at the expense of prolonged simulation time or OPNET® code optimization to reduce simulation time. In general, the results are indicative of the trends associated with a parameter change rather than being an absolute measure of performance.

5.0 Static Metric Values

In this section, the routing metrics are kept constant. Consequently, Statistics Reports sent by the SNACs are ignored by the RMGs and the routers use the initial metric values set for the circuits during the start-up phase of the simulation.

5.1 Default IS-IS Timer Values

The first simulation results were obtained with the values listed in Table 2. The table lists the default values of the main ISO 10589 timers except for the two Hello transmission intervals that have been arbitrarily scaled up by a factor of 10. Initial simulation runs (not presented here) have shown that the default Hello transmission rates constitute too much routing data overhead for the low bandwidth subnetworks of the CSNI transit domain.

Examples of time series recorded during these simulation runs are shown in Figure 4. The figure illustrates how updated-LSPs propagate throughout the network during the last 500 seconds of a simulation run.

The topology contains 11 RIBs (one for each IS). As soon as one LSP is regenerated, 10 RIBs become de-synchronized and need to be updated. There is one instance in Figure 4 (top) where more than one LSP were being re-generated concurrently. This occurs around the 4825th second of the simulation where all RIBs (11) of the network were divergent. As a LSP is being flooded throughout the network, RIBs resynchronize one by one and eventually all RIBs become synchronized again.

The black strips on the middle graphic in Figure 4 represent the time intervals where the network is in transition. The duration of these intervals is the network convergence time, as defined in Section 3.0, and is given by the bottom graphic in Figure 4. In this figure, the convergence time varies between 6 and 23 seconds. Over the whole 5000 seconds that this simulation run lasted, the convergence time was distributed as shown by the 2 second frequency histogram in Figure 5. The cumulative distribution function shown on the same page indicates that half of the time, RIB updates were completed in less than 15 seconds. It is interesting to note that the function is quite asymmetric around x=15 seconds which is most likely due to the combined presence of relatively fast and slow subnetworks.

The Time-Averaged Divergency of the network is obtained by applying eq. (2) to the network Divergency time series. This is equivalent to summing the convergence intervals, dividing the result by the duration of the simulation (5000 seconds) and multiplying by 100%. The result is shown in Figure 6. When no message is sent, the network spends 20% of its time going through transient periods due to periodic LSP regenerations. This value increases almost linearly at the rate of 0.3% per each network load percentage increase. When the load exceeds 60%, the Time-Averaged Divergency decreases although this effect is probably attributable to a Congestion Control (CC) reaction. To minimize the effects of the congestion signals, a short Source Quench "blocking" time of 1.0 second was used.

TABLE 2. Parameter set "A": Static metrics, default ISO 10589 timer values

PARAMETER		VALUE[1]
IS	Max. LSP Generation Interval (min)	15
	Min. LSP Generation Interval	30
	Min. LSP Transmission Interval	5
	Min. Broadcast LSP Transmission Interval	0.033
	Complete SNP Transmission Interval	10
	Partial SNP Transmission Interval	2
	IS-IS Hello Transmission Interval	$3 \rightarrow 30$
	DR IS-IS Hello Transmission Interval	$1 \rightarrow 10$
RMG	Min. Metric Update Interval (min)	n/a
	Min. Metric Delta (mu)	n/a
SNAC	Statistics Report Interval	n/a
	Routing PDU TTL	128
Other	Nominal Load (%)	variable

1. All parameters expressed in seconds unless otherwise specified.

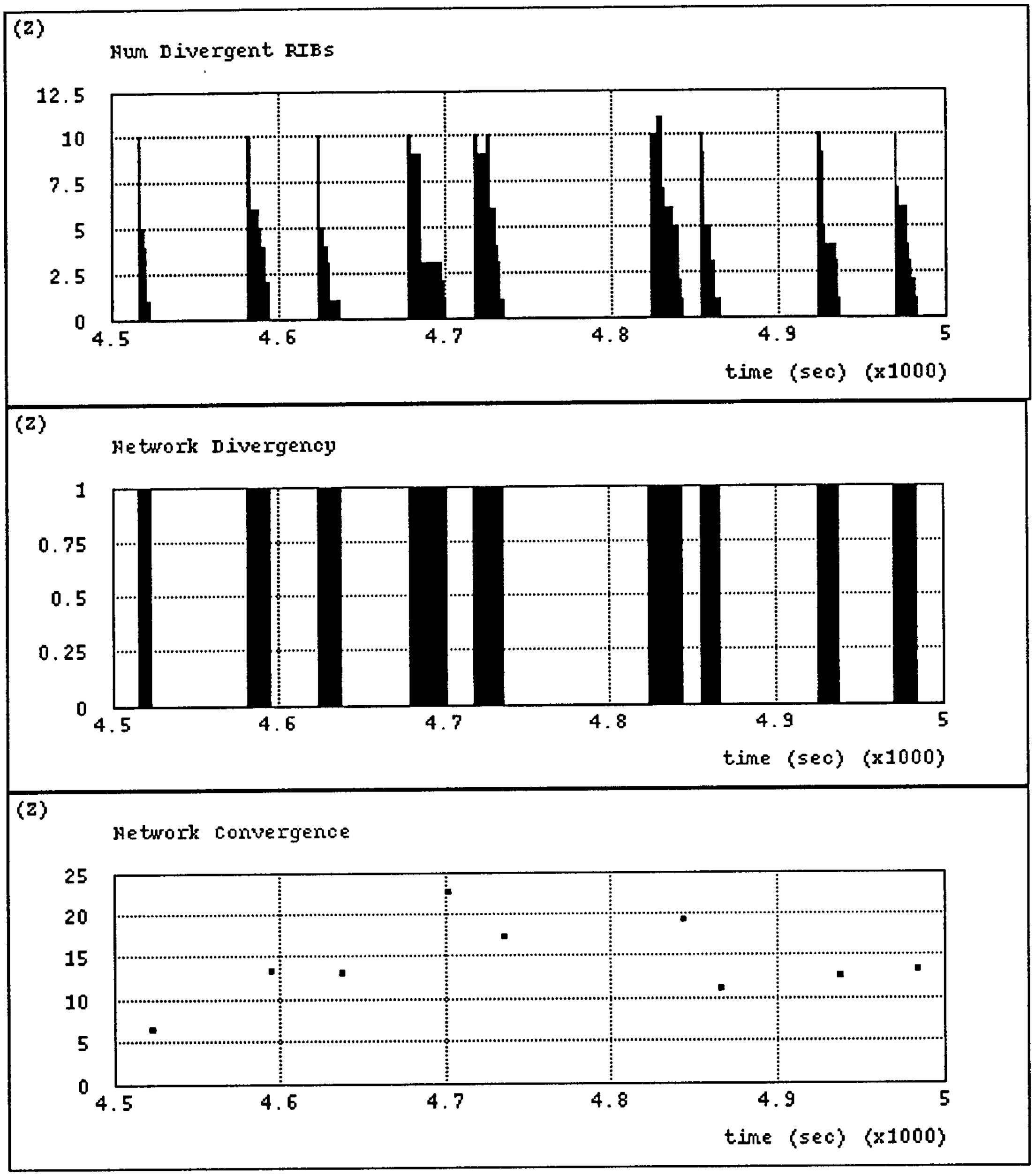


FIGURE 4. Example of RIB transient periods due to LSP propagation throughout the transit domain

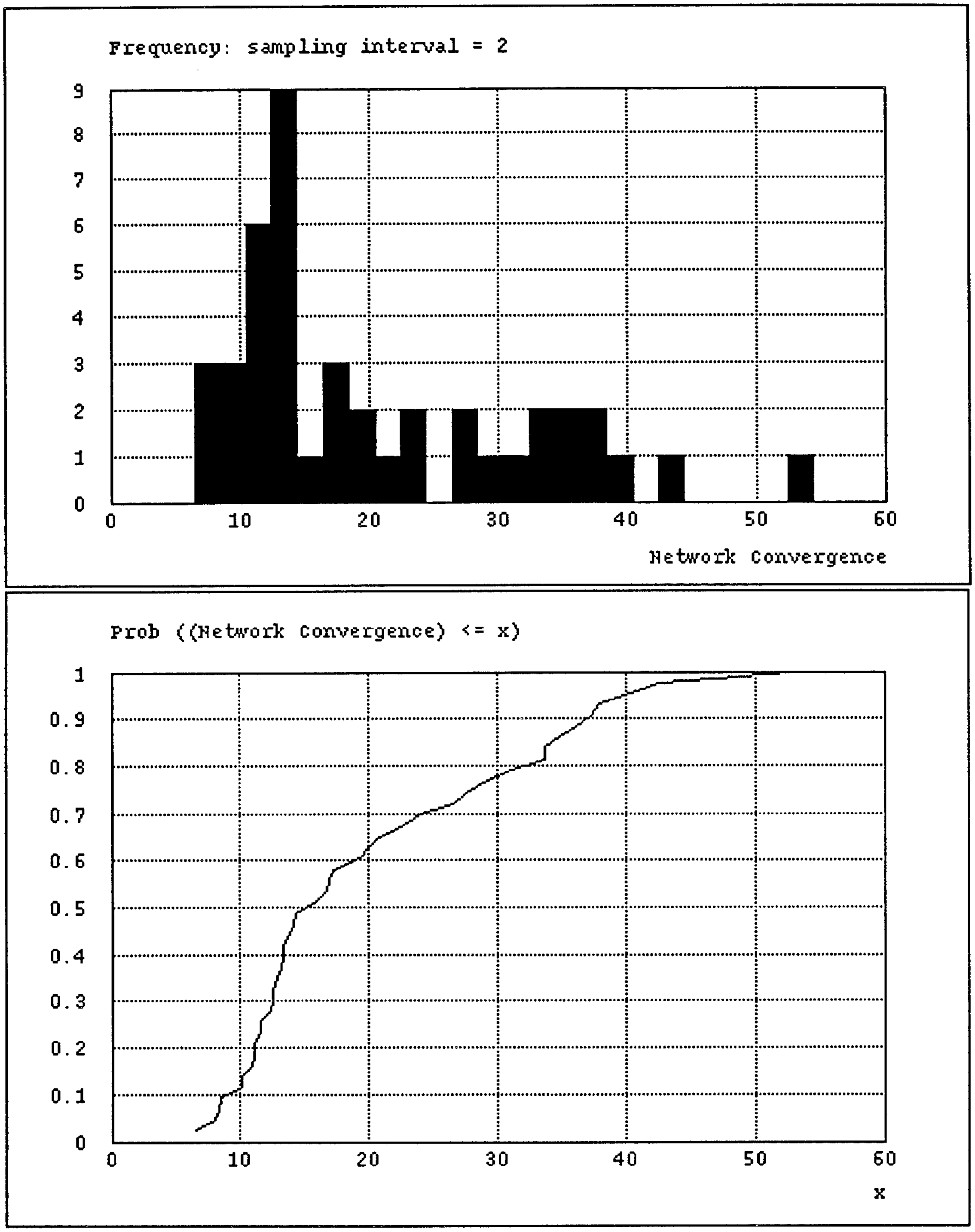


FIGURE 5. Example of histogram and CDF computed from network Convergence time

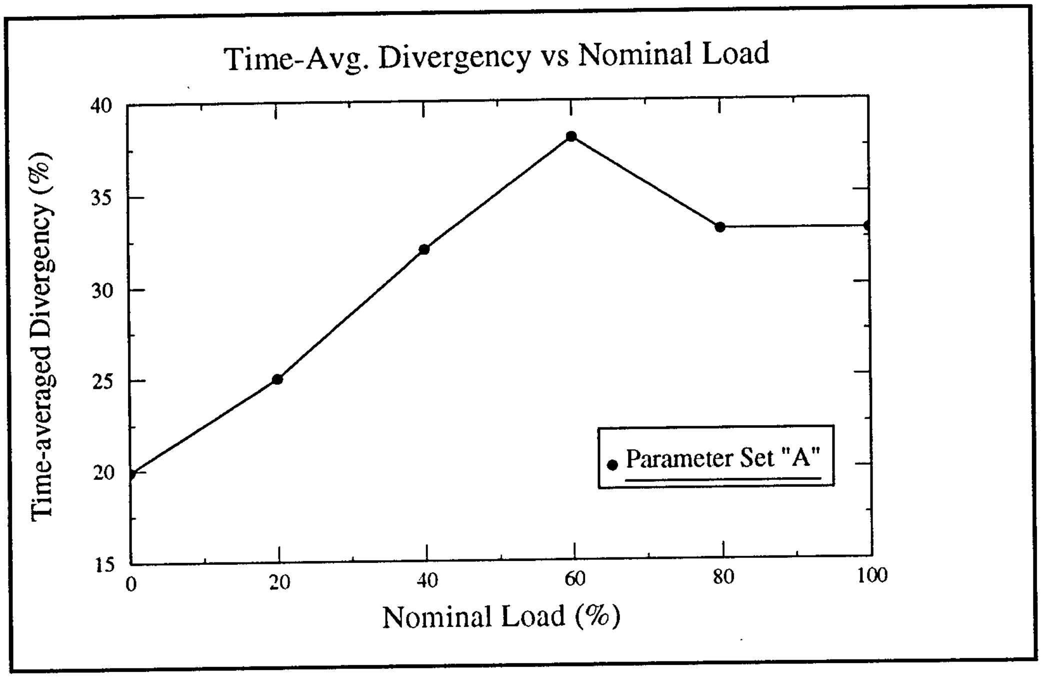


FIGURE 6. Network performance when using static metrics and default ISO 10589 timer values

5.2 Sensitivity Analysis

In an attempt to improve the performance obtained when using the default ISO 10589 timer values, a sensitivity analysis was conducted and its results are presented in this section.

Before this could be done however, one thing needed to be verified: does it matter which IS is elected to perform the pseudonode duties on a broadcast subnetwork? Designated ISs generate additional LSPs on behalf of the LAN and one could think that some ISs in a topology are better positioned than others to perform this role. The IS located at the CRC for example is the only IS attached to all the broadcast subnetworks of the topology being studied and is the IS having the most number of circuits (6).

Various combinations of ISs were elected (forced election) as shown in Table 3. A Time-Averaged Divergency of 19.9% was previously obtained when the IS at US3 (NRL) is the Designated IS for the two UHF Satcom subnetworks and, the IS at US1 (NRaD), the Designated IS over the SHF Satcom subnetwork. Results show that it does not matter much which IS is elected although longer simulation runs would be needed to obtain a better statistical confidence level to confirm this result. For the remainder of this Part-I, the configuration listed last in Table 3 (DSCS-STC, FLTSAT1-STC, FLTSAT7-CA) will be used.

TABLE 3.

DESIGNATED IS			
DSCS	FLTSAT1	FLTSAT7	Time-averaged Divergency (%)
STC	UK1	CA	20.7
STC	UK1	CA	20.7
US1	US3	US3	19.9
UK1	UK1	US1	19.8
US1	STC	US1	19.8
US1	UK3	US3	19.6
CA	CA	CA	19.4
STC	STC	CA	19.3

The sensitivity analysis will now be presented.

The *maximumLSPGenerationInterval* timer is periodic. It defines the maximum cycle duration for updating a LSP, whether the metric values have changed or not. This is required since every LSP has a maximum lifetime. This timer guarantees the re-incarnation of a LSP before it dies. Through flooding, the new LSP overwrites the old LSP in every RIB. The maximum timer value is 20 minutes, as specified by the protocol constant *MaxAge*. In practice, the timer should not be set to this maximum value since some finite amount of time is needed to flood the new LSPs before the old ones expire. Table 4 shows the simulation results when the timer value is increased from 15 to 18 minutes in 1 minute increment steps. A longer LSP regeneration cycle reduces the Time-Averaged Divergency.

The *minimumLSPGenerationInterval* timer prevents a source IS[1] from re-generating its own LSP(s) faster than the time interval set for this timer. This timer was not included in the sensitivity analysis to maintain a common base for comparison.

The *minimumLSPTransmissionInterval* timer applies to point-to-point circuits only (see footnote on page 9). A same LSP cannot be retransmitted on a given circuit unless the time interval set for this timer has elapsed. This timer is periodic. Any newly arrived LSP or locally generated LSP must wait until the end of the timer cycle before being transmitted on the circuit. Table 5 shows that by reducing the timer interval, faster convergence time can be achieved.

1. Every IS is responsible for (re-)generating one LSP, or more if needed, containing the routing costs associated with the transmit circuit of its active links.

TABLE 4.

MAX. LSP GENERATION INTERVAL	Time-averaged Divergency (%)
15 min.	19.9
16 min.	20.1
17 min.	19.9
18 min.	16.2

TABLE 5.

MIN. LSP TX. INTERVAL	PARTIAL SNP TX INTERVAL	Time-averaged Divergency (%)
7.0	2.8	22.9
6.0	2.4	22.6
5.0	2.0	19.9
4.0	1.6	17.8
3.0	1.2	17.5

TABLE 6.

MIN. BROADCAST LSP TRANSMISSION INTERVAL	Time-averaged Divergency (%)
.033	19.9
.10	19.3
.33	20.9
1.0	19.9

The *PartialSNPInterval* timer controls the amount of time between sending partial Sequence Number PDUs. PSNPs are used by non Designated-ISs over broadcast circuits to request transmission of a specified set of LSPs from the Designated-IS. Over point-to-point circuits, PSNPs are used to acknowledge a set of LSPs. The time interval for this timer is by default 60% less than the time interval specified for the *minimumLSPTransmissionInterval* timer. The same scaling factor was used to match the changes made to the *minimumLSPTransmissionInterval* and the *PartialSNPInterval* timer values as shown in Table 5

The *minimumBroadcastLSPTransmissionInterval* timer applies to broadcast circuits only. As mentioned in Section 2.1.1, the routing functions provided for the ISO 8802 LANs do not adequately cover the CSNI broadcast subnetwork requirements. The recommended default timer value is 33 ms, which is adequate for the Level 1 LLC circuits but inadequate for the Level 2 SNAC circuits. Being a CLNS-MO attribute, all broadcast circuits share this "global" timer setting. To preserve the LLC1 subnet performance, this timer cannot be changed sufficiently to make significant difference in the Level 2 subnet performance, as shown in Table 6. This is because PDUs pile up in the SNAC queues rather than being broadcasted immediately as is the case for the high bandwidth 8802 LANs.

The *CompleteSNPInterval* timer is also periodic and only used over broadcast circuits. Furthermore, only Designated-ISs are allowed to periodically send CSNPs. In principle, a short timer interval will help disseminate the most up-to-date Link State information. However, this is mainly true for ISs that are joining the net. For those that have been attached to the LAN for some time, any new LSP is captured right away as it is multicasted on the LAN for the first time. This explains why the values recorded in Table 7 show little variation as the timer interval is increased to reduce the routing data overheads.

TABLE 7.

COMPLETE SNP TRANSMISSION INTERVAL	Time-averaged Divergency (%)
10	---
12	16.9
15	16.9
20	15.8
25	17.0
30	17.2

The *iSISHelloTimer* controls the generation of IIH PDUs over point-to-point and broadcast circuits. The *dRISISHelloTimer* is a CLNS-MO attribute and supersedes the *iSISHelloTimer* when the IS has been elected Designated-IS on a broadcast circuit. The *iSISHelloTimer* is a circuit attribute and a value can therefore be set independently for each circuit. IIH PDUs sent over broadcast circuits are padded to the maximum CSNI SnSDU size of 1420 bytes and can constitute a significant routing traffic load for the low bandwidth broadcast subnets of the transit domain. Table 8 shows that some improvement in convergence time can be obtained when the Designated ISs send Hello PDUs less frequently.

TABLE 8.

DR IS-IS HELLO TRANSMISSION INTERVAL	Time-averaged Divergency (%)
10	19.9
15	18.1
20	17.0
25	18.0
30	17.6

The changes discussed above are summarized in Table 9. For easy comparison, the values used in the previous simulation runs are also indicated. The simulation was re-run with the new values and the results are plotted in Figure 7. As a whole, the changes have significantly improved the Time-Averaged Divergency although it still is surprisingly high, especially since no data is being sent across the network. When the network is heavily loaded, the network is in a transient state for about one third of the time.

TABLE 9. Parameter set "B": Static metrics, optimum timer values determined by sensitivity analysis

	PARAMETER	VALUE[1]
IS	Max. LSP Generation Interval (min)	15 → 18
	Min. LSP Generation Interval	30
	Min. LSP Transmission Interval	5 → 3
	Min. Broadcast LSP Transmission Interval	0.033 → 0.1
	Complete SNP Transmission Interval	10 → 20
	Partial SNP Transmission Interval	2 → 1.33
	IS-IS Hello Transmission Interval	3 → 30 → 60
	DR IS-IS Hello Transmission Interval	1 → 10 → 30
RMG	Min. Metric Update Interval (min)	n/a
	Min. Metric Delta (mu)	n/a
SNAC	Statistics Report Interval	n/a
	Routing PDU TTL	128
Other	Nominal Load (%)	variable

1. All parameters expressed in seconds unless otherwise specified.

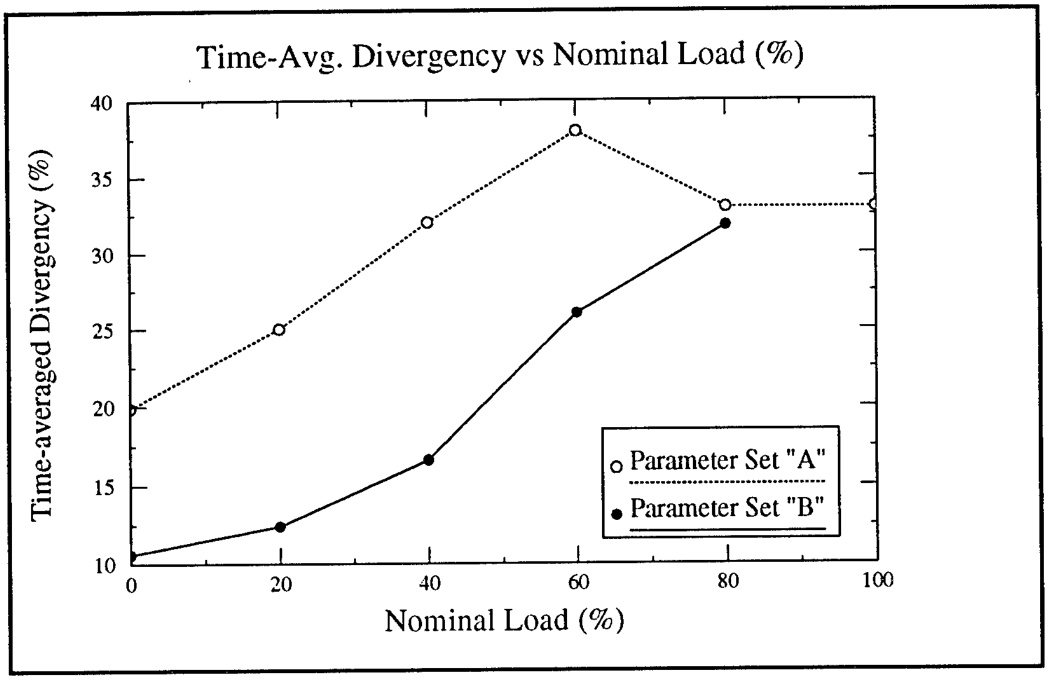


FIGURE 7. Network performance when using static metrics and sensitivity analysis results

6.0 Dynamic Metric Update

In Section 5.0, the RMG was disabled to prevent any metric change and to exercise the OSI routing functions within the quasi-static environment foreseen by the standard. In this section, the RMG will be enabled and the metrics allowed to change according to the QoS computed dynamically by the SNACs and the adaptation scheme developed by CSNI.

6.1 Asynchronous Metric Generation

The simulator will be run with the IS timer values determined by the sensitivity analysis. This time however, the SNACs are configured to generate Statistics (QoS) Reports at periodic intervals of *StatisticsReportInterval*=30 seconds. The 33 SNACs of the topology are randomly activated during the first 30 seconds of the simulation to avoid time synchronization of the Reports. These Reports are processed by the RMGs which must decide whether to update the metrics or not. The conditions necessary for triggering a metric update are set as follows:

1. the previous metric update is at least *minimumMetricUpdateInterval* seconds old. This parameter remains constant during a simulation run but varies from run to run.
2. the metric must have changed by an amount equal to or exceeding *minimumMetricDelta* metric unit[1] (mu). This parameter is kept constant for all the simulation runs. A value of 1 mu is used when the capacity metric is being tested and a value of 5 if it is the delay metric. The other two metric deltas need not be specified since every link of the topology is assigned a fixed expense and error metric value for the duration of the simulation.

A summary of the parameter assignments is given in Table 10. Shaded areas identify changes relative to the configurations described previously.

The network performance is plotted in Figure 8 (upper graphic). Although the figures obtained are high (Time-averaged Divergency >35%), the general trends are as expected i.e. more frequent link cost updates cause the network to undergo more frequent transient periods of suboptimal routing. Also, as more traffic is sent through the network, these transients become longer because of the routing information being delayed by the non-routing transmissions already in progress.

Figure 8 also indicates that under some conditions the QoS updates keep coming faster than the time interval needed by the network to complete the Link State advertisements for updating all the RIBs with the previous QoS values. Where the Time-Averaged Divergency reaches 100%, the network is constantly in a state of suboptimal routing.

There is one result in Figure 8 that needs further explanation. When the nominal load is at 20% and the *minimumMetricUpdateInterval* is 10 minutes or less, the Time-Averaged Divergency is 100%. For the same metric update interval, higher loads apparently produce better results. This simulation run was repeated and analysed more carefully to discover that the HF link NL→FR1 (ref. Figure 1) had accumulated a time lag of about 120 seconds which caused the RIB at IS21 in France to remain desynchronized for the duration of the simulation. Under light network loading conditions, the queue of the SNAC (HFEM4001) in the Netherlands predominantly fills up with routing information PDUs (LSPs, PSNPs, etc.). When the French router receives this information, the latter is significantly outdated. Similarly, most LSP acknowledgments received by the Netherlands IS from the French IS are old and cause up-to-date routing information to be resent. This problem solves itself when the network load is increased because without queuing preemption, some LSPs must be discarded by the Netherlands' IS when the SNAC queue fills up. Less outdated routing information is then sent to the French node and RIB resynchronization can be reestablished. These results show that it is preferable to discard routing information than to send it if it is outdated. A way to achieve this is by reducing the routing PDU Time-to-Live as will be verified in Section 6.4.

---

1. The CSNI RMG software code implements a relative metric change comparison. When the metric update intervals are relatively long (exceeding 60 seconds), the use of an absolute or relative *minimumMetricDelta* becomes less relevant.

6.2 Synchronous Metric Generation

The current CSNI RMG design specifications do not define any special rule for handling the metric updates of multiple SNACs attached to the same RMG. Statistics Reports are processed asynchronously as they are received and the metrics associated with the corresponding circuit are updated according to, and as soon as, the two conditions listed in Section 6.1 are met. For comparison purposes, the simulation scenario described in Section 6.1 was run again, with the same configuration parameters, except that two more conditions for triggering a metric update were added to the RMG model:

3. all metric updates performed by an RMG must occur on a common local clock tick of *minimumMetricUpdateInterval* seconds;
4. all metric updates performed by any RMG throughout the topology must occur on a common global clock tick of *minimumMetricUpdateInterval* seconds.

Condition 3 causes some additional delays before reporting a change in QoS however, it indirectly reduces the number of instances a LSP must be regenerated and flooded to inform the network of the link cost changes. The reason is that the metric changes reported to an IS by an RMG are most likely inserted into the same LSP since a single LSP can hold the link state information of more than 100 neighbours. Flooding a single LSP containing all the changes is more efficient than flooding the same LSP each time the QoS of any one of its circuits changes.

Condition 4 may or may not be a desirable thing to do. If most links were to undergo a cost change all at the same time then the network might be prone to routing decision errors during this "short", but intensive, adaptation period. This aspect has not been investigated further in this report.

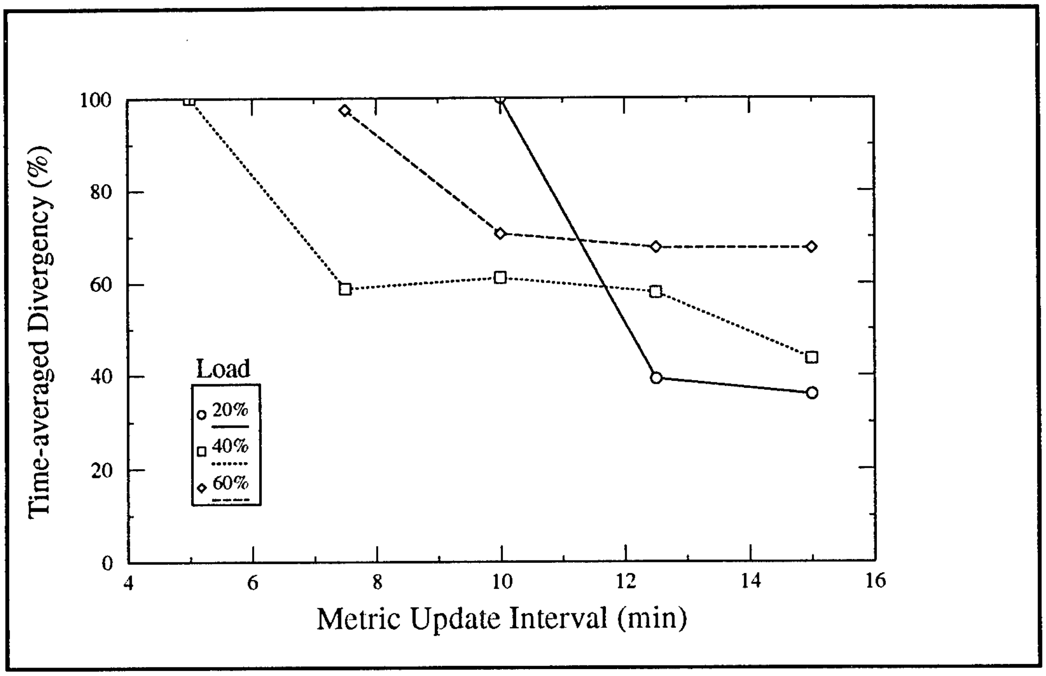
The performance obtained when a synchronous metric generation scheme is used are shown in Figure 8 (lower plot). By comparison with Figure 8 (upper plot), the results indicate that time synchronization of the metric updates at the RMGs can significantly reduce the time spent by the network adapting to QoS-based routing changes.

TABLE 10. Parameter set "C": Dynamic metric update

	PARAMETER	VALUE[1] A → B	VALUE[1] C
IS	Max. LSP Generation Interval (min)	15 → 18	18
	Min. LSP Generation Interval	30	30
	Min. LSP Transmission Interval	5 → 3	3
	Min. Broadcast LSP Transmission Interval	0.033 → 0.1	0.1
	Complete SNP Transmission Interval	10 → 20	20
	Partial SNP Transmission Interval	2 → 1.33	1.33
	IS-IS Hello Transmission Interval	3 → 30 → 60	60
	DR IS-IS Hello Transmission Interval	1 → 10 → 30	30
RMG	Min. Metric Update Interval (min)	n/a	variable
	Min. Metric Delta (mu)	n/a	{1,5,∞,∞}
SNAC	Statistics Report Interval	n/a	30
	Routing PDU TTL	128	128
Other	Nominal Load (%)	variable	variable

1. All parameters expressed in seconds unless otherwise specified.

(Parameter Set "C")



(Parameter Set "C", RMGs synchronized)

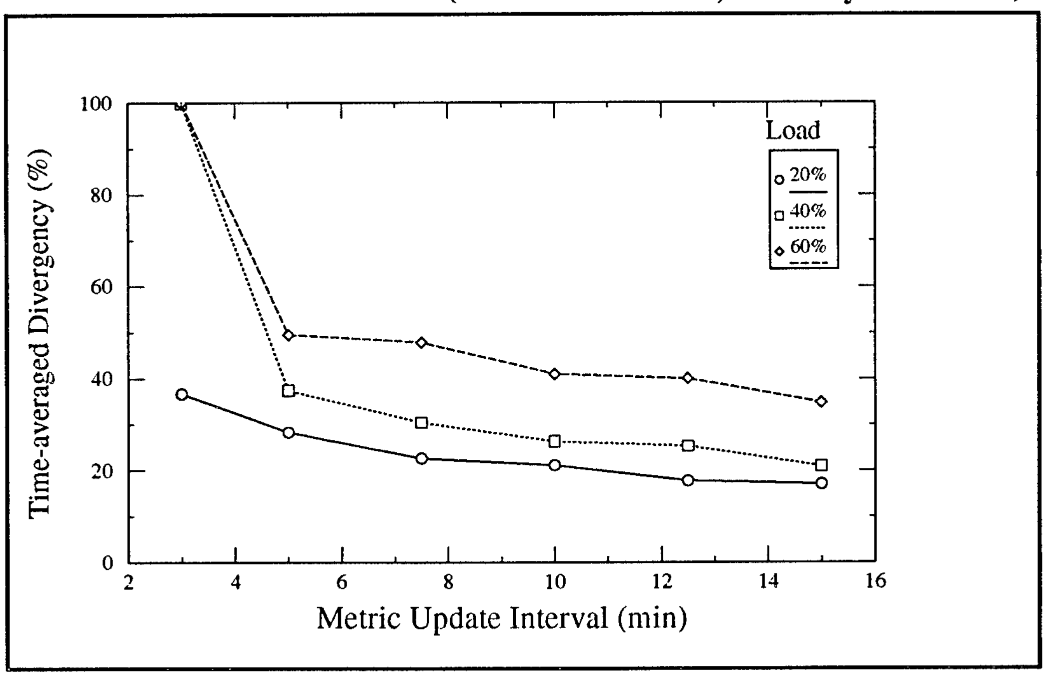


FIGURE 8. Network performance achieved with QoS-driven metric generation

6.3 Reduced Connectivity-Reporting Rate

All the results produced so far have one characteristic in common: they were all produced using a *minimumLSPGenerationInterval* of 30 seconds. The *minimumLSPGenerationInterval* timer controls the amount of time a source IS must wait before regenerating one of its own LSPs, if needed. A LSP can be regenerated for different reasons but there are two that are of specific interest for this discussion:

1. a change in circuit metric; and,
2. a change in adjacency status (up/down event).

The *minimumMetricUpdateInterval* timer of the RMGs provides a mechanism for limiting the rate of change of the circuit metrics. Adjacencies can go up or down in an unpredictable manner and when they do, it is important that the state of these destinations be made known rapidly throughout the network. For these reasons, it has been so far judged more appropriate to limit the rate of change in circuit metrics with the *minimumMetricUpdateInterval* timer and that of the adjacency connectivity with the *minimumLSPGenerationInterval* timer.

Despite these considerations, some simulation runs will now be performed with a *minimumLSPGenerationInterval* of 90 and 180 seconds as indicated in Table 11. The *iSISHelloTimer* of the UHF Satcom circuits will also be increased to 180 seconds. Both of these changes cause additional delays in reporting the connectivity changes.

The results are shown in Figure 9. Doubling the *minimumLSPGenerationInterval* from 90 to 180 seconds does not significantly improve the performance. These results were obtained with a nominal load set at 20% and can be compared with the corresponding curve of the upper plot in Figure 8. The anomaly detected at *minimumMetricUpdateInterval* equals 10 minutes or less which was commented upon Section 6.1 is no longer present in Figure 9. A longer time interval between LSP regeneration allows every IS more time to achieve synchronization of its RIB.

TABLE 11. Parameter set "D": Dynamic metric update, reduced Connectivity-Reporting rate

	PARAMETER	VALUE[1] A → C	VALUE[1] D
IS	Max. LSP Generation Interval (min)	15; 18	18
	Min. LSP Generation Interval	30	90 & 180
	Min. LSP Transmission Interval	5; 3	3
	Min. Broadcast LSP Transmission Interval	.033; 0.1	0.1
	Complete SNP Transmission Interval	10; 20	20
	Partial SNP Transmission Interval	2; 1.33	1.33
	IS-IS Hello Transmission Interval	3; 30; 60	60, 180[2]
	DR IS-IS Hello Transmission Interval	1; 10; 30	30
RMG	Min. Metric Update Interval (min)	variable	variable
	Min. Metric Delta (mu)	{1,5,∞,∞}	{1,5,∞,∞}
SNAC	Statistics Report Interval	30	30
	Routing PDU TTL	128	128
Other	Nominal Load (%)	variable	20

1. All parameters expressed in seconds unless otherwise specified.
2. UHF Satcom circuits only.

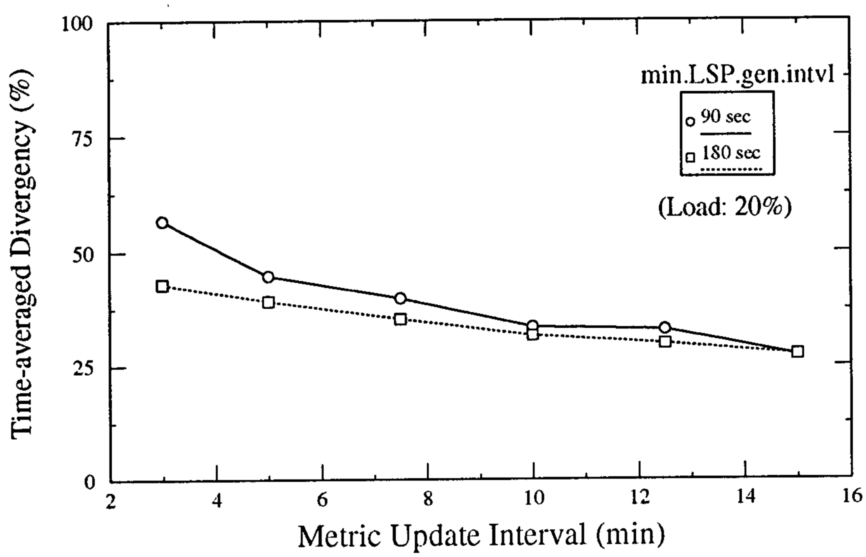


FIGURE 9. Network performance achieved with reduced Connectivity-Reporting rate

6.4 Routing PDU TTL

The last results to be presented in Part-I of this study were obtained by varying the Time-to-Live value assigned to the routing PDUs that the router forwards to the SNACs via a Subnetwork Unit Data Request primitive. Note that the TTL field is part of the primitive, not a field of the NPDU itself.

Table 12 summarizes the parameter assignments made for generating the results plotted in Figure 10. A *minimumLSPGenerationInterval* of 180 seconds was temporarily retained for limiting the LSP generation rate. Similarly, the link cost update rate was limited by the *minimumMetricUpdateInterval* timer to one change or less per 10 minutes.

The results indicate that the Time-Averaged Divergency of the network can be improved by reducing the TTL value when the network load is high (60%).

TABLE 12. Parameter set "E": Dynamic metric update, reduced TTL of routing PDUs

	PARAMETER	VALUE[1] A → D	VALUE[1] E
IS	Max. LSP Generation Interval (min.)	15; 18	18
	Min. LSP Generation Interval	30; 90/180	180
	Min. LSP Transmission Interval	5; 3	3
	Min. Broadcast LSP Transmission Interval	.033; 0.1	0.1
	Complete SNP Transmission Interval	10; 20	20
	Partial SNP Transmission Interval	2; 1.33	1.33
	IS-IS Hello Transmission Interval	3; 30; 60/180[2]	60/180[2]
	DR IS-IS Hello Transmission Interval	1; 10; 30	30
RMG	Min. Metric Update Interval (min)	variable	10
	Min. Metric Delta (mu)	{1,5,∞,∞}	{1,5,∞,∞}
SNAC	Statistics Report Interval	30	30
	Routing PDU TTL	128	variable
Other	Nominal Load (%)	20	variable

1. All parameters expressed in seconds unless otherwise specified.
2. UHF Satcom circuits only.

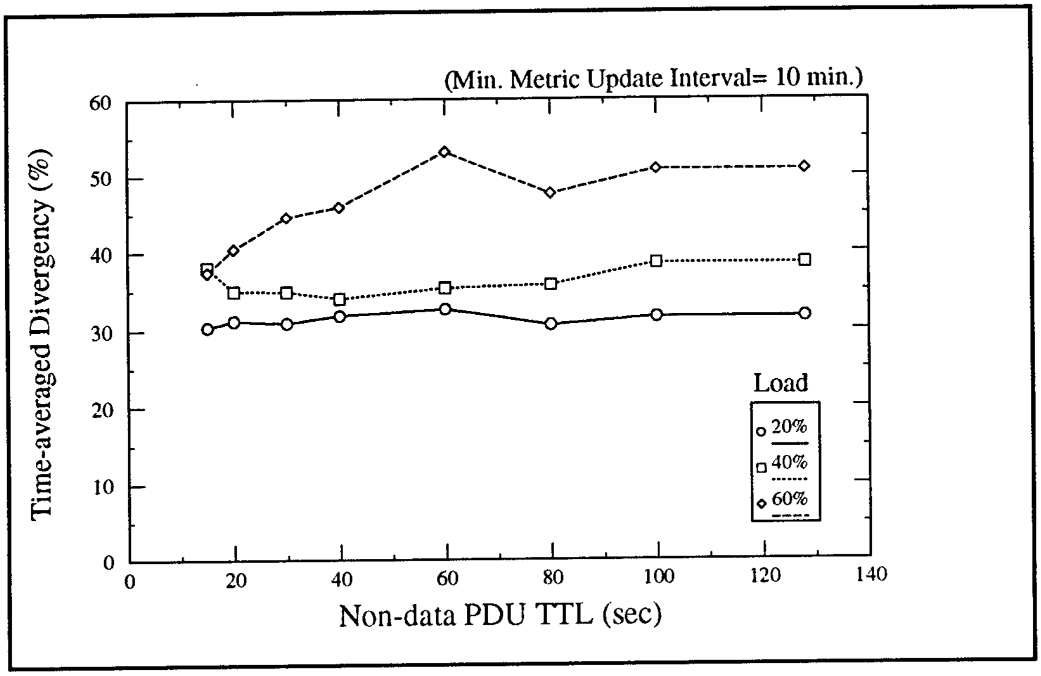


FIGURE 10. Network performance achieved with reduced routing PDU TTL value

PART-II: EFFICIENCY

In Part-I of this study, several limitations of the IS-IS protocol were identified. CSNI Network instability due to RIB de-synchronization was observed and the IS-IS/RMG/SNAC timers adjusted to minimize the amount of time the network requires to adapt to each QoS change and reach a stable, global and deterministic routing decision state.

The results obtained in Part-I will be the starting point of this second part of the study. In Part-II, the amount of routing data overhead due to the IS-IS protocol will be estimated, then minimized by readjusting some of the timers and system parameters. In most instances, the IS-IS data overhead and routing function adaptation speed, including connectivity-reporting rate (neighbour greeting), are tied together and require some trade-off if "good" overall system performance is to be achieved. The choices made by the author are somewhat arbitrary but should nevertheless provide useful indications of what to expect when operating the CSNI testbed under analogous conditions.

7.0 Link/Routing Data Overhead Estimate

Based on the results obtained in Part-I, the parameter values listed in Table 13 are proposed to begin this routing data overhead analysis. The main characteristics of a first simulation scenario are:

- routing traffic only, no ES data generation (Nominal Load = 0%);
- the metric changes are limited to one update per 10 minutes;
- reduced routing PDU TTL (30 seconds).

Note also that the capacity and delay metric deltas are set to a value[1] of 50%.

A summary of the resulting routing data flow is presented in Table 14. The link number and packet flow direction can be derived from the "source SNAC" label listed in the first column (see Figure 1 for topology and SNAC label descriptions). The second column lists the average aggregate of Level 2 routing data overhead plus link protocol overhead, expressed in bits per second, measured at a SNAC output. The third column expresses these averages as a percentage of the nominal transmission capacity available on the corresponding link.

1. Most of the results in Part-I were produced before that value was specified by the CSNI System study group. This explains why different *minimumMetricDelas* were used in that part of the study.

TABLE 13. Parameter set "F": Link/Routing Data Overhead analysis

	PARAMETER	VALUE[1]
IS	Max. LSP Generation Interval (min)	18
	Min. LSP Generation Interval	90
	Min. LSP Transmission Interval	3
	Min. Broadcast LSP Transmission Interval	0.1
	Complete SNP Transmission Interval	20
	Partial SNP Transmission Interval	1.33
	IS-IS Hello Transmission Interval	60/180[2]
	DR IS-IS Hello Transmission Interval	30
RMG	Min. Metric Update Interval (min)	10
	Min. Metric Delta (%)	{50,50,∞,∞}
SNAC	Statistics Report Interval	30
	Routing PDU TTL	30
Other	Nominal Load (%)	0

1. All parameters expressed in seconds unless otherwise specified.
2. UHF Satcom circuits only.

TABLE 14. Overheads as per configuration listed in Table 13.

Source SNAC	Link/Routing Data Overhead (bps)	(%)[1]
HFNA1000	56.9	2.4
HFNA1001	62.2	2.6
HFNA7100	63.4	2.6
HFNA7101	53.4	2.2
HFNA7300	55.0	2.3
HFNA7301	51.6	2.2
HFTA1005	54.1	2.3
HFTA5003	56.6	2.4
HFTA6100	63.8	2.7
HFTA7303	55.5	2.3
HFEM2100	152.9	12.7
HFEM2200	152.6	12.7
HFEM4001	183.9	15.3
HFEM6201	179.4	15.0
NATO3000	54.7	0.09
NATO3001	61.1	0.10
NATO4000	46.9	0.07
NATO5000	74.5	0.12
NATO5001	73.5	0.11
NATO6102	63.3	0.10
SHF1002	254.9	2.1
SHF5002[2]	549.3	4.6
SHF6103	249.7	2.1
SHF7102	244.1	2.0
UHF1004[2]	612.5	38.3
UHF7103	177.8	11.1
UHF7304	165.9	10.4
UHF1003	151.4	18.9
UHF5004[2]	589.9	73.7
UHF6101	151.8	19.0
UHF6200	105.4	13.2
UHF6300	109.6	13.7
UHF7302	142.0	17.8

1. Percent of Nominal Link Capacity as per Table 1
2. Designated IS circuit

TABLE 15.

Circuit Type		Label[1]
Point-to-Point	any	HFNA, HFTA, HFEM, NATO
Broadcast	non Designated-IS circuits only	UHF, SHF
	Designated-IS circuits only	UHF_{dr}, SHF_{dr}
	non Designated-IS circuits only, FLTSAT1 and FLTSAT7 respectively	UHF_1, UHF_7
	Designated-IS circuits only, FLTSAT1 and FLTSAT7 respectively	$UHF_{1,dr}$, $UHF_{7,dr}$

1. The subscript *dr* stands for Designated Router (Designated IS).

The results listed in Table 14 have been regrouped by circuit type and plotted as shown in Figure 11. In these figures, vertical columns represent the range of overhead observed over the subnets/circuits of a same type. The circuits are identified by labels as shown in Table 15. A number of observations can be made from these two figures and the data in Table 14.

For the point-to-point circuits:

- the absolute amount of overhead measured over the HF European subnets is at least three times higher than the one measured over the HF North American and Trans Atlantic subnets;
- the amount of overhead measured over the NATO IV subnets is negligible (<<1%) and the overhead over the HF North American and HF Trans Atlantic subnets is acceptable (<3%);

For the broadcast circuits:

- the amount of overhead over non Designated-IS circuits increases with the number of subnet members;
- routers having to handle the pseudonode functions consume up to 3.5 times more overhead bandwidth than those which do not;
- the overhead of the Designated-IS circuits over the UHF Satcom subnets (IS50 at STC and IS10 in Canada) represents an excessive fraction (73% and 38%) of the nominal link capacity of these circuits.

The above observations suggest that the parameters should be readjusted to reduce the amount of overhead generated over:

- the HF European subnet and;
- the UHF Satcom subnets (Designated IS circuits mainly).

These two problems will be handled separately in Sections 8.1 and 8.2 respectively as part of the optimization work described next.

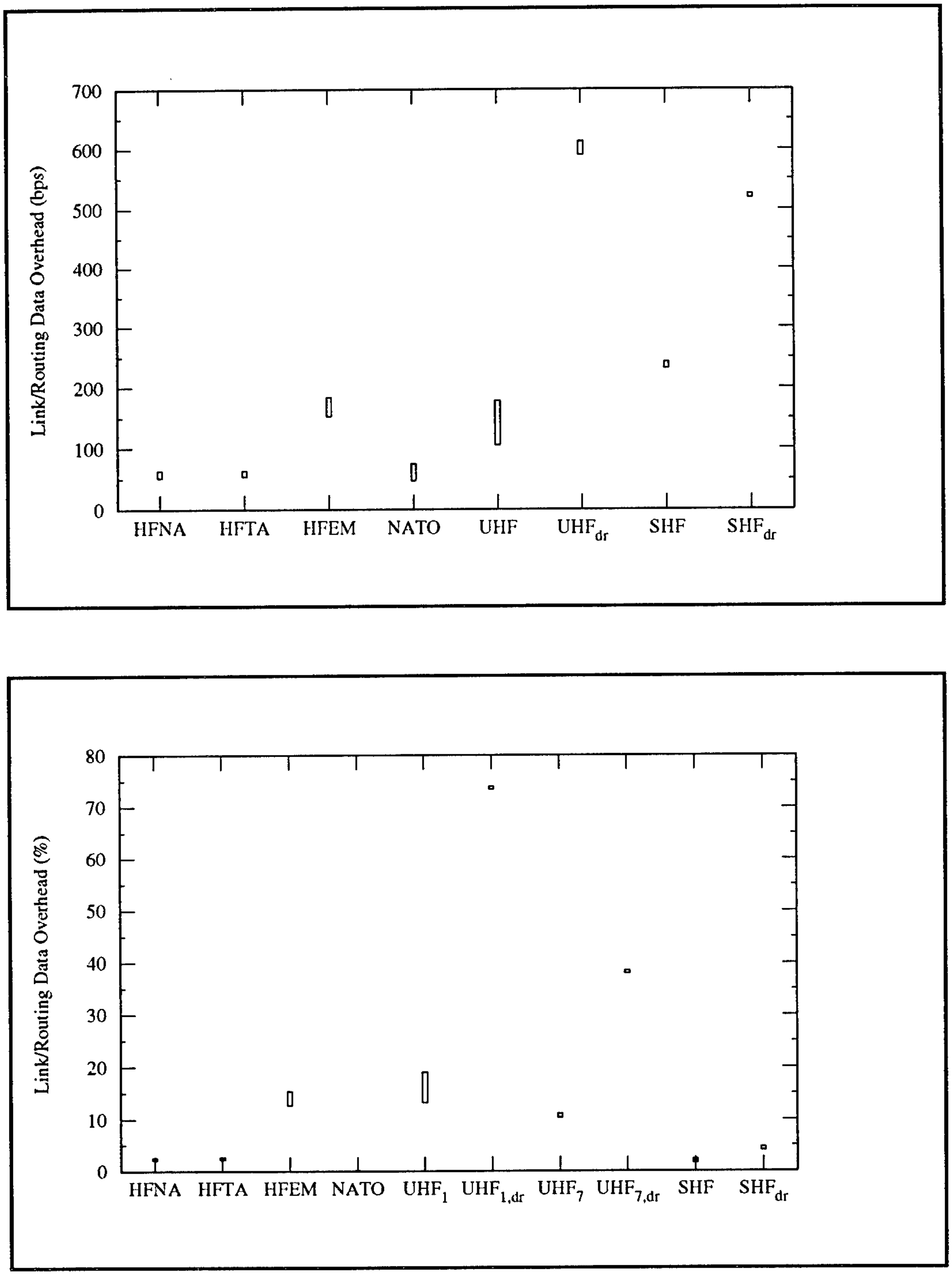


FIGURE 11. Overheads obtained with parameter values listed in Table 13

8.0 Routing Data Overhead Reduction

Routing data overheads can be reduced in several ways. The approach taken here is to avoid making any change that may significantly reduce system performance globally. This selective overhead reduction will be accomplished by adjusting a subset of the ISO 10589 timers. The RMG and SNAC parameters will remain constant. Sections 8.1 and 8.2 identify the timers that are most relevant to a particular circuit type and discuss, subnet by subnet, the main options for reducing the overhead. Specific changes in parameter values are later proposed in Section 8.3.

8.1 Point-to-Point Circuits

Of the eight ISO 10589 timers listed in Table 13, only two are specifically designed to control the amount of routing traffic over point-to-point circuits. These are the *minimumLSPTransmissionInterval* and *PartialSNPInterval*[1] timers. The former handles the retransmission frequency of LSPs on the circuit whereas the latter, the frequency of acknowledgments for the LSPs received. The operation of both timers is inter-related:

1. LSPs sent on a point-to-point circuit must be acknowledged by a Partial Sequence Number PDU; and,
2. a LSP is periodically re-sent until it is acknowledged by a PSNP.

For these reasons, the following relationship can be established between these two timers:

$$PartialSNPInterval = k * minimumLSPTransmissionInterval \tag{4}$$

where k is a constant. Using the default ISO 10589 timer values, k=0.4 (40%). This proportionality constant was preserved throughout this study, except for the cases where the *PartialSNPInterval* timer value would have been less than 1.33 seconds. In such cases, the *PartialSNPInterval* timer was set at 1.33 seconds so as to remain above the 1 second minimum time interval specified by the standard when a 25% timer jitter[2] is applied.

Point-to-point IIH PDUs also contribute to the routing data overhead. The load they impose on the subnet is however on average constant and much more predictable. For example, assuming a typical IIH PDU size of 36 bytes, a timer jitter of 25% and an average link layer protocol overhead of 10%, the overhead can be approximated by:

$$\text{PP IIH PDU Overhead (bps)} \approx \frac{36 \text{ bytes} \times 8 \text{ bits/byte} \times 1.10}{iSISHelloTimer \text{ sec} \times (1 - 0.25/2)} \approx \frac{360}{iSISHelloTimer} \tag{5}$$

For an *iSISHelloTimer* setting of 60 seconds as was used in the previous simulation scenario, the point-to-point IIH PDU overhead would be about equal to 6 bps.

1. This timer is also used over broadcast circuits of non-Designated IS.
2. Routing architectural constant specified by the standard.

8.1.1 HFNA and HFTA Subnets

The link/routing data overhead estimates obtained in Section 7.0 for the HF North American and HF Trans Atlantic subnets do not exceed 3% (72 bps) of the nominal circuit capacity (2400 bps). This result is judged quite acceptable. The amount of overhead could in fact be increased by 0.5% or 1% (12-24 bps) if better IS to IS connectivity (faster detection of adjacency Up/Down events) is desired. This would correspond to tripling (6→18 bps) or quintupling (6→30 bps) the amount of overhead due to the generation of IIH PDUs which in turn can be achieved, according to eq. (5), by reducing the *iSISHelloTimer* setting of these circuits to about 20 or 12 seconds respectively.

8.1.2 HF European Subnet

In this simulation, the HF European subnet is modeled as two static point-to-point half-duplex circuits: the links FR1⇔NL and FR2⇔UK2 (see Figure 12 for a partial map of the topology). The link/routing data overhead measured over these circuits varies between 152 and 184 bps. These figures are three times higher than the ones measured over the HFNA and HFTA circuits, although all circuits were set with the same IS configuration values. To explain such a difference in the results, the simulation was repeated using the original parameter values (Table 13) except for the four ISs (IS21, IS22, IS40 and IS62) of the two HF European links that were configured with higher values of *minimumLSPTransmissionInterval* and *PartialSNPInterval* as summarized in Table 16.

The resulting routing data flow is shown in Figure 13. Link numbers and data flow directions are identified by the source SNAC labels listed in the legend of the upper plot. The overhead of the HF European links remains nearly constant when the *minimumLSPTransmissionInterval* value exceeds 5 seconds. A slight reduction in overhead (~1% of link capacity) could be achieved by increasing the timer to 5 seconds but at the expense of a similar increase in Time-Averaged Divergency.

This result suggests that the link layer protocol itself, not the ISO 10589 timer settings, is mainly responsible for the large link/routing data overhead difference observed between the HF European and HFNA/HFTA subnets. This was confirmed by running the same simulation scenario once more but temporarily configuring the HF European circuits to operate at 1200 bps in full-duplex mode. Overheads similar to those measured over the 2400 bps full-duplex HFNA/HFTA circuits were obtained. The Time-Averaged Divergency of the network did not change significantly.

The *minimumLSPTransmissionInterval* and *PartialSNPInterval* parameters are attributes of the CLNS-MO. The setting of these timers affects all point-to-point circuits that are under control of the same IS. This characteristic creates some routing traffic dependency between the links. Figure 13 shows that this side effect is present on the NATO IV NL⇔STC link although the point-to-point IS timers at STC (IS50) were not modified during this simulation. As indicated by the NATO5000 curve, the link/routing data overhead more than doubles in the STC→NL direction when the NL's *minimumLSPTransmissionInterval* timer is changed from 3 seconds to 30 seconds in an attempt to

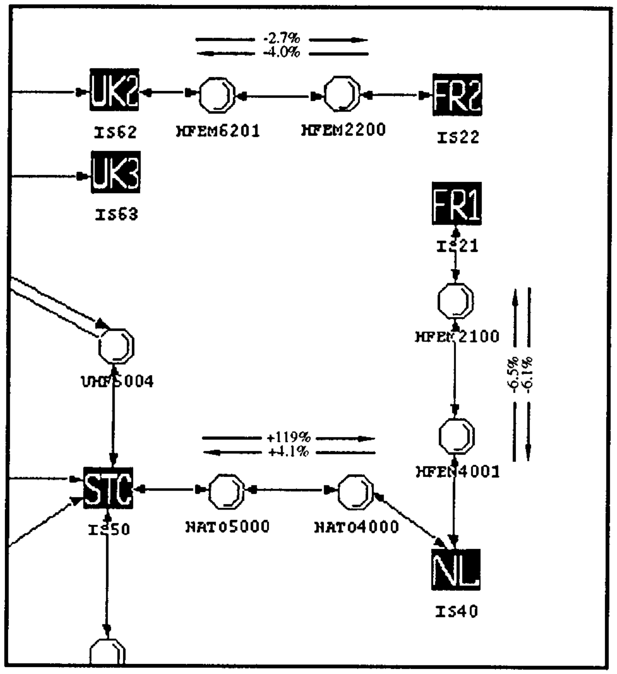


FIGURE 12. Variation in absolute link/routing data overhead when *minimumLSPTransmission-Interval* is increased to 30 seconds from 3 seconds over the HF European circuits.

reduce the overhead on the FR1⇔NL link. This also causes the Time-Averaged Divergency of the network to double. The large overhead increase is caused by IS50 re-sending LSPs at short time intervals (3 seconds) and IS40 waiting up to 0.4×30 = 12 seconds before acknowledging receiverships. Both FR1's and NL's RIBs gradually become time desynchronized with regard to the other RIBs of the network. Figure 13 summarises the overhead changes when the *minimumLSPTransmissionInterval* is increased to 30 seconds from 3 seconds.

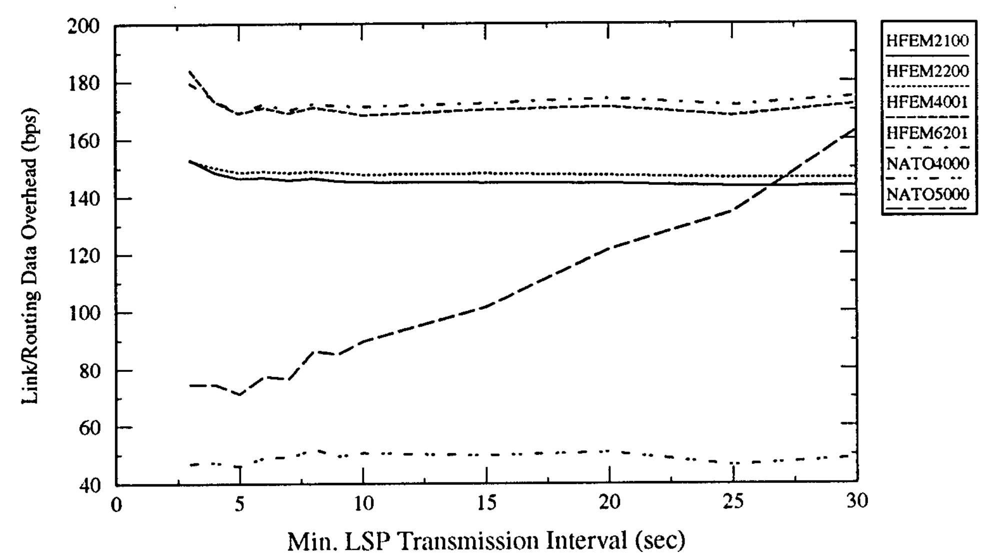
8.1.3 NATO IV Subnets

The amount of link/routing data overhead measured over the NATO IV subnets is relatively negligeable (<<1%). This is due to the large bandwidth of these subnets. No attempt will here be made to try to reduce it further.

TABLE 16. Parameter set "G": Configuration for HF European traffic analysis

PARAMETER		VALUE[1]
IS	Max. LSP Generation Interval (min)	18
	Min. LSP Generation Interval	90
	Min. LSP Transmission Interval	
	- FR1, FR2, NL, UK2	variable[2]
	- other ISs	3
	Min. Broadcast LSP Transmission Interval	0.1
	Complete SNP Transmission Interval	20
	Partial SNP Transmission Interval	
	- FR1, FR2, NL, UK2	variable[2]
	- other ISs	1.33
	IS-IS Hello Transmission Interval	60/180[3]
	DR IS-IS Hello Transmission Interval	30
RMG	Min. Metric Update Interval (min)	10
	Min. Metric Delta (%)	{50,50,∞,∞}
SNAC	Statistics Report Interval	30
	Routing PDU TTL	30
Other	Nominal Load (%)	0

1. All parameters expressed in seconds unless otherwise specified.
2. As per eq. (4) on page 37.



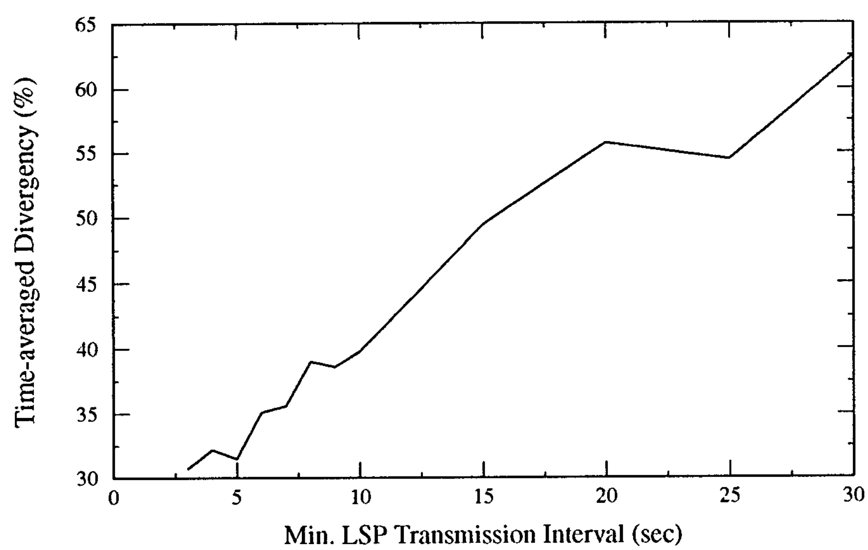


FIGURE 13. Overheads and network Divergency resulting from increase in *minimumLSPTransmissionInterval* timer value of HF European circuits

8.2 Broadcast Circuits

Three ISO 10589 timers are specifically designed to control the amount of routing traffic over broadcast circuits. These are the *minimumBroadcastLSPTransmissionInterval*, the *CompleteSNPInterval* and the *PartialSNPInterval* timers.

The *minimumBroadcastLSPTransmissionInterval* timer ensures a minimum time separation between the PDU arrivals on the LAN. This timer is of little use for the CSNI transit domain circuits since the SNACs do their own PDU queueing and a much different MAC protocol than the ISO 8802 is used by these subnets.

The *CompleteSNPInterval* timer is used by the Designated IS, i.e. by the router on the LAN, that has been elected to carry on the pseudonode duties, in particular that of sending Link State and Hello PDUs on behalf of the LAN. Section 5.2 showed that when there is no change in IS to IS connectivity (adjacency Up/Down events), this timer interval can be increased to reduce the routing data overheads without causing additional delay in network RIB convergence.

The *PartialSNPInterval* timer is needed by the non-Designated ISs of the LAN. On broadcast circuits, there is no explicit acknowledgment sent by the ISs when LSPs are received. While a complete list of the network LSPs is sent periodically by the Designated IS, PSNPs are used by non-Designated ISs to request from the Designated IS the missing or outdated LSPs. Any change of the *PartialSNPInterval* timer configuration must be made by taking into consideration the effects such changes could have on the point-to-point circuits that may be present at the same IS.

Hello PDUs sent over narrow band broadcast circuits can represent a significant percentage of the overall routing data overhead. This is because IIH PDUs are padded to the maximum LSP buffer size for which the circuits are configured. For CSNI, this size has been fixed at 1420 bytes to avoid encapsulation of fragmented NPDUs within IP PDUs.

As for point-to-point circuits, the load imposed by the transmission of IIH PDUs is on average constant and predictable. Assuming a timer jitter of 25% and an average link layer protocol overhead of 5%, the overhead can be approximated by:

$$\text{LAN IIH PDU Overhead (bps)} \approx \frac{1420 \text{ bytes} \times 8 \text{ bits/byte} \times 1.05}{iSISHelloTimer \text{ sec} \times (1 - 0.25/2)} \approx \frac{13700}{iSISHelloTimer} \qquad (6)$$

For example, an *iSISHelloTimer* setting of 60 seconds would produce an overhead of about 228 bps.

In the event that an IS becomes elected to fulfil the pseudonode functions, the load imposed by the transmission of IIH PDUs by the Designated IS is somewhat different. According to Section 10.1 of the standard, jitter must be applied to all periodic timers to avoid the routing traffic becoming synchronised which could cause overloading of both the transmission medium and the systems receiving the PDUs. The standard also states in Section 8.4.4 that at least 1 second must elapse between the transmissions of LAN IIH

PDUs on the same circuit. The recommended default *dRISISHelloTimer* interval being 1 second, the standard states in the footnote of Section 8.4.4 that

> "in this case jitter is not applied, since it would result in intervals of less than one second".

Retix interpreted this statement in a narrow sense and implemented the *dRISISHelloTimer* without any jitter, regardless of the value the timer is set at. Assuming 0% jitter and an average link layer protocol overhead of 5%, the load can be approximated by:

$$\text{DIS IIH PDU Overhead (bps)} \approx \frac{1420 \text{ bytes} \times 8 \text{ bits/byte} \times 1.05}{dRISISHelloTimer \text{ sec}} \approx \frac{12000}{dRISISHelloTimer} \qquad (7)$$

For a *dRISISHelloTimer* setting of 30 seconds as was used in the previous simulation scenario, the IIH PDU overhead over a broadcast circuit handling the Designated IS functions would be about equal to 400 bps.

8.2.1 UHF Satcom Circuits

In the previous simulation scenario, the *iSISHelloTimer* of the UHF circuits was set at 180 seconds, a time interval 60 times higher than the default value recommended by the standard. According to eq. (6), the broadcast IIH PDU overhead amounts to 76 bps i.e. about 4.8% of the link capacity available per member of the FLTSAT7 subnet and 9.5% for the FLTSAT1 subnet. These figures, despite being only estimates, indicate clearly that the mandatory padding of the broadcast Hello PDUs and default *iSISHelloTimer* rate are very expensive in terms of:

1. bandwidth; and,
2. link delay[1].

The routing data overhead is even higher on the Designated IS circuits. One reason is of course the relatively short *dRISISHelloTimer* interval (30 seconds) that was selected for the pseudonodes. Another reason is the transmission of CSNPs by the Designated IS. In the previous simulation scenario, the *CompleteSNPInterval* timer was set at 20 seconds. It is worth mentioning that having *dRISISHelloTimer* > *CompleteSNPInterval* challenges the whole purpose of sending (DR-) IIH PDUs at all. One would think that if the Designated IS is able to send SNPs faster than it can say Hello then it should be alive! For comparison purposes, the default standard values for these two timers would normally cause 10 Hellos to be sent before a full Sequence Number report is released by the Designated IS. As far as the broadcast circuits are concerned, these figures illustrate that the protocol is being "stretched" to perform outside the large bandwidth environment for which it was originally conceived.

The routing data overhead due to IIH PDUs was further investigated by simulation by varying the *CompleteSNPInterval* or the *dRISISHelloTimer* parameters while the other

1. At nominal link capacity (4800 bps), every LAN IIH PDU ties up the subnet for more than 2.4 seconds.

parameters were kept constant as listed in Table 17. The results are plotted as shown in Figure 14. Increasing the time interval of either one of these two timers causes the overhead to decrease monotonically. Significant overhead reduction (up to 50%) is achievable by reducing (quadrupling) the IIH PDU generation rate of the Designated ISs. The gain achieved by increasing the CSNP transmission interval is much less because of the small size of the Sequence Number PDUs. Depending on the amount of traffic foreseen for the network, it may be advisable to trade some of the adjacency Up/Down notification speed for higher data throughput capability. Note that neither one of these two timers directly affects the convergence time of the network as long as the adjacencies remain Up or stay Down. For example, the Time-Averaged Divergency during this last simulation fluctuated by less than +0/-3%.

8.2.2 SHF Circuits

The link/routing data overhead estimate listed in Table 14 for the Designated IS circuit of the DSCS III subnet represents 4.6% (550 bps) of the nominal circuit capacity (12000 bps). The overhead does not exceed 2.1% on the other IS circuits of the subnet. These figures are quite small compared to those measured for the UHF Satcom subnets and may be found quite acceptable in many network operational environments. There is one problem though. In this simulation, the router at STC (IS50) has been elected Designated IS on both subnets, DSCS III and FLTSAT1. This means that both Designated IS circuits must share the same CLNS-MO attributes, in particular the timer settings for the *CompleteSNPInterval*, the *PartialSNPInterval* and the *dRISISHelloTimer*. As pointed out in Section 8.1.2, this characteristic of the protocol creates some routing traffic dependency between the links and some of the freedom necessary for optimizing the routing traffic on a per link basis is lost. In this case, the options are:

1. have the router at STC resign its Designated IS function on one of the two LANs and, optimize the two circuits independently;
2. have the router at STC keep its Designated IS function on both LANs but find a timer setting that will accommodate more or less, both circuits.

Data are produced in the next section for the first option only.

TABLE 17. Parameter set "H": Broadcast Subnets

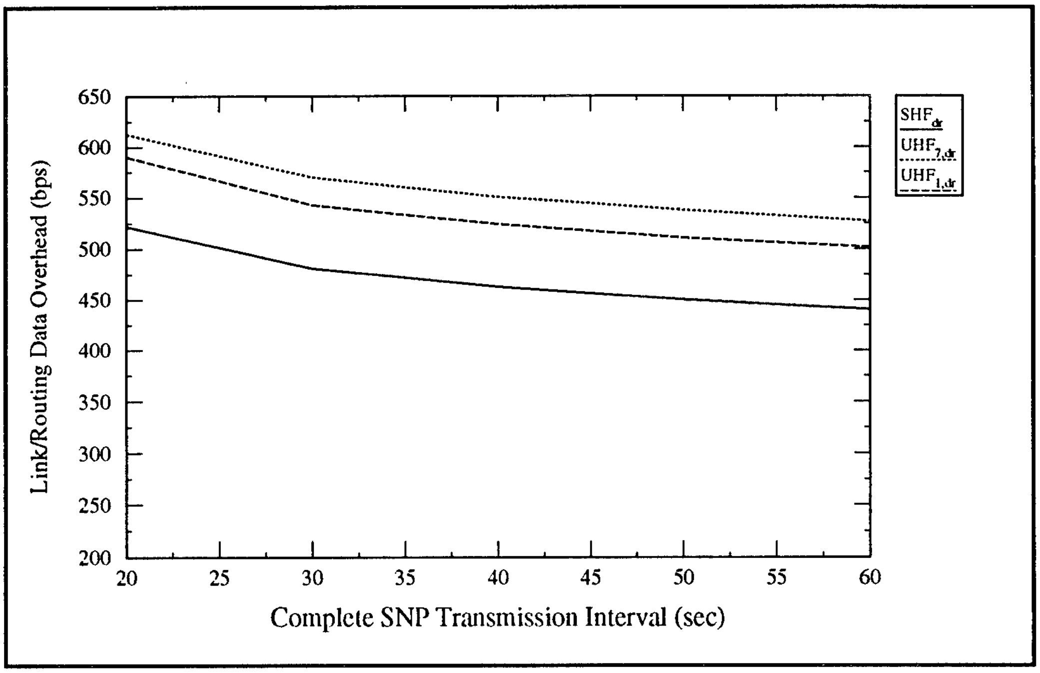
	PARAMETER	VALUE[1]
IS	Max. LSP Generation Interval (min)	18
	Min. LSP Generation Interval	90
	Min. LSP Transmission Interval	3
	Min. Broadcast LSP Transmission Interval	0.1
	Complete SNP Transmission Interval	variable[2]
	Partial SNP Transmission Interval	1.33
	IS-IS Hello Transmission Interval	60/180[3]
	DR IS-IS Hello Transmission Interval	variable[4]
RMG	Min. Metric Update Interval (min)	10
	Min. Metric Delta (%)	$\{50,50,\infty,\infty\}$
SNAC	Statistics Report Interval	30
	Routing PDU TTL	30
Other	Nominal Load (%)	0

1. All parameters expressed in seconds unless otherwise specified.

2. The DR IS-IS Hello Transmission Interval was set at 30 seconds while the Complete SNP Transmission Interval was increased by steps.

3. UHF Satcom circuits only.

4. The Complete SNP Transmission Interval timer was set at 20 seconds while the DR IS-IS Hello Transmission Interval was increased by steps.



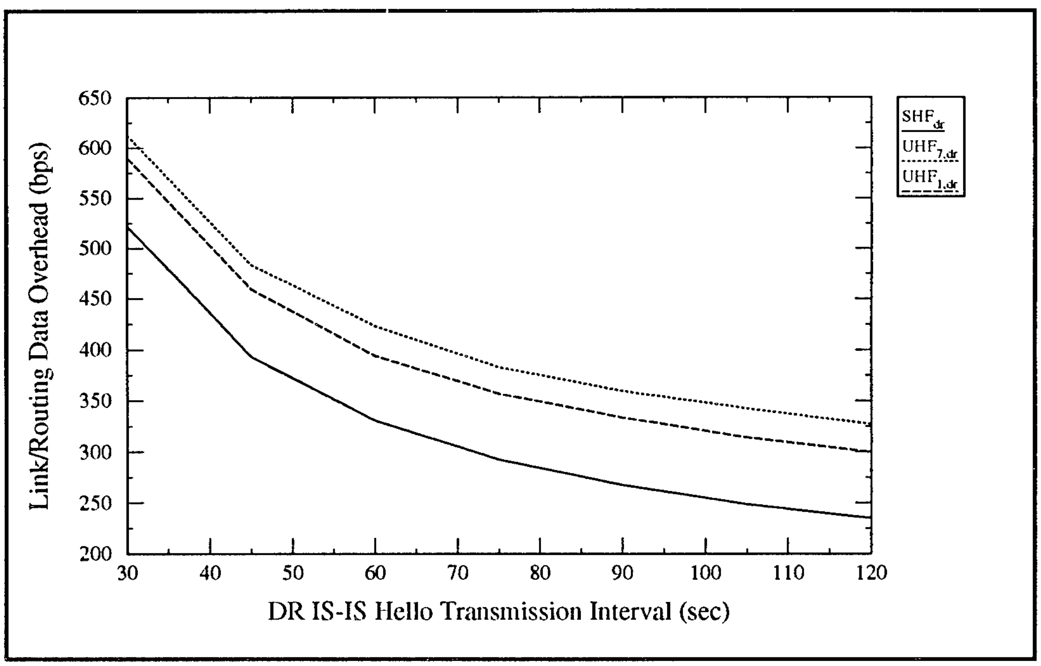


FIGURE 14. Overheads measured over the DIS circuits of the broadcast subnets

8.3 Specific Changes

In Sections 8.1 and 8.2, a number of alternatives have been identified to reduce the IS-IS data overhead. In this section, specific changes in parameter values are proposed.

A summary of the revised parameter assignments is given in Table 18. As before, shaded areas identify changes relative to a configuration previously described, in this case the configuration used at the start of the overhead analysis (parameter set "F", Table 13).

In this simulation scenario, no IS is allowed to perform the Designated IS function on more than one subnet. Based on the data of Table 3 in Section 5.2, the router at NRaD (IS71) was elected to perform the pseudonode function the router at STC (IS50) was previously performing over the DSCS III subnet. This allows more flexibility in adjusting the routing traffic overhead.

The bulk of the overhead reduction is obtained by increasing the IIH transmission interval of the Designated ISs. This obviously slows down the adjacency Up/Down notification speed. As far as the Level 2 IS to IS connectivity is concerned, this approach is tolerable since the routers of the CSNI transit domain are likely to remain in the Up state at all times. The same approach is however less than desirable when the Level 1 ES to IS connectivity is taken into consideration. Most routers of the transit domain also manage Level 2 and Level 1 LLC circuits where several End Systems are likely to be used sporadically for transmitting and receiving user data. When low bandwidth broadcast subnets are involved, the ISO 10589 protocol forces one to trade some of the connectivity-reporting speed for higher throughput capability.

The performance obtained when the simulator is run with the values listed in Table 18 is shown in Table 19. The overhead reduction is most significant over the UHF Designated IS circuits (reduction of 17.3% and 47.7%) although the overhead of these circuits still remains high (21% and 26% of nominal link capacity). Overall, 18.2% of the FLTSAT 1 subnet capacity is needed to convey the routing data information of its 6 routers. On the FLTSAT 7 subnet, 14.4% of the available capacity is used by 3 routers. These numbers are considerably higher than the overheads obtained for other subnets, except the HF European subnet (13.4%).

TABLE 18. Parameter set "I": Overhead reduction

	PARAMETER	VALUE[1] F	VALUE[1] I
IS	Max. LSP Generation Interval (min)	18	18
	Min. LSP Generation Interval	90	90
	Min. LSP Transmission Interval		
	- FR1, FR2, NL, UK2	3	5
	- other ISs	3	3
	Min. Broadcast LSP Transmission Interval	0.1	0.1
	Complete SNP Transmission Interval		
	- CA[2], US1[3]	20	30
	- STC[4]	20	40
	- other ISs	20	20
	Partial SNP Transmission Interval		
	- FR1, FR2, NL, UK2	1.33	2
	- other ISs	1.33	1.33
	IS-IS Hello Transmission Interval		
	- UHF Satcom circuits	180	180
	- other circuits	60	60
	DR IS-IS Hello Transmission Interval		
	- CA[2], US1[3]	30	80
	- STC[4]	30	180
RMG	Min. Metric Update Interval (min)	10	10
	Min. Metric Delta (%)	{50,50,∞,∞}	{50,50,∞,∞}
SNAC	Statistics Report Interval	30	30
	Routing PDU TTL	30	30
Other	Nominal Load (%)	0	0

1. All parameters expressed in seconds unless otherwise specified.
2. IS elected on the FLTSAT7 subnet
3. IS elected on the DSCS III subnet
4. IS elected on the FLTSAT1 subnet

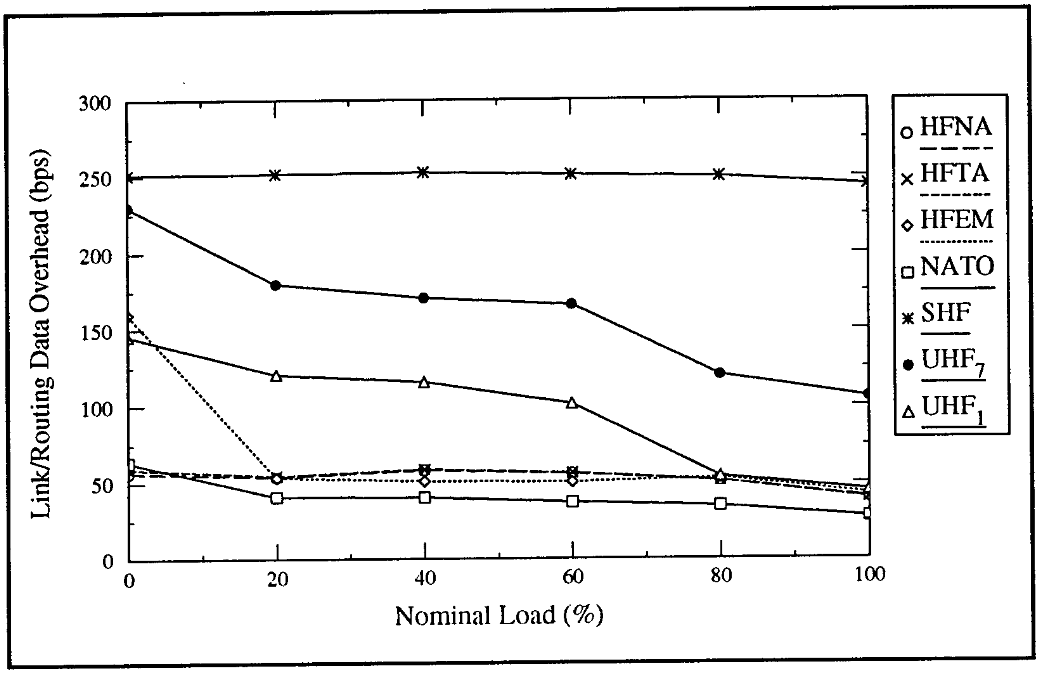
TABLE 19. Overheads after reduction (|change| ≥ 0.5% highlighted)

Source SNAC	Link/Routing Data Overhead (bps)	(%)[2]	Change[1] (%)
HFNA1000	56.4	2.4	-
HFNA1001	59.6	2.5	-0.1
HFNA7100	61.7	2.6	-
HFNA7101	54.3	2.3	+0.1
HFNA7300	57.0	2.4	+0.1
HFNA7301	51.2	2.1	-0.1
HFTA1005	58.5	2.4	+0.1
HFTA5003	56.5	2.4	-
HFTA6100	64.7	2.7	-
HFTA7303	58.0	2.4	+0.1
HFEM2100	146.9	12.2	-0.5
HFEM2200	149.5	12.5	-0.2
HFEM4001	171.2	14.3	-1.0
HFEM6201	172.5	14.4	-0.6
NATO3000	59.3	0.09	-
NATO3001	57.8	0.09	-0.01
NATO4000	47.5	0.07	-
NATO5000	75.4	0.12	-
NATO5001	73.1	0.11	-
NATO6102	69.7	0.10	+0.01
SHF1002	255.3	2.1	-
SHF5002[3]	249.0	2.1	-2.5
SHF6103	248.6	2.1	-
SHF7102[4]	252.4	2.1	+0.1
UHF1004[4]	336.4	21.0	-17.3
UHF7103	181.0	11.3	+0.2
UHF7304	173.2	10.8	+0.4
UHF1003	152.0	19.0	+0.1
UHF5004[4]	208.3	26.0	-47.7
UHF6101	154.2	19.3	+0.3
UHF6200	108.3	13.5	+0.3
UHF6300	110.0	13.8	+0.1
UHF7302	141.6	17.7	-0.1

1. Data in adjacent column minus data in last column of Table 14.
2. Percent of Nominal Link Capacity as per Table 1
3. Was the DSCS III Designated IS circuit in previous configurations
4. Designated IS circuit

Figure 15 (top plot) provides some insight into how the link/routing data overhead varies when the network is gradually loaded with user data. For the full-duplex subnets, the general trend is that the overhead remains nearly constant regardless of the network loading. For the half-duplex subnets (HFEM and UHF Satcom) the results are rather different. The overhead registered over the HFEM circuits reaches a level comparable to what is measured over the HFNA/HFTA circuits once the loading exceeds 20%. Under light circuit loading (<20%), the simulated link layer protocol itself introduces significantly more overhead than the routing data generated by the router. This characteristic is also present although to a lesser extent, over the UHF Satcom subnets. For both of them, there is a reduction of the overhead when the load increases from 0% to 20%. Thereafter, the overhead remains nearly constant until the load reaches 60%. Past this level, the subnets become overloaded and unable to forward in a timely manner all the routing data that are generated by the routers. This has a striking effect on the system stability as shown by the bottom plot of Figure 15. When the nominal load reaches 60%, the Time-Averaged Divergency goes up abruptly. It has nearly doubled by the time the load reaches 80%. At this level, the RIBs are de-synchronized for as much as three quarters of the time. The reason why the divergency falls back when the load is increased from 80% to 100% will be given in a moment.

To wrap-up this analysis and verify that the changes proposed under this section are in-line with the network stability analysis described in Part-I of this study, the simulation scenario of Table 18 was repeated for different values of *minimumMetricUpdateInterval* and *nominalLoad*. The results are summarized in Figure 16. The network routing stability is maintained over a loading range of 0% to 100% and a range of metric generation rates of 3 to 15 minutes. When the load is less than 60%, the Time-Averaged Divergency is almost linearly related to the *minimumMetricUpdateInterval*. Above 60%, the Divergency tends to level out, regardless of the metric generation rate. This is due to an over-saturation of the UHF subnets. The mean Time-Averaged Divergency over the metric generation rate range calculated from Figure 16 is 85.1% when the load is at 80% and 76.9% when the load is at 100%. As pointed out above, the Divergency reaches a peak then falls back when the load is increased from 60% to 100%. This result can best be explained with the help of Table 20. Over a simulation interval of 5000 seconds, up to 28 metric changes can occur over each link if the RMGs are configured to accept a change in QoS every 3 minutes. The table lists how many of these changes actually occurred over each circuit type. For example, when the load was at 0%, 25, 26 or 27 metric changes were recorded over each one of the HFNA circuits. The reduction in Divergency is caused by the large reduction in the number of metric (QoS) changes when the load approaches 100%. The SNAC queues become full, the transmitters are always busy and the metric values remain unchanged. The only exception to this is the NATO circuits which record 18 to 26 metric changes when the load is at 100%. This can be attributed to the Source Quench mechanism which, although brief (SQ time-out set at 1 sec) and nearly inoperative most of the time (SQ scale set a 0.1), still allows enough draining of the NATO SNAC queues to avoid over-saturation of these relatively high transmission rate subnets.



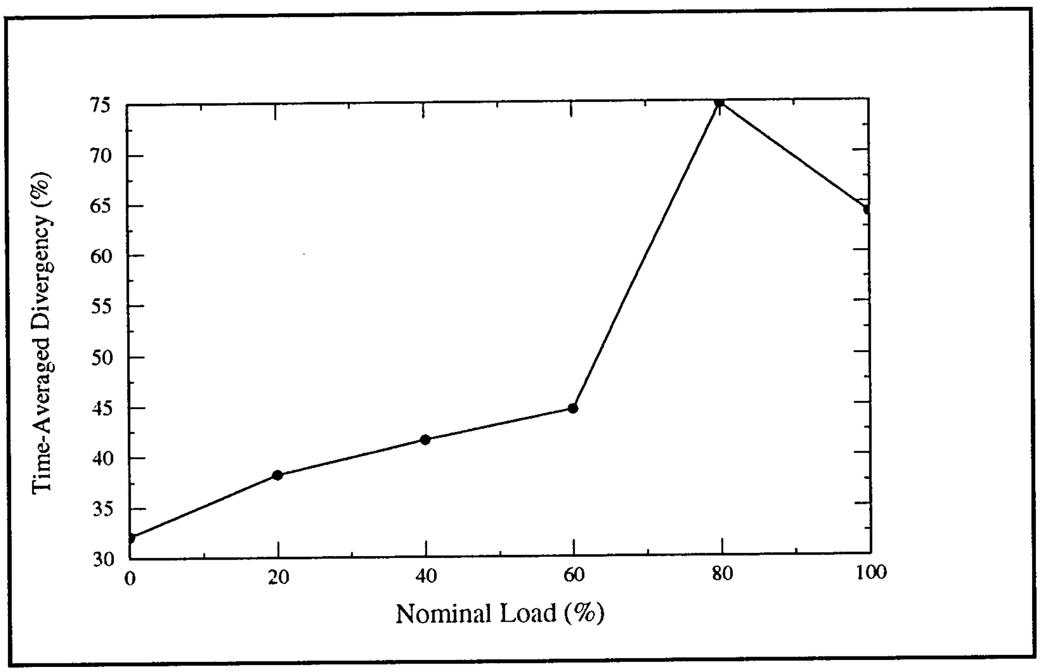


FIGURE 15. Overhead and network Divergency resulting from increase in network loading

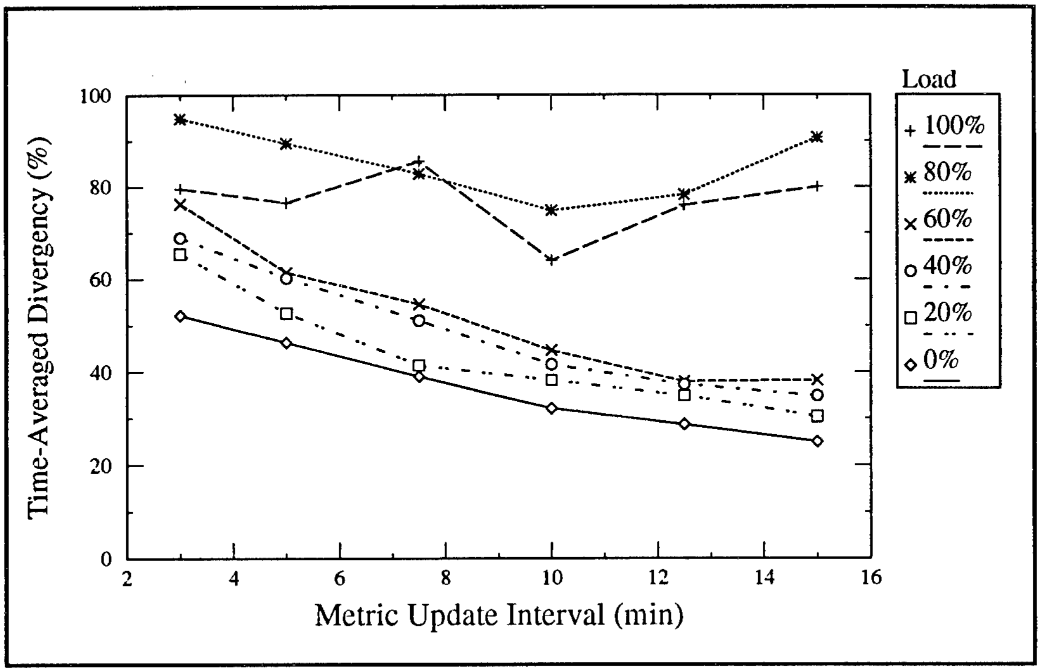


FIGURE 16. Network Divergency achieved with selected QoS-driven metric generation rates

TABLE 20. Number of metric updates produced per circuit over a 5000 second simulation interval and for a *minimumMetricUpdateInterval* set at 3 minutes

Load (%)	HFNA	HFTA	HFEM	NATO	SHF	UHF_7	UHF_1
0	25-27	23-27	5,6,18,19	1	1-2	27-28	26-28
20	25-27	24-27	26-27	1	1	27-28	27-28
40	23-27	24-27	26-28	1	1	27-28	23-27
60	26-27	26-27	26-28	1	1	24-26	2-7
80	24-28	27-28	28,27,2,2	1	3	1-2	1
100	1-2	1-2	5,3,3,1	18-26	1	1-2	1

The values in the two last columns of Table 20 suggest further that the FLTSAT 7 subnet is already saturated when the load is at 80% and that the FLTSAT 1 subnet is about to reach saturation when the load is at 60%. Such significant reduction in available throughput is caused by the relatively large routing data overhead of these subnets. Finally, the numbers in Table 20 indicate also that the SHF subnet metric does not appear to undergo much change. This is only so because the sampling of the network load is too thin to give a complete picture of what is really happening over this subnet. It is worth noting that, in general, the metrics change nearly as often as the RMGs allow them to change or else they rarely change at all.

PART - III

9.0 Summary

This study has identified several important characteristics of the ISO 10589 protocol that may negatively impact the effectiveness of a routing scheme based on QoS. In fact, this study shows that the protocol must be "stretched" to perform over the low bandwidth broadcast subnet environment, like for example, the CSNI UHF Satcom subnets. This should not be interpreted as being a negative statement about the protocol or the QoS-driven routing scheme. In several aspects, the protocol is being asked to perform outside the functional environment that it is designed for or that it can currently support. On the other hand, the CSNI QoS-driven routing scheme cannot produce demonstrable effectiveness that is any better than the limits imposed by the protocol.

Two characteristics of the ISO 10589 protocol that are particularly relevant to this study and that can help one appreciate the context in which the CSNI QoS-driven routing scheme is being implemented and tested are:

1. Adaptability and Stability — The standard explicitly states in Section 6.6.1 that the protocol does not adapt to traffic changes or traffic history. It does not automatically modify routes based on global traffic load. It stabilizes in finite time to "good routes", provided no continuous topological changes occur. The period of adaptation to topological changes in the domain is a function of the domain diameter and data link speeds. Such statements clearly indicate that dynamic routing as conceived by CSNI or otherwise, was not within the design scope of the IS-IS protocol. The protocol was designed for static or quasi-static (QoS) routing.

2. Efficiency — When used over broadcast subnetworks, the protocol is intended to operate on ISO 8802 LANs or other similar high bandwidth subnets. The standard (Section 6.5.1) recognizes that there are broadcast subnetworks that may not be adequately covered at this time; the low bandwidth broadcast subnets used in CSNI probably fall into this category.

This study confirms indeed that the current version of the ISO 10589 protocol imposes severe limitations viz. demonstrating the usefulness of a dynamic routing scheme. Other limitations relating to the metric generation/update process itself were also identified in the body of this document. These limitations will now be succinctly reviewed.

- Most default ISO 10589 timer values are inappropriate for the CSNI network environment. These values would cause very poor throughput performance and lead to network instability.
- The routing functions provided in the protocol for the ISO 8802 LANs do not adequately cover the CSNI broadcast subnetwork requirements, e.g.

- The *minimumBroadcastLSPTransmissionInterval* timer is adequate for the Level 1 LLC circuits but inadequate for the Level 2 SNAC circuits of the transit domain;

- The LAN IIH PDUs are unnecessarily and mandatorily padded.

- In an effort to reduce the overhead over the UHF broadcast subnets, the *dRISISHelloTimer* interval must be made larger than the *CompleteSNPInterval*. Such a timer setting challenges the whole purpose and concept of sending DIS IIH PDUs at all over these subnets.

- The LAN IIH PDU size constitutes excessive routing data overheads for the low bandwidth broadcast subnets (UHF Satcom) of the transit domain. Even though the protocol is able to accommodate increased routing functions, this capability was not exploited by the CSNI network designer. For example, the standard defines only 9 PDU types, 3 of which are IIH PDUs. A non-standard, non-padded LAN IIH PDU type could have been defined to reduce routing data overhead and improve connectivity-reporting rate over these subnets.

  Despite the timer adjustments performed to reduce the amount of overhead, the FLTSAT 1 subnet could not be tuned to obtain satisfactory throughput performance and a reasonable connectivity-reporting rate. This subnet becomes overloaded when the user's throughput reaches 60% of the nominal link capacity (480 bps) and the connectivity-report rate is around 3 minutes.

- The protocol has an intrinsic LSP regeneration cycle of 15 minutes (20 min. max.). This cycle is asynchronous which means that it is not required to synchronise the regeneration of the individual LSPs. In a dynamic routing environment, this protocol characteristic provides a Time-Averaged Divergency floor which cannot be reduced. This floor rises as the network load is increased.

- Most ISO 10589 timers are attributes of the CLNS-MO. This is a major constraint since heterogeneous subnets require different timer settings for optimum network performance. Not having the capability to adjust some of the ISO 10589 timers on a per link (circuit) basis causes:

  1. some routing traffic dependency between the links,
  2. difficulty in adjusting timers to suit all links,
  3. sub-optimal routing data overhead control.

- Greater flexibility in adjusting the protocol timers is obtained when no IS is allowed to perform the Designated IS functions on more than one subnet of the transit domain.

- It does not significantly matter which IS among the members of a LAN is elected to perform the pseudonode duties of the LAN.

- For a network of the size of the CSNI testbed, this study shows that it is possible to maintain the network routing stability over a loading range of 0 to 100% and a range of metric generation rates of 3 to 15 minutes. For much larger networks, it is likely that longer and more frequent transient periods may put these networks in a constant state of sub-optimal routing.
- Results show that it is preferable to discard routing information than to send it if it is outdated. This can be achieved by adjusting the routing PDU Time-to-Live timer of the SNACs. The gain made in doing so reduces the Time-Averaged Divergency significantly when the network load is high (60%).
- Time synchronization of the metric updates at the RMGs can significantly reduce the time spent by the network to adapt to QoS-based routing changes.
- Because of the limitations of the protocol, a trade-off must be made between the system throughput capability and the system connectivity-reporting speed. This constraint is particularly severe over the UHF Satcom subnets.
- Routers having to handle the pseudonode functions on behalf of a LAN can consume significantly more overhead bandwidth than those which do not. This overhead can be reduced by:

  1. reducing the connectivity-reporting speed (reducing the DIS IIH PDU transmission rate), and;
  2. increasing the network convergence time when a new IS is joining the net (reducing the CSNP transmission rate).

  This approach requires that a specific IS be "pre-elected" on a LAN by assigning it a higher LAN Level 2 DIS priority value than the one assigned to any other IS members of the LAN. Note that the adaptability of the protocol to topological changes is reduced by doing so.

  When the link protocol is based on bandwidth reservation such as the UHF Satcom protocol, a greater share of the bandwidth could also be assigned on demand to the Designated IS of the LAN. This aspect has not been considered in this study.
- Even after applying the measures enumerated in the previous paragraph, the overhead of the Designated IS circuits over the UHF Satcom subnets still represent an excessive fraction (21% and 26%) of the nominal link capacity of these circuits.
- The routing functions provided in the protocol for point-to-point subnetworks are adequate and cover well the CSNI HFNA/HFTA/NATO link requirements. The routing data overhead over these links is less than 3%.
- When a subnetwork becomes overloaded, the network stability can be seriously compromised. The performance of the QoS-driven routing scheme can be significantly altered during such periods and is difficult to assess even by simulation.

- There is a reduction in the number of metric (QoS) changes generated by the RMGs when the network load approaches 100%. This tends to reduce the routing traffic overhead and improve the Time-Averaged Divergency. On the other hand, if a subnetwork becomes saturated well before the others then this network may negatively impact the routing adaptation speed of the whole network.
- The functions currently used for calculating and updating the metrics become less appropriate as the metric update interval exceeds the SNAC queue draining time. Further work is needed to determine whether there is a benefit in attempting to predict metric values rather than relying on short-term statistics.
- Considering that relatively long QoS update intervals (3 minutes or more) are required to maintain network stability, a Statistics Report Interval of 30 seconds was judged to be adequate and used throughout this study. This timer setting provides long enough time intervals to gather meaningful statistical information but is yet not too long to avoid dragging too much outdated historical data.

10.0 Recommendations

QoS-based dynamic routing over small network topologies is most effective when the link costs can be updated frequently i.e. usually every 60 seconds at most, depending on the SNAC buffer size and maximum (tolerable) link delay. As the update interval is increased, the QoS-based routing performs more and more like any conventional static routing scheme. Unfortunately, this report shows that for a network of the size of the CSNI testbed, a QoS update interval of 3 minutes stresses the routing adaptation capability of the protocol.

Several measures for substantially reducing the routing stability and overhead problems observed in this study have been mentioned in the previous sections. These include:

1. Parameter Optimization — Because of the limitations of the protocol, a trade-off must be made between the system throughput capability and the system connectivity-reporting speed. The IS, RMG and SNAC parameter values recommended for implementing and testing a full deployment of the CSNI testbed are listed in the last column of Table 18 in Section 8.3.
2. Metric Synchronization — Time synchronization of the metric updates at the RMGs can significantly reduce the time spent by the network to adapt to QoS changes. It would be desirable to modify the RMG metric generation mechanism to synchronize the metric updates on a local clock tick of *minimumMetricUpdateInterval* seconds.
3. IIH PDU Padding Removal — The LAN IIH PDU size constitutes excessive routing data overheads for the low bandwidth broadcast subnets. The protocol should be amended to operate without the padding.

11.0 Conclusion

This study has identified several limitations imposed by the IS-IS protocol when the RIBs are modified to adapt to QoS changes of the subnets. It has determined the amount of time the network requires to adapt to each QoS change and reach the state where all RIBs are synchronized. It has estimated the amount of routing data overhead that can be expected over the CSNI links. It has investigated how to set the ISO 10589 protocol timers for best controlling the rate of link state advertisements within the CSNI network and to minimize the routing data overhead over the links.

Overall, this simulation study shows that the April 1992 edition of the ISO 10589 protocol is not well suited to a dynamic routing and low-bandwidth broadcast subnetwork environment. In several aspects, the protocol is being asked to perform outside the functional environment that it is designed for or that it can currently support. It is doubtful that the QoS-driven routing concept could successfully be applied to heterogeneous radio networks exceeding several times the size of the CSNI demonstrator without making any change to the protocol and to the QoS-driven routing concept itself.

12.0 References

[1] ISO/IEC 10589 :1992 (E), "*Information technology — Telecommunications and information exchange between systems — Intermediate system to Intermediate system intra-domain routeing information exchange protocol for use in conjunction with the protocol for providing the connectionless-mode Network Service (ISO 8473)*", 1st Edition, 30 April 1992

[2] Retix, "*OSI Route Software Reference Manual*", PN 1080261-00-B, Santa Monica, CA, April 1993

[3] STC SP-12, "*Proceedings of the Communications Systems Network Interoperability (CSNI) Symposium on Communications Internetworking*", Shape Technical Centre File Ref: 9980, The Hague, Sept. 1995

This page is left blank intentionally

ANNEX A — Link QoS Control Loop

This annex gives an overview of how the CSNI QoS-driven dynamic metric generation mechanisms relate to the routing information database (RIB) update mechanisms provided by the ISO 10589 standard.

A.1.0 Closed-Loop Control System

The QoS-driven routing scheme is a non-linear closed-loop (feedback) control system[1] (see Figure 17). In the body of this report, the performance of such system is analysed by simulation rather than by analytical means. The link QoS control loop is composed of three sub-systems: a SNAC, a RMG and an IS. By their combined actions, the system attempts to maintain the highest possible Quality of Service and prevent to a lesser extent, overload of the subnet. If for example, the free link capacity goes down because of a sudden increase in user traffic, the RMG produces a new metric value to indicate that from now on a lower capacity is available on this circuit. The IS advertises this information by inserting the link metrics into a Link State PDU and flooding this LSP on the local circuits, including the SNAC circuit where the QoS change originated. The IS, like every other IS receiving this new Link State information, also recalculates the Shortest-Path-First (SPF) tree so as to redirect the user traffic towards better circuits or paths if necessary. These two actions will normally cause a reduction in the net amount of traffic forwarded through the original SNAC circuit and gradually reestablish the QoS of the link. Obviously, if a SNAC is on a path to a destination where there is no alternate route then the IS has no choice but to forward the traffic for that destination through that SNAC. In such a case, if the user traffic was to become excessive, the QoS-driven routing feedback mechanism would fail to reduce the circuit load, even though a low QoS value would be advertised for the link. The Congestion Control and Congestion Avoidance mechanisms should then, as they would normally do at every other time, prevent the circuit from being overloaded.

As for the Congestion Control scheme, the loop response time of the QoS-driven routing scheme must be controlled carefully to achieve the best system performance. Each sub-system in Figure 17 contains a number of parameters that can be adjusted for this purpose. The simulation models developed for the CSNI project include most of them and these are taken into consideration in this study. The reader is referred to [3] for a description of the RMG design or discussion on QoS measurement issues. The remainder of this annex gives an overview of the timing (response time) constraints relating to the QoS-driven routing control loop.

1. A topology typically contains one loop for every SNAC circuit.

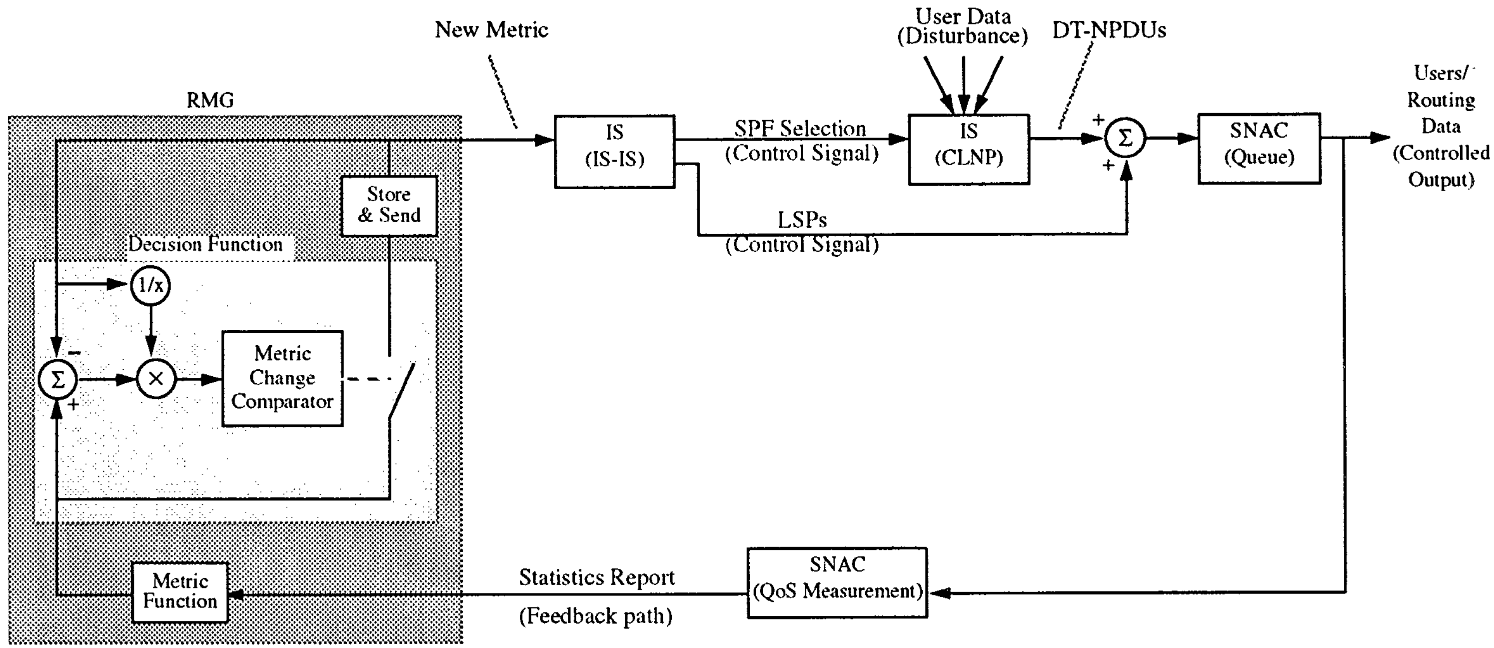


FIGURE 17. Closed-loop block diagram of QoS-driven routing scheme

A.2.0 Loop Response Time

A number of timers regulate how fast the control system reacts to a change in QoS. The first of these is the Statistics Report Interval timer located on the feedback path of the loop. Every SNAC periodically generates a report indicating the currently available QoS. The reports issued from different SNACs (circuits) and delivered to a same RMG are asynchronously received in time (Figure 18). QoS data are converted to metric values which are compared with the metric values last sent to the router. Whenever the difference exceeds some user-specified threshold, the new metrics are forwarded to the router unless the previous metric update is less than *minimumMetricUpdateInterval* seconds old. In such a case, the metric update is delayed until the time interval set for the timer has elapsed.

A metric update issued by a RMG to an IS does not cause an immediate Link State change remotely, or even locally. The ISO 10589 protocol handles the QoS change in two steps:

1. Generation (or re-generation) of a LSP;
2. Transmission of a LSP.

The *minimumLSPGenerationInterval* timer holds down any metric change i.e. does not update a source LSP, unless the previous change made to that LSP is older than the time interval set for this timer. So, once the timer has elapsed, or later if the RMG has not yet reported any metric change, a LSP is regenerated by inserting into it the new metric values and assigning this LSP a new Sequence Number. At this time, the *minimumLSP-GenerationInterval* timer is also re-initialized to limit the rate of regeneration of this LSP. A Send Routing Message (SRM) flag is also set to initialize the second step of the QoS update i.e. the transmission of that LSP.

Once a LSP has been regenerated, the router handling the SNAC attachment where the QoS change occurred contains in its RIB a LSP with a Sequence Number and Link State information that must then be made known to all ISs of the network for proper link quality routing. The flooding process requires that a copy of the newly generated LSP be transmitted on every local circuit that has at least one adjacency in the Up state. On point-to-point circuits, the transmission occurs when the periodic *minimumLSPTransmissionInterval* timer has expired. The LSP is re-flooded until acknowledged by the neighbour IS. On broadcast circuits, the LSP ID number must first be drawn from a lottery selection process amongst all LSPs needed to be transmitted, before it can be transmitted. The lottery draw is periodic and repeated every *minimumBroadcastLSPTransmissionInterval* seconds.

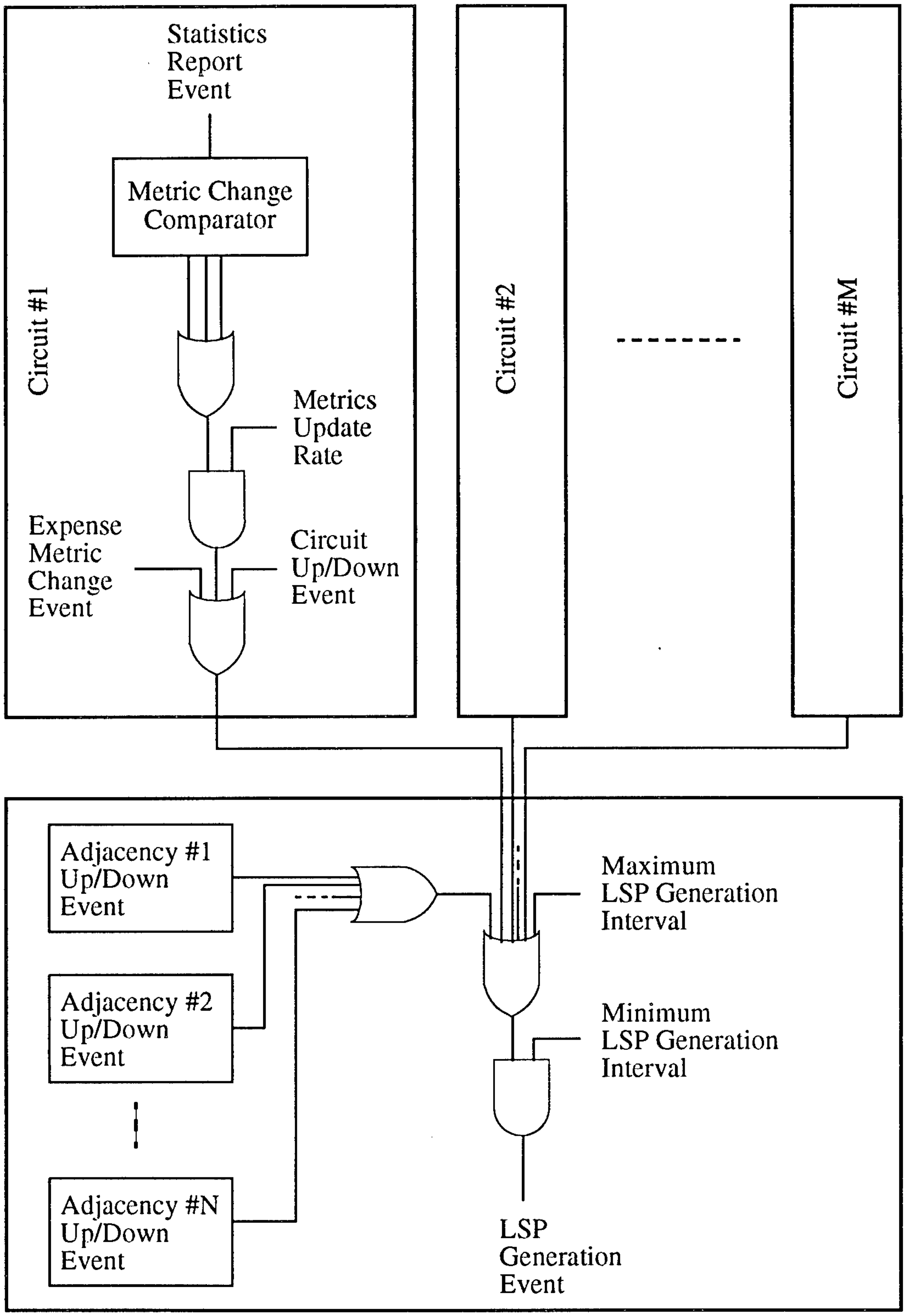


FIGURE 18. Simplified representation of main mechanisms leading to the generation of a LSP

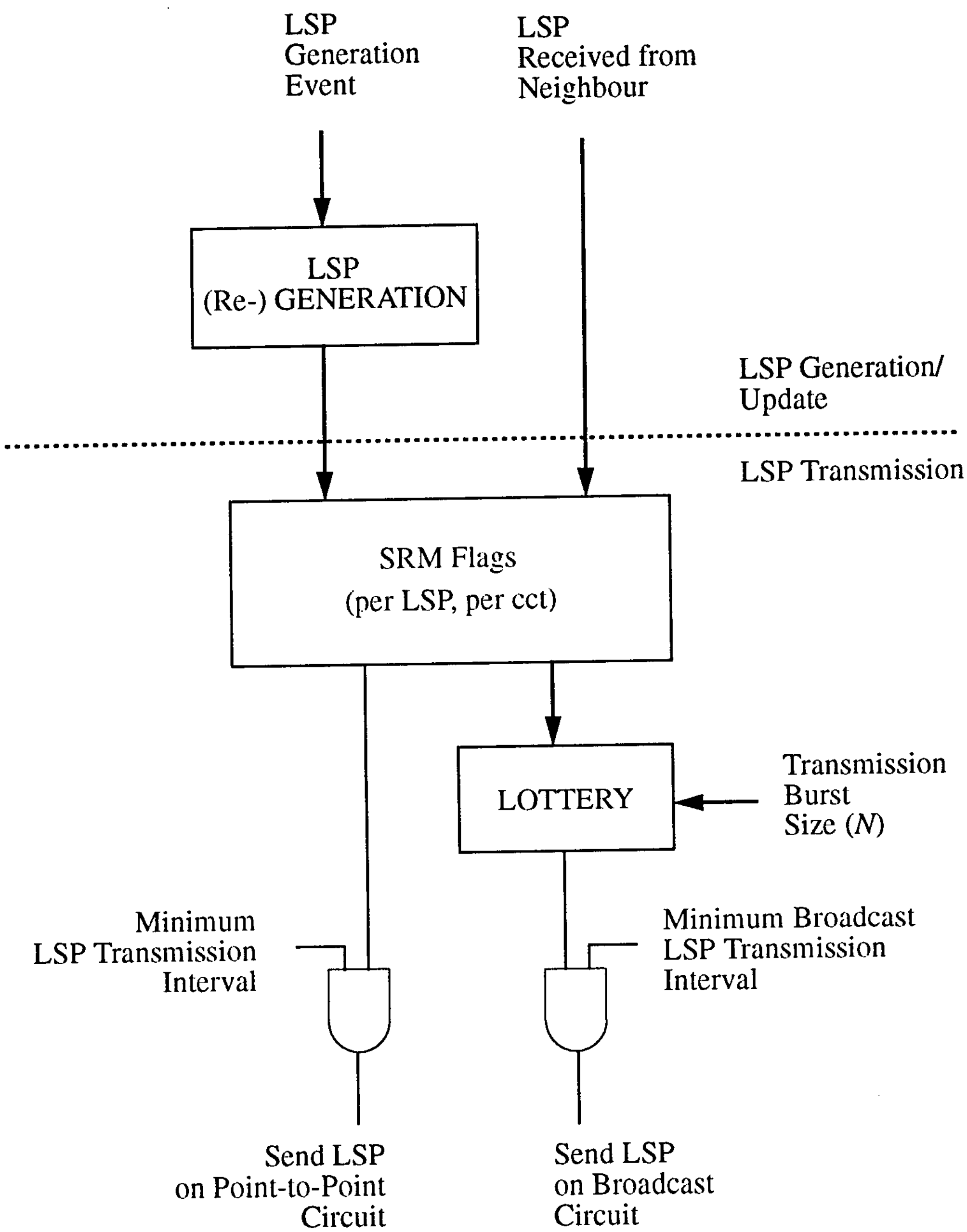


FIGURE 19. Simplified representation of main mechanisms leading to the transmission of a LSP

Retix's implementation is a variant of this scheme where up to *N* LSPs (*N*=5) are randomly picked and transmitted back-to-back every *N* x *minimumBroadcastLSPTransmissionInterval* seconds. This technique is used to reduce the processing overhead.

From this discussion, it should be apparent that the loop response time cannot be any shorter than the sum of the delays incurred at every step along the QoS measurement - LSP flooding migration path. All timers being asynchronous with one another, some timers being periodic others aperiodic, some with jitter others without, can make each one of these delays fluctuate quite a bit. Let us simply state for now:

$$\text{Loop Response Time} > f(S_P, M_A, G_A, T_J) \qquad (8)$$

where the subscript *A* denotes an aperiodic timer, the subscripts *J* and *P* a periodic timer with and without jitter respectively, *f*() a function, and;

S = *StatisticsReportInterval*;

M = *minimumMetricUpdateInterval*;

G = *minimumLSPGenerationInterval*;

T = *minimumLSPTransmissionInterval* or *minimumBroadcastLSPTransmissionInterval* depending whether the circuit is point-to-point or broadcast respectively.

Further work is needed to derive a detailed mathematical expression of the loop response time.

ANNEX B — List of Acronyms

APDU	application protocol data unit
BER	bit error rate
bps	bits per second
CA	Canada
CC	congestion control
CLNP	connectionless network protocol
CLNS	connectionless network service
CLNS-MO	connectionless network service - management object
CLTP	connectionless transport protocol
CLTS	connectionless transport service
COTS	commercial off-the-shelf
CSNI	Communications System Network Interoperability
CSNP	complete sequence number protocol data unit
DA	dynamically assigned
DIS	designated intermediate system
DR, dr	designated router
DT-NPDU	data - network protocol data unit
DSCS	Defense Satellite Communication System
EHF	extra high frequency
ES	end system
ES-IS	end system to intermediate system
FLTSAT	Fleet Satellite
FR	France
GE	Germany
HF	high frequency
HFEM	HF European mesh
HFNA	HF North American
HFTA	HF Trans Atlantic
IIH	IS to IS hello protocol data unit
IS	intermediate system
IS-IS	intermediate system to intermediate system
ISO	International Standards Organization
LAN	local area network
LLC	logical link control
LLC1	logical link control type 1
LSP	link state protocol data unit
MAC	medium access control

MAN	metropolitan area network
Mbps	mega bits per second
MHS	message handling system
MNC	multinet controller
MSG	message generator
mu	metric unit
NATO	North Atlantic Treaty Organization
NL	The Netherlands
NPDU	network protocol data unit
NRL	Naval Research Laboratory
NSAP	network service access point
OSI	open systems interconnection
PDU	protocol data unit
PP	point-to-point
PSNP	partial sequence number protocol data unit
QoS	quality of service
R&D	research and development
RIB	routing information base
RMG	routing metric generator
SATCOM	satellite communications
SHF	super high frequency
SNAC	subnet access controller
SNAcP	subnetwork access protocol
SNDCP	subnetwork dependent convergence protocol
SNP	sequence number protocol data unit
SNPA	subnetwork point of attachment
SNPDU	subnetwork protocol data unit
SQ	source quench
SnSAP	subnetwork service access point
SnSDU	subnetwork service data unit
SPF	shortest-path-first
SRM	send routing message
STC	Shape Technical Centre
SVC	switched virtual circuits
TDG	tactical data generator
TTL	time-to-live
UDP	user data protocol
UK	The United Kingdom
US	The United States
VHF	very high frequency

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. Establishment sponsoring a contractor's report, or tasking agency, are entered in section 8.)

COMMUNICATION RESEARCH CENTRE
3701 CARLING AVENUE,
OTTAWA, ONTARIO K2H 8S2

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.)

A STUDY OF ISO 10589 PROTOCOL IN A DYNAMIC QoS BASED ROUTING ENVIRONMENT (U)

4. AUTHORS (Last name, first name, middle initial)

CLAUDE BILODEAU

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

APRIL 1996

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

66 PAGES

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

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.)

CRC REPORT

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

1410-233

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)

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

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.)

CRC-RP-96-001

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
( ) Distribution limited to defence departments and defence contractors; further distribution only as approved
( ) Distribution limited to defence departments and Canadian defence contractors; further distribution only as approved
( ) Distribution limited to government departments and agencies; further distribution only as approved
( ) Distribution limited to 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 Availabilty (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 offical languages unless the text is bilingual).

A COMPUTER SIMULATION MODEL OF THE CSNI NETWORK TOPOLOGY WAS DEVELOPED TO ASSESS THE EFFECTIVENESS OF QoS-DRIVEN ROUTING IN HETEROGENEOUS RADIO SUBNETWORKS. THE SIMULATED TOPOLOGY CONTAINS 11 ROUTERS, 33 SUBNETWORK ACCESS CONTROLLERS AND AS MANY END SYSTEMS FORMING A SINGLE IS-IS ROUTING DOMAIN OF 13 SUBNETS. THE STUDY INVESTIGATES HOW TO SET THE ISO 10589 PROTOCOL TIMERS AND QoS MEASUREMENT PERIOD TO BEST CONTROL THE RATE OF LINK STATE ADVERTISEMENTS WITHIN THE NETWORK AND TO MINIMIZE THE ROUTING DATA OVERHEAD OVER THE LINKS. SIMULATION RESULTS SHOW THAT THE ISO 10589: 1992(E) PROTOCOL IMPOSES SEVERAL LIMITATIONS WHEN THE ROUTING INFORMATION BASES ARE DYNAMICALLY MODIFIED TO ADAPT TO QoS CHANGES OF THE SUBNETS.

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 is not possible to select indexing terms which are Unclassified, the classification of each should be indicated as with the title.)

CSNI
ISO 10589
DYNAMIC ROUTINH
PROTOCOL SIMULATION
QUALITY OF SERVICE
OPNET SIMULATION
NETWORK INTEROPERABILITY
NETWORK SIMULATION