NORTH ATLANTIC TREATY ORGANISATION

RESEARCH AND TECHNOLOGY ORGANISATION

AC/323(IST-030)TP/33

www.rto.nato.int

RTO TECHNICAL REPORT **TR-IST-030**

Information Management over Disadvantaged Grids

(Gestion des informations sur des maillages désavantagés)

Final Report of the RTO Information Systems Technology Panel Task Group IST-030/RTG-012.

Published December 2007

NORTH ATLANTIC TREATY ORGANISATION

RESEARCH AND TECHNOLOGY ORGANISATION

AC/323(IST-030)TP/33

www.rto.nato.int

RTO TECHNICAL REPORT **TR-IST-030**

Information Management over Disadvantaged Grids

(Gestion des informations sur des maillages désavantagés)

Final Report of the RTO Information Systems Technology Panel Task Group IST-030/RTG-012.

The Research and Technology Organisation (RTO) of NATO

RTO is the single focus in NATO for Defence Research and Technology activities. Its mission is to conduct and promote co-operative research and information exchange. The objective is to support the development and effective use of national defence research and technology and to meet the military needs of the Alliance, to maintain a technological lead, and to provide advice to NATO and national decision makers. The RTO performs its mission with the support of an extensive network of national experts. It also ensures effective co-ordination with other NATO bodies involved in R&T activities.

RTO reports both to the Military Committee of NATO and to the Conference of National Armament Directors. It comprises a Research and Technology Board (RTB) as the highest level of national representation and the Research and Technology Agency (RTA), a dedicated staff with its headquarters in Neuilly, near Paris, France. In order to facilitate contacts with the military users and other NATO activities, a small part of the RTA staff is located in NATO Headquarters in Brussels. The Brussels staff also co-ordinates RTO's co-operation with nations in Middle and Eastern Europe, to which RTO attaches particular importance especially as working together in the field of research is one of the more promising areas of co-operation.

The total spectrum of R&T activities is covered by the following 7 bodies:

- AVT Applied Vehicle Technology Panel
- HFM Human Factors and Medicine Panel
- IST Information Systems Technology Panel
- NMSG NATO Modelling and Simulation Group
- SAS System Analysis and Studies Panel
- SCI Systems Concepts and Integration Panel
- SET Sensors and Electronics Technology Panel

These bodies are made up of national representatives as well as generally recognised 'world class' scientists. They also provide a communication link to military users and other NATO bodies. RTO's scientific and technological work is carried out by Technical Teams, created for specific activities and with a specific duration. Such Technical Teams can organise workshops, symposia, field trials, lecture series and training courses. An important function of these Technical Teams is to ensure the continuity of the expert networks.

RTO builds upon earlier co-operation in defence research and technology as set-up under the Advisory Group for Aerospace Research and Development (AGARD) and the Defence Research Group (DRG). AGARD and the DRG share common roots in that they were both established at the initiative of Dr Theodore von Kármán, a leading aerospace scientist, who early on recognised the importance of scientific support for the Allied Armed Forces. RTO is capitalising on these common roots in order to provide the Alliance and the NATO nations with a strong scientific and technological basis that will guarantee a solid base for the future.

Published December 2007

ISBN 978-92-837-0082-1

Table of Contents

List of Figures/Tables

List of Acronyms

ACK	Acknowledgement
ARDS ADM	Artillery Regimental Data System Advanced Development Model
ARM	ATCCIS Replication Mechanism
ATCCIS	Army Tactical Command and Control Information System
BER	Bit-Error Rate
C2	Command and Control
C2IEDM	Command and Control Information Exchange Data Model
C2IS	Command and Control Information System
CCM	CORBA Component Model
CORBA	Common Object Request Broker Architecture
CSMA/CA	Collision Sense Multiple Access / Collision Avoidance
DACCIS	Danish Army Command and Control Information System
DBMS	Database management system
DC	District of Columbia
DCE	Distributed Computing Environment
DCOM	Distributed Component Object Model
DP	Data Provider
DR	Data Receiver
DRDC	Defence Research and Development Canada
FEC	Forward error correction
FGAN	Forschungsgesellschaft für Angewandte Naturwissenschaften
FhG	Fraunhofer-Gesellschaft
FKIE	Forschungsinstitut für Kommunikation, Informationsverarbeitung und Ergonomie
FOKUS	Fraunhofer Institut Offene Kommunikationssysteme
GPS	Global Positioning System
HQ	Headquarters
HTML	Hyper Text Markup Language
IEEE	Institute of Electrical and Electronics Engineers
IM	Information Management
IP	Internet Protocol
ISO	International Standards Organization
IST	Information Systems Technology
JMS	Java Message Service
JTRS	Joint Tactical Radio System
JXTA	Juxtapose
Kbps	Kilobits per second
LAN	Local Area Network
LBTB	Low Bandwidth Test Bed

LC2IEDM	Land Command and Control Information Exchange Data Model
LOS	Line of sight
MAC	Media Access Control
MANET	Mobile Ad Hoc Network
Mbps	Megabits per second
MIP	Multilateral Interoperability Programme
MTIR	MIP Tactical C2IS Interoperability Requirement
MTU	Maximum Transmission Unit
MVD	Majority Vote Detect
NATO	North Atlantic Treaty OrganiSation
NSA	National Security Agency
ODB	Operational Database
OFDM	Orthogonal Frequency Division Multiplexing
OGSA	Open Grid Services Architecture
OMG	Object Management Group
OSF	Open Software Foundation
OSI	Open Systems Interconnection
OWG	Operational Working Group
P2P	Peer to Peer
P2PS	Peer-to-Peer System
PDU	Protocol Data Unit
PfP	Partners for Peace
QoS	Quality of Service
RM	Replication mechanism
RMI	Remote Method Invocation
RTCORBA	Real-Time CORBA
RTG	Research Task Group
RTL	Replication Transport Layer
RTO	Research and Technology Organisation
SGML	Standard Generalized Markup Language
SINCGARS	Single Channel Ground and Airborne Radio System
SIP	Session Initiation Protocol
STN	Simulated Tactical Node
TCP	Transport Control Protocol
TD	Technology Demonstration
TNO	Nederlandse Organisatie voor Toegepast Natuurwetenschappelijk Onderzoek
UDP	User Datagram Protocol
UHF	Ultra High Frequency
VHF	Very High Frequency
WANODE	Wide Area Network for Operational Data Exchange
XML	Extensible Markup Language
zlib	Name of open-source lossless data compression algorithm

Acknowledgements

The Chairman would sincerely like to thank the members of IST-030/RTG-012 for their contributions to the Task Group's success. All Task Group members found themselves with serious demands on their time yet managed to find the time to advance Task Group objectives. The Chairman is very appreciative of their effort, perseverance and collaborative approach.

Thanks are also due to the home organizations of the Task Group members for providing the time and financial support necessary for the Task Group members to participate in six meetings and three associated workshops over four years. I am grateful to the organisations that hosted these meetings/workshops. They are, in order of hosting:

- RTO Headquarters, Neuilly-sur-Seine, France;
- Military Institute of Technology, Warsaw, Poland;
- DRDC Valcartier, Québec City, Canada;
- FGAN/FKIE, Wachtberg, Germany; and
- Naval Research Laboratory, Washington, DC, USA.

The Chairman would like to say a special word of thanks to Dr. J. Grosche and management at FGAN/FKIE who agreed to nominate Mr. Schmeing as a replacement for Dr. Fassbender when the latter left FGAN in 2002. Due to its small size, the Task Group could not have continued without this support. He is also very indebted to Herr Karlheinz Wagner and Frau Ute Spickermann of FGAN/FKIE and to BarbaraJo Cox and members of Ray Cole's Branch at NRL for their splendid logistic/administrative support of the second and third workshops. These workshops would not have happened without their timely and energetic support.

Of course, workshops are only as good as the people who attend them. Task Group members would join me in thanking the participants in the three informal Task Group workshops for their enthusiastic participation and excellent presentations. A special word of thanks is due to the Keynote Speakers at the first and third workshops, Dr. Sam Chamberlain of U.S. Army Research Laboratory and Professor Michael Pursley of Clemson University, whose strong presentations set the stage for two very successful events.

The Chairman would also like to thank the members of the IST Panel for their support and guidance, and members of the Panel Executive, particularly LCol Alain Gouay, LCol Patrick Prodhome, and Aysegul Apaydin for their generous support, particularly during the Task Group meetings at RTO Headquarters.

Finally, the Chairman would like to thank DRDC Valcartier, Mr. Jean-Claude St-Jacques and the other members of the 'High Capacity Tactical Communications Network' Technology Demonstration Project for contributing the simulation results presented in Chapter 7.

The Task Group hopes that its study has served to make a contribution to the understanding of a complex but under-appreciated problem of importance to all NATO forces, namely managing data flow over disadvantaged tactical wireless communications networks in a way that takes into account network state, battlefield state and operational importance of the data being transmitted.

Allan Gibb, Ph. D.
Chairman IST-030/RTG-012
February 2007

IST-030/RTG-012 Task Group Membership List

CANADA

Dr. Allan Gibb (Chairman)*
c/o Jean-Claude St-Jacques
System of Systems Section
DRDC – Valcartier
2459 Pie XI Blvd North
Val-Bélair, Québec G3J 1X5
Tel.: +1-418-844-4000 (ext. 4376)
Fax.: +1-418-844-4538
e-mail: jean-claude.st-jacques@drdc-rddc.gc.ca

GERMANY

(until December 2002)
Dr. Heinz Faßbender
FGAN / FKIE†

†Present address:
FH Aachen
FB 05
Eupener Str. 70
D-52066 Aachen
Tel.: +49-241-6009-1913
e-mail: fassbender@fh-aachen.de

(since August 2003)
Mr. Michael Schmeing
FGAN / FKIE
Neuenahrer Strasse 20
D-53343 Wachtberg-Werthhoven
Tel.: +49-228-9435-593
Fax.: +49-228-9435-685
e-mail: schmeing@fgan.de

POLAND

Dr. Jaroslaw Michalak
Military University of Technology
Institute of Communications Systems
Kaligiesko 2
00-908 Warsaw 49
Tel: 48-22-683-7733
Fax: 48 22 683 9038
e-mail: jmichalak@wel.wat.waw.pl

UNITED STATES

Dr. Jeffrey E. Wieselthier
Information Technology Division
Code 5521
Naval Research Laboratory
Washington, DC 20375
USA
Tel: +1-202-767-3043
Fax : +1-202-767-1191
email: wieselthier@itd.nrl.navy.mil

* The Task Group Chairman, Dr. Allan Gibb, retired from active government service on March 30, 2007. He can be contacted at allan.gibb@sympatico.ca. All requests for additional copies of this CD-ROM should be directed to Jean-Claude St-Jacques at the above address.

Information Management over Disadvantaged Grids

(RTO-TR-IST-030)

Executive Summary

This report summarizes a four-year study carried out by NATO RTG-012/IST-030 Research Task Group on the problem of "Information Management over Disadvantaged Grids". Such disadvantaged grids (e.g., tactical ad hoc military radio networks) are characterized by low bandwidth, variable throughput, unreliable connectivity, and energy constraints imposed by the wireless communications grid that links the nodes. The objective of managed information exchange is to support the commander's ability to execute command and control by providing a timely flow of accurate, relevant information. The highly mobile tactical military environment creates several challenges not endemic to either strategic or civilian environments. The Task Group studied managed information exchange in this communications environment from three different perspectives within a system architecture: the application level, the middleware level and the network level.

The Task Group limited its scope to land-based data exchange on the tactical battlefield (below brigade level) where all nodes are mobile and the exchange medium is combat net radio. The impact of security requirements was considered to be outside the scope of the study. Two alternative approaches to tactical information exchange, namely data replication and formal messaging, are addressed in this report. The Task Group focused primarily on data replication, since it was felt that it offers the most potential for minimizing bandwidth demands, by propagating database changes only, and for maximizing interoperability, by exchanging data based on an agreed formal data schema. The report concludes that asynchronous replication mechanisms are best for this type of communications environment, that an "all-informed" data distribution scheme may offer advantages, and that the replication mechanism and network need to cooperate to ensure that priority is given to maintaining consistent values across the net for those data judged to be of highest operational value (where necessary, at the expense of data of lower value).

Due to the highly variable quality of the tactical communications channels and the unpredictable nature of the tactical battlefield, it is argued that dynamic adaptation to rapid changes in either the communications or battlefield environment is required to achieve optimal information exchange. This adaptation is possible only if some information about the current status of the network is available to the middleware and/or application layer in each participating node. Concerning middleware design, the report concludes that next-generation middleware must meet several new requirements to satisfy the operating challenges in the tactical domain. The most important of these are context awareness, adaptivity, and the ability to function with acceptable levels of performance in both non-disadvantaged and disadvantaged communication environments.

This report identifies special characteristics of ad hoc wireless networks that differentiate them fundamentally from wired or cellular networks, notably the lack of an infrastructure and the fact that the set of network links and their capacities are not determined a priori. The report discusses how designing tactical ad hoc wireless networks using cross-layer techniques rather than traditional layered design principles can provide performance benefits throughout the whole system (radio physical layer through application layer). However, the report discourages the complete abandonment of layers.

The report also discusses ‘energy-efficient’ and ‘energy-constrained’ modes of operation, modes crucial to tactical networks involving dismounted soldiers due to limitations on battery weight. The authors conclude that techniques based on minimization of total energy expenditure do not necessarily perform well when batteries cannot be replaced. Energy-constrained operation leads to strong coupling among functions at several network layers, and consequently can benefit from the use of cross-layer network protocols.

Results are presented from simulations of a tactical scenario in which exchange of position updates over a single tactical radio subnet is accomplished via data replication based on an all-informed distribution model. The results illustrate the positive impact that application-layer information management techniques that reduce payload size, or limit offered load through application of context-sensitive business rules, can have on information flow over disadvantaged tactical communication grids.

The authors’ overall conclusion is that, for optimal information exchange performance in the tactical wireless domain, systems need to be designed from a holistic perspective. All levels of a system architecture (application/database, middleware and network) must be designed to work cooperatively to manage the information flow. This report attempts to identify required attributes that must be present at each level to enable this cooperative behaviour.

Gestion des informations sur des maillages désavantagés

(RTO-TR-IST-030)

Synthèse

Ce rapport résumé une étude de quatre ans menée par le groupe de recherche RTG-012/IST-030 de l'OTAN sur le problème de la « Gestion des informations sur des maillages désavantagés ». Ces maillages désavantagés (ex. : réseaux radio militaires ad hoc tactiques) se caractérisent par une bande passante étroite, un rendement variable, une connectivité peu fiable et des contraintes d'énergie imposées par le maillage de communications radio qui relie les nœuds. Le but de l'échange géré des informations est de soutenir la capacité du chef à exécuter Commandement et Contrôle en fournissant à temps un flux d'informations précises et appropriées. L'environnement militaire, tactique hautement mobile, crée plusieurs défis peu naturels aux environnements stratégiques et civils. Le Groupe de Recherche (RTG) a étudié l'échange géré d'informations dans cet environnement de communications à partir de trois perspectives différentes au sein d'une architecture système : le niveau Application, le niveau Intermédiaire et le niveau Réseau.

Le groupe de recherche a restreint son domaine à l'échange terrestre de données sur le champ de bataille tactique (en-dessous du niveau Brigade), là où tous les nœuds sont mobiles et le milieu d'échange est la radio de combat en réseau. Les impacts des exigences de sécurité ont été considérés comme en-dehors de l'étude. Deux approches alternatives à l'échange d'informations tactiques, à savoir : la duplication des données et la messagerie formelle, sont traités dans ce rapport. Le groupe de recherche s'est principalement concentré sur la duplication des données, car il a été perçu qu'elle offrait le plus de potentiel pour minimiser les exigences de la bande passante, en ne propageant que les échanges de bases de données, tout en maximisant l'interopérabilité, en échangeant des données basées sur un schéma convenu et formel. Ce rapport conclut que les mécanismes asynchrones de duplication sont les mieux adaptés à ce type d'environnement de communications, qu'un programme généralisé d'informations par distribution de données peut présenter des avantages, et que le mécanisme de duplication et le réseau doivent coopérer pour s'assurer que priorité est donnée au maintien de valeurs logiques à travers le réseau pour les données jugées de la plus haute valeur opérationnelle (si nécessaire, aux dépens de données de moindre valeur).

Du fait de la qualité très variable des canaux de communications tactiques et de la nature imprévisible du champ de bataille tactique, il est soutenu que l'adaptation dynamique à des changements rapides, soit dans l'environnement de communications soit sur le champ de bataille, est nécessaire pour obtenir l'échange optimal d'informations. Cette adaptation n'est possible que si certaines informations sur l'état en cours du réseau sont disponibles au niveau de la couche Intermédiaire et/ou Application de chaque nœud participant. Pour ce qui est de la conception au niveau Intermédiaire, le rapport conclut que ce niveau de Nouvelle Génération doit répondre à plusieurs exigences nouvelles pour satisfaire les défis de fonctionnement dans le domaine tactique. Le plus important d'entre eux est la connaissance du contexte, l'adaptabilité et la faculté à fonctionner avec des niveaux acceptables de performance, à la fois dans des environnements non désavantagés et désavantagés.

Ce rapport identifie les caractéristiques spéciales des réseaux sans-fil ad hoc qui les différencient fondamentalement des réseaux filaires ou cellulaires, notamment par le manque d'infrastructure et le fait que l'ensemble des liens du réseau et leur capacité sont déterminées a priori. Ce rapport discute de la manière de concevoir des réseaux tactiques radio ad hoc au moyen de techniques « inter-couches » plutôt

qu'à partir des principes traditionnels « en couches », et comment cela peut présenter des avantages en termes de performances dans tout le système (couche Radio Physique jusqu'à couche Application). Ce rapport décourage toutefois l'abandon total des couches.

Ce rapport discute aussi des modes de fonctionnement « peu gourmands en énergie » et « à énergie contrôlée ». Ces modes sont cruciaux pour les réseaux tactiques mettant en jeu des soldats à pieds à cause des limites de poids de la batterie. Les auteurs concluent que les techniques basées sur la minimisation des dépenses totales en énergie ne fonctionnent pas si bien lorsque les batteries ne peuvent être remplacées. Le contrôle de l'énergie mène à un fort couplage entre les fonctions au niveau de plusieurs couches du réseau, et peut donc profiter de l'utilisation de protocoles de réseaux inter-couches.

Les résultats sont présentés à partir de simulations d'un scénario tactique dans lequel l'échange de mises à jour de position sur un seul sous-réseau radio tactique est effectué par l'intermédiaire de la duplication de données basée sur un modèle généralisé de distribution. Les résultats illustrent l'impact positif que des techniques de gestion d'informations au niveau Application qui diminuent ou limitent la charge proposée en appliquant des règles commerciales sensibles au contexte, peuvent avoir sur le flux d'informations par rapport aux maillages de communications tactiques désavantagées.

La conclusion générale des auteurs est que pour des performances optimales d'échange d'informations dans le domaine de la radio tactique, les systèmes doivent être conçus à partir d'une perspective holistique. Tous les niveaux d'une architecture système (Application/Base de Données, niveau Intermédiaire et Réseau) doivent être conçus pour travailler en coopération et gérer le flux d'informations. Ce rapport tente d'identifier les attributs nécessaires qui doivent être présents à chaque niveau pour permettre un comportement coopératif.

Chapter 1 – INTRODUCTION

Mobile communication is an important military requirement. Voice communications still occupy a pre-eminent place in Army operations. Present-generation digital data communications at the tactical level (below brigade) are accomplished using radio systems designed primarily with voice in mind. Data throughput tends to be very limited (less than one Kbit/second is not uncommon) and highly variable. Digital command and control systems offer the promise of increased battlefield awareness. To deliver on this promise, the communication backbone must be capable of distributing relevant sets of digital data among participating command, control and information system (C2IS) nodes accurately and with a timeliness that permits friendly commanders to act within the decision cycle of the enemy commanders. Satisfying data distribution requirements of completeness, accuracy and timeliness when the communication system is characterized by low and variable throughput and highly unreliable connectivity represents a considerable challenge. Realistically, the limitations of the mobile wireless communications network will make it impossible to satisfy fully all of these requirements all of the time. Dynamic trade-offs between these factors will be required. A key factor in managing these tradeoffs is a set of adaptive protocols within each C2IS node. These protocols must exploit current information about the constantly-evolving situation picture contained in the node's database, as well as information about the current state of the communications network, with the goal of optimizing the timeliness and relevance of information passed between nodes. Commercial products do not provide protocols with the sophistication required for the demanding wireless military environment. In general, the products assume the presence of reliable high bandwidth links. This assumption is not valid on the tactical battlefield.

The Research and Technology Organisation's (RTO) Information Systems Technology (IST) Panel recognized the challenge inherent in distributing timely and relevant tactical information as digital data over a disadvantaged communication grid (i.e., over a mobile wireless communication network characterized by low and variable throughput, unreliable connectivity and energy-constrained nodes). In order to address that problem, the Panel authorized in October 1999 the formation of an Exploratory Team on Information Management over Disadvantaged Grids. The Exploratory Team met at RTO Headquarters in Paris in May 2000 and concluded that the problem of Information Management over Disadvantaged Grids should be addressed through formation of a Task Group under the IST Panel.

Task Group 12 on 'Information Management over Disadvantaged Grids' was formed in January 2001. The Task Group consisted of four countries: Canada, Germany, Poland and United States, with the Chairman being provided by Canada. The objective of the Task Group was the following:

> Investigation of adaptive information management schemes, implemented in the nodes of tactical command and control systems, to mitigate the effects of low bandwidth, variable throughput, unreliable connectivity and energy-constrained nodes imposed by the mobile wireless communications grid that links the command and control nodes.

The Task Group limited the scope of its study to the tactical wireless domain for a Land Force operating in a "national" context (i.e., issues related to multinational coalition interoperability were not addressed). The Task Group also decided that the impact of security requirements on information exchange protocols was a large topic that lay beyond the scope of the Task Group's mandate. Therefore, security-specific considerations (for example, managing exchange across classified and unclassified domains) were not addressed in this study.

Chapter 2 – BACKGROUND

2.1 PROBLEM FRAMEWORK

Analysis for formulating the Task Group Programme of Work was initiated during two meetings of Exploratory Team 014 and was completed during the first and second meetings of TG-012. A summary of that analysis follows:

Four different frameworks for analyzing the problem being addressed by the Task Group were considered:

1) Architectural

   Adaptive information management strategies would be classified according to the level in a system architecture where they are applied, namely:

   - Application or application database;
   - Middleware layer; and
   - Communications network layers.

2) Command Level

   Adaptive information management strategies would be classified according to the command level in the tactical domain where they are applied:

   - Brigade;
   - Battalion;
   - Company; and
   - Section.

3) Conflict Intensity

   Adaptive information management strategies would be classified according to the different intensities of conflict where they would be used, namely:

   - High intensity;
   - Medium intensity;
   - Low intensity; and
   - Operations other than war.

4) Tactical Radio Characteristics

   Adaptive information management strategies would be classified according to the nature of the radio systems over which the data is being passed, namely:

   - Data only; and
   - Integrated voice and data.

The Task Group concluded that the architectural framework provided the most logical framework for analyzing the problem. The factors of command level, conflict intensity and radio characteristics can each influence the nature of information strategies that are employed and the effectiveness of those strategies. However, it was felt that the level in a system architecture where the techniques are applied has the most direct bearing on the nature of techniques selected, the way in which the techniques are implemented, and their ultimate effectiveness.

The Task Group decided that it should limit its attention to national (as opposed to coalition) forces employing land-based tactical radio systems (Army, Marines)[1] deployed in mobile nodes. Further, it was agreed that attention would be restricted to battalion level and below, in particular:

a) A battalion command net employing vehicle-mounted radios; and

b) A small dismounted unit (for example, a section of infantry) employing man-portable radios.

The Task Group further concluded that its attention should be focused on scenarios of high-intensity conflict, since that is the environment in which the demand on the communication system is most intense and is therefore the environment in which adaptive information management strategies can provide the most added value for the military user.

2.2 PROGRAMME OF WORK

It was decided that the programme of work which best served the Task Group goals would comprise the following key components:

1) A series of three workshops. Each workshop would:
   - Address a key issue in information management at a different level of C2IS architecture (application, middleware, communications network);
   - Be associated with a Task Group meeting;
   - Be of one – two day duration; and
   - Be open to all NATO (but not PfP) nations and involve, where possible, invited experts on the issue.

2) Sharing and analysis of results from one or more national experiments.

2.3 OVERVIEW OF WORKSHOPS

As part of its programme of work, the Task Group undertook to organize three workshops, each addressing a key issue in information management at a different level of C2IS architecture (application, middleware, communications network). The first workshop was held at DRDC – Valcartier in Canada in September 2002 on the topic of 'Data Replication over Disadvantaged Tactical Communication Links'. The second workshop was held at FGAN/FKIE in Germany in August 2003 on the topic of 'Role of Middleware in Systems Functioning over Mobile Wireless Networks'. The third workshop was held at Naval Research Laboratory in the United States in June 2004 on the topic of 'Cross-Layer Issues in the Design of Tactical Mobile Ad Hoc

[1] In this document, the general term Land Force will be used to refer to either Army or Marine elements employing land-based tactical radios.

Wireless Networks: Integration of Communication and Networking Functions to Support Optimal Information Management'.

A summary of the objective of each workshop is provided in the following three sections.

2.3.1 Data Replication over Disadvantaged Tactical Communication Links

The objective of this workshop was to address the problem of replicating data among distributed databases over disadvantaged (unreliable, low bandwidth or energy-constrained) mobile wireless military communication networks. Links in such networks are generally characterized by extremely limited and highly variable data throughput. For the foreseeable future, tactical data communications networks are unlikely to be able to distribute all of the timely information required to support global situation awareness. One consequence is that the data in adjacent databases on the battlefield will not be fully consistent much of the time. The challenge was to find ways for the military user in the tactical domain to exploit the information in databases effectively when traditional consistency expectations are unrealistic.

Application layer information exchange issues are discussed in Chapter 4. A copy of the technical programme for the data replication workshop can be found at Annex A.

2.3.2 Role of Middleware in Systems Functioning over Mobile Wireless Networks

The objective of this workshop was to address the role of middleware in disadvantaged military communication networks. In contrast to the cellular networks that are the standard for commercial applications, these military networks will typically be of an ad hoc or infrastructureless nature. The challenge was to find ways to use middleware effectively to enable and support intelligent information distribution in such communications environments.

Issues associated with the role of middleware are discussed in Chapter 5. A copy of the technical programme for the workshop on role of middleware can be found at Annex B.

2.3.3 Cross-Layer Issues in the Design of Tactical Mobile Ad Hoc Wireless Networks: Integration of Communication and Networking Functions to Support Optimal Information Management

The objective of this workshop was to address the challenge of 'top-to-bottom' information management in tactical command and control systems functioning over mobile ad hoc networks. Crucial issues relating to communication and networking in hostile environments include the characteristics of the physical channel, media-access mechanisms, routing, network control structures, and data structures. These issues span the entire layered structure from the physical layer up through the application layer. The conventional way to implement networks is to use a strictly layered structure, in which these issues are addressed separately, while defining interfaces to higher and lower layer functions. Although such an approach may be appropriate for wired networks, the characteristics of wireless networks suggest that improved performance may be obtained by addressing these issues in a coordinated fashion. This workshop addressed alternative approaches to various aspects of the communication/networking problem, and assessed the potential benefits that may be achieved through the vertical integration of layer functionality. Employing results obtained from the first two workshops, this workshop also addressed how these changes can support improved collaboration with business applications residing on the network to achieve information exchange that is sensitive to, and constantly optimized for, the changing needs of the military user in the tactical domain.

Networking issues associated with mobile ad hoc networks are discussed in Chapter 6. A copy of the technical programme for the workshop on cross-layer issues can be found at Annex C.

2.4 OVERVIEW OF CONTRIBUTIONS FROM NATIONAL EXPERIMENTS

The Task Group sought to organize one or more experiments, or to analyze available results of one or more experiments, that would validate and demonstrate the potential for adaptive information management schemes, implemented at the application/database layer within tactical C2 nodes, to minimize the impact of the low and variable throughput and unreliable connectivity of a mobile wireless military communications grid. After studying several options, the Task Group decided that it would focus its efforts on exploiting the low-bandwidth test bed being installed at DRDC Valcartier in Canada. The test bed was designed to support a set of national experiments planned under the Canadian High Capacity Tactical Communications Network Technology Demonstration project. The objectives of these experiments were very closely aligned with the overall objective of RTG-012. Canada offered to make some of its experimental results available to the Task Group.

At the Task Group's first meeting, the possibility was discussed of having the Task Group members observe experiments to be conducted using Poland's 'Wide Area Network for Operational Data Exchange' (WANODE). WANODE is a prototype system for tactical data exchange that uses a data replication mechanism to replicate data changes between mobile or fixed tactical nodes. The data replication mechanism is a custom mechanism that is not based on the ATCCIS Replication Mechanism (see Annex D). It was subsequently decided that observation of WANODE experiments would not be possible for security reasons. In October 2003, Poland offered the possibility of contributing some results from WANODE to support the Task Group's work. However, in late 2004, it was learned that performance metrics obtained using WANODE would not be available in time for inclusion in the Task Group's final report.

Similarly, a German simulation tool called FIT was considered but it was concluded that it did not provide the level of functionality required for Task Group work.

Experiments conducted using the Canadian Low Bandwidth Test Bed, and lessons learned from those experiments, are discussed in Chapter 7.

Chapter 3 – ARMY TACTICAL COMMAND, CONTROL AND COMMUNICATIONS ENVIRONMENT

The issues discussed in Chapters 4, 5 and 6 require a clear understanding of the constraints imposed by the tactical command, control and communications environment. An overview of this communications environment is described in the present chapter, based on an Army command and control structure. Marine forces deployed ashore tend to be smaller in size, but face most of the same challenges when employing land based radio systems to coordinate their actions on the ground.

3.1 MILITARY COMMAND AND CONTROL SYSTEM STRUCTURE

Command and control systems must support three types of relationships: command, support and proximity. The following extract is taken from [1]:

> "Command relationships exist whenever one unit or formation commander is a direct subordinate of another…Command requires a rich bi-directional exchange of information between the higher headquarters and the subordinate headquarters. The purpose of this exchange is to pass command information (plans, orders, task organization, battlefield geometry, alerts, warnings and status) between the two headquarters. This type of information exchange follows the parent-child relationship. The superior (parent) headquarters supplies directive information to the subordinate (child) headquarters – higher to lower; the subordinate (child) headquarters provides status information to the superior (parent) headquarters – lower to higher.
>
> Support relationships are a particular type of command relationship. Support relationships are established when one organization must aid, protect, complement or sustain another force. In the context of Command and Control (but not, for example, fire control) these organizations have the same requirements for information exchange as command relationship. Support relationships are of two types: Direct Support and General Support.
>
> Proximity relationships exist when units with no direct command or support relationship are operating in proximity to each other and must exchange non-command information in order to establish and maintain situational awareness. Examples of this type of relationship could be the flank coordination of adjacent tank and infantry battalions or the forward passage of lines of an armoured regiment through a mechanized infantry battle group. In proximity relationships, information flows horizontally between the headquarters of the units involved as peers, not parents/children. Units involved in proximity relationships may be subordinate to different higher headquarters."

The military command and control structure is hierarchical (Figure 3-1). A headquarters at a certain command level will be parent of one or more subordinate headquarters, and will itself be subordinate (child) to a higher headquarters. The communication infrastructure and flow of information over that infrastructure reflects this command hierarchy. A commander at a given level will generally be required to maintain information from one level up and two levels down in the command hierarchy, as well as from flanking formations with which he has a proximity relationship.

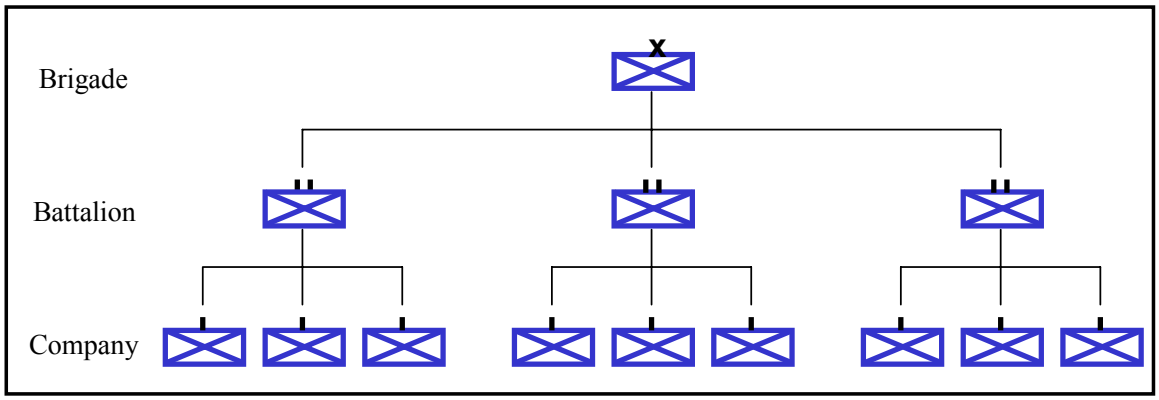


Figure 3-1: A Typical Hierarchy of Command Headquarters.

3.2 COMMAND AND CONTROL COMMUNICATIONS INFRASTRUCTURE

The command and control communications infrastructure is generally organized as a series of hierarchical subnetworks. A brigade command subnet will include a node at each of the three battalion headquarters under the brigade's command. A battalion command subnet may have a dozen or more nodes, but will include a node at each of the three company command posts under the battalion's command. Links from battalion headquarters to brigade or higher headquarters are typically provided by relatively reliable and high bandwidth satellite links or dial trunk systems employing wire or wideband microwave links. The Task Group focused its attention on the Army battlefield environment forward of a battle group or battalion[1] headquarters in which all communication occurs between mobile nodes equipped with a combat net radio operating in either the Very High Frequency (VHF) or Ultra High Frequency (UHF) band. A radio subnetwork consists of a set of radios tuned to a common assigned frequency. The subnet is linked to an adjacent subnet through a relay or gateway node that is common to both subnets. The relay node contains two or more radios, each tuned to a different subnet frequency. A verbal message received at the relay node from a sending node on one subnet is recorded in writing by a human operator, and then retransmitted verbally on the appropriate subnet to reach the destination addressee(s) on that subnet, if required. A gateway performs the same function for a data transmission, except that caching of the received transmission, and retransmission on the target subnet(s), are handled automatically by the gateway.

3.3 COMBAT NET RADIO COMMUNICATIONS ENVIRONMENT

The discussion in this section is taken from [2]. Forward of battalion, communication occurs over line-of-sight (LOS) radios operating in the Very High Frequency (VHF) or Ultra High Frequency (UHF) bands. Most of these radios are vehicle-mounted, but man-portable versions are also employed. At the lowest echelon, such as an infantry section on foot engaged in urban warfare, soldiers may use short-range radios operating in the High Band UHF. These radios have ranges of 400 meters or less. UHF radios are used for medium-range LOS wireless communication from 400 meters to 15 kilometres. VHF Radios must be used where non-LOS communication beyond 15 kilometres is required.

[1] Battalions are either mechanized battalions (two armoured companies, one infantry company) or infantry battalions (two infantry companies, one armoured company). A battle group is a battalion augmented with other assets (usually engineer or artillery assets).

The vast majority of VHF combat net radios have base bandwidths of 16 kilobits per second (Kbps) half duplex[2]. UHF radios can have base bandwidths as high as 288 Kbps full duplex. High Bandwidth UHF radios can have base bandwidths as high as 11 Mbps. In spite of their limited bandwidth, combat net radios operating in the VHF band still predominate because of their greater range and beyond-LOS capability.

The values quoted above apply to the physical layers of the network, the lowest layer of the seven-layer OSI model (Figure 3-2). The reality is that the useable throughput at the application layer is a fraction of this base rate. Factors such as forward error correction (FEC), encryption overhead (e.g., crypto synchronization sequence), acknowledgement request and retransmission mechanisms at data link and transport layer, and media access control mechanisms, are responsible for this reduction. Moreover, the effective throughput at any time may fall well below this maximum value due to variations in the performance of the physical radio channel caused by real-world factors such as terrain interference, atmospheric interference, multi-path (reflections) and prolonged fading. Effective throughput can become zero for periods of time for certain links, or for the entire subnet (in the case of imposed radio silence).

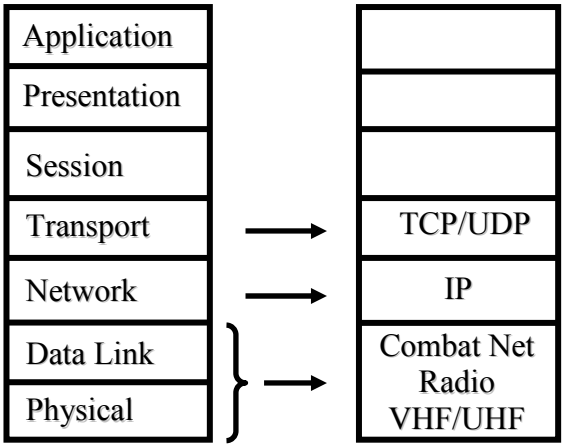


Figure 3-2: Seven-Layer ISO Network Reference Model.

A VHF combat net radio with a base rate of 16 Kbps may have a maximum effective throughput as low as 1 Kbps due to the factors quoted above. If ten users share the radio subnet, the maximum effective throughput per user will be only 100 bps. This figure assumes a data-only network with no voice contention for the channel. The reality is that the residual bit-error rate (BER) (i.e., the BER after error-control coding) for these radio channels can be as high as 10^{-5}.

One known UHF combat net radio can have a base rate as high as 288 Kbps. However, UHF radios have shorter ranges than VHF radios since the higher frequency UHF waves are more susceptible to the real-world factors quoted above. For the case considered, field measurements suggest that the maximum effective throughput at the top of the data link layer in a tactical environment would be approximately 16 Kbps for a link with a 22 Kbps base rate and approximately 80 Kbps for a link with a 100 Kbps base rate. For a subnet with 10 users, maximum effective throughput per user would be 1.6 Kbps and 8.0 Kbps respectively. Residual BER are similar to rates experienced with VHF radios.

[2] Half duplex means that the radio can either send or receive on the same channel, but cannot simultaneously send and receive. Full duplex means that the radio can simultaneously send and receive on the same channel.

In the Internet world, the transport protocol of choice for most applications is the Transport Control Protocol (TCP). TCP provides reliable packet delivery using a sequencing and positive acknowledgement scheme and is a connection-oriented protocol as it establishes a connection with each recipient. Although point-to-multipoint addressing schemes are being developed for TCP, in the radio domain the vast majority of systems still use a point-to-point addressing scheme for TCP. Therefore, in this domain, if a transmission is intended for *N* recipients, it must be sent *N* times. In a highly bandwidth-constrained wireless environment, the communication overhead associated with the use of a connection-oriented transport protocol like TCP is generally unacceptable. As well, the congestion-control mechanisms used by TCP were developed for the wired domain. TCP interprets latency and packet loss as evidence of congestion, to which its reaction is to throttle back the offered load to the network. This results in a significantly lower TCP throughput than the network can actually provide. In low bandwidth, high latency, and relatively high error networks like those found in the tactical radio domain, many TCP connections would be terminated due to these effects. User Datagram Protocol (UDP) is a connectionless alternative to TCP. However, UDP is unreliable since it does not employ sequencing or acknowledgement. Custom middleware operating just above UDP may be required to optimise reliability of packet transmissions in a connectionless, high BER tactical wireless environment.

Chapter 4 – APPLICATION LAYER INFORMATION EXCHANGE ISSUES

Digitization of Command and Control information seeks to facilitate increased operational tempo by reducing the need to slow down, or stop, in order to regain situational understanding. In order to realize this operational objective, information must continue to flow as the forces move throughout the battlespace. Since commanders may need to receive and disseminate situational awareness and execution information from any location in the area of operations, all of the sources and destinations of information should remain available to the commander at all times. The objective of managed information exchange is to support the commander's ability to execute command and control by providing a timely flow of accurate, relevant information.

Computers assist human beings by maintaining and manipulating large amounts of data that represent meaningful information to the human beings that use the computers. For information to be managed and manipulated by a computer, however, it must be highly structured. This structure typically takes the form of a structured data schema implemented in a database. The structure is important to permit the computer to deal with the data. However, the structure also serves to preserve a context for the data. Preservation of context (structure) for data is critical as the data are manipulated and exchanged.

The operational benefits of digital information technology employed in support of military command and control are derived mainly from the speed with which computers can access, retrieve, process and display structured data that have been entered into them. However, the operational benefits are also highly dependent upon the ability to exchange structured data of operational importance between computers with the speed necessary to support operational tempo.

4.1 STRUCTURED MESSAGING

Over voice radio, the standard means of communication for decades has been verbal transmission of structured military messages (example: "0 this is 6, Contact Grid 247653, six enemy tanks using routes in to the built up area and concentrating in that location oriented South. Continuing to observe. Time of contact 1430 hours. Over."). In the conversion from voice to digital technology, one option is to preserve the structured message paradigm in digital form. It has the advantage that it is a paradigm with which armies are already familiar. However, in a bandwidth-constrained data communication environment, use of structured military messages has two important disadvantages. First, a structured message must, by definition, be semantically complete (i.e., in a form sufficiently complete to be fully understood by the recipient without resort to external references). Second, the message structure itself creates a certain amount of overhead (header, field separators, etc.). The need for semantic completeness can create a tendency on the part of users to recall a previously sent message, modify the fields whose content have changed, and resend the entire message (including fields whose values have not changed). Sending of redundant information and the message structure itself combine to create a significant communication overhead that can limit the throughput of operationally important new data in a bandwidth-challenged radio environment.

4.2 DATA REPLICATION

An alternative to the use of structured messages is to perform information exchange via direct database-to-database exchange. Using this approach, one or more database transactions modify attributes in relational

tables in a local database using some combination of insert, update and delete operations. The local database transactions are packaged for transmission and copied using data replication middleware to recipient nodes. At the receiving end, the database transactions are unpackaged and applied directly to the databases in the recipient nodes. Such database-to-database exchanges can be more bandwidth-efficient than structured message exchanges because they can have minimal overhead and transmit only data values that have changed.

The term 'replication' refers to the systematic propagation and maintenance of copies of data between datastores within a distributed computing environment. The field of replication has two domains; file replication and data replication. Most commercial products address one of these two but not both. File replication is concerned with the copying of complete files while data replication operates at the level of database transactions.

In database systems, a transaction is a set of database operations that must succeed or fail as a single unit. A transaction can be considered a logical unit of work that transforms the database from one consistent state to another consistent state. Data replication middleware is used to copy database transactions from a source database to multiple replicate databases. In the military context of the present workshop, the databases in question are contained in computers located in stationary or mobile command headquarters, in military vehicles, or on dismounted soldiers. From the communications perspective, each of these entities is considered a node on a military command and control network.

4.2.1 Data Replication in a Bandwidth-Constrained Wireless Environment

4.2.1.1 Synchronous versus Asynchronous Replication

The distinction between a synchronous and an asynchronous communication paradigm is common to many topics dealing with data exchange. Although the topics may appear to employ separate usages of these terms, the usages have some key points in common. The fundamental characteristic of the synchronous paradigm is that all participants of the communication are engaged at the same time. Usually, synchronous communication or data exchange protocols include an immediate acknowledgement of the reception and some schema for retransmission in case of failure. If either or both of these characteristics are absent in a communication protocol, the protocol is called 'asynchronous'. 'Asynchronous' may refer to the fact that no (immediate) acknowledgement is sent back to the sender and/or to the fact that sender and receiver engage in the communication at different times.

The following paragraphs examine the use of an asynchronous communication protocol by data replication middleware, while Chapter 5 discusses use of synchronous and asynchronous protocols by middleware in general.

Data replication middleware is used to copy database transactions from a source database to multiple replicate databases. In a commercial application (e.g., a financial system), the transactions are applied synchronously (i.e., simultaneously), using a protocol known as two-phase commit. The protocol ensures that either all participating databases commit the transaction at the same instant, or none do. Use of such a protocol produces what is termed 'tight' consistency of database content. Tight consistency means that data values are consistent (i.e., the same) across source and replicate databases at all times.

The two-phase commit protocol is a two-step protocol in which each replicate database sends a positive acknowledgement back to the source database for each step. In a bandwidth-constrained environment, use of such a protocol generates such a significant communication overhead that it can seriously impede throughput of operational data. In this case, asynchronous replication must be employed. Under asynchronous replication

the transaction(s) are applied to the source database before being replicated. Replicated transactions are applied to the replicate databases sometime after they are applied to the source database. Use of asynchronous replication produces what is termed 'loose' consistency of database content[1]. Loose consistency means that modified data values always become consistent across source and replicate databases, but only after a certain time delay (latency).

In the very low bandwidth regime associated with combat net radio, where the data throughput of the radio network can be chronically less than the load offered to the network, it is necessary to relax traditional consistency expectations. In this situation, at any point in time, a percentage of data values in the source and replicate databases will be inconsistent, producing a condition that has been dubbed 'lazy' consistency. Since one can achieve neither tight nor loose consistency of full database content in this situation, the challenge shifts to preservation of consistency for those attributes judged to be of highest operational importance. This problem is inherently more challenging than the problem of achieving tight or loose consistency. In the latter two cases, the end state is inherently stable, i.e., the database content will always evolve toward a state of consistency, albeit with some delay. In the case of chronically inadequate and variable bandwidth, the end state is uncertain and unstable. Data delivery is on a 'best effort' basis and there is no assured 'audit trail' for the database changes that are propagated. A continual deterioration in the consistency of database content is likely, unless steps are taken to manage information flow in a proactive manner.

In this low bandwidth environment, the data replication and transport mechanisms must include intelligent information management protocols capable of adapting to the variable throughput of the radio network and to changing battlefield priorities.

4.2.2 Desirable Characteristics of the Data Replication Service

In his keynote address at the data replication workshop, Chamberlain [3] from U.S. Army Research Laboratory discussed characteristics that a data replication service should have to function well in the tactical radio domain. He noted that the variation in bandwidth is as important as the actual bandwidth values in determining data throughput. Zero is a valid value for throughput in this domain. A command and control node must continue to function even when the data throughput equals zero. He proposed the use of local predictive algorithms to estimate changes in data values (e.g., vehicle position) during these periods.

To be useful, the information exchange service must support rapid selection and dissemination of information and hands-off operation. According to Chamberlain, the service should have three characteristics:

1) Automatic (hands-off, context-sensitive);
2) Adaptive (context-based, responsive to changing bandwidth); and
3) Affordable (treated as part of a Distributed Computing Environment).

The goal is to balance or 'tune' information management to available network resources.

Chamberlain's proposal for what he terms 'Model-Based Battle Command' is based upon the following four principles:

1) Use data abstractions as the medium of exchange (data replication);

[1] For this to be true, it is essential that the asynchronous replication mechanism preserve transactional semantics. Database-to-database replication of transaction events as they occur in the source database generally preserves transactional semantics whereas a complete or incremental periodic refresh (snapshot) mechanism does not.

2) Control exchange by active database triggers (provides context-sensitivity since triggers link the decision to replicate to the state (i.e., value) of certain database elements);

3) Allow reasonably different perceptions of battlefield (accept 'lazy consistency' of database content); and

4) Treat database synchronization as the realistic control of differing perceptions (management of 'lazy consistency' rather than enforcement of loose/tight consistency of database content).

4.2.2.1 Network Awareness

A key element in the above approach is the ability to monitor or measure bandwidth so that synchronization efforts can be adapted to current communication resources. To achieve this objective, according to Chamberlain a communications model, incorporating information such as network performance and connectivity data, should be part of the data schema. As well, network performance statistics should be passively collected (e.g., channel access delay, round-trip time, queue delay, number of failed versus successful transmissions), analyzed, and results made available to the application through the communications model in the data schema. Chamberlain also proposes that innovative protocol mechanisms be considered to permit more direct performance feedback to the application and more efficient data transmission. Possible protocol mechanisms include: stack cognizance (sharing of information between ISO layers, e.g., transport and data link layer), 'just-in-time' packet construction (packing several small (highest-priority) application Protocol Data Units (PDU) into a Maximum Transmission Unit (MTU)), and overhearing (the direct passage between the network and transport layer to the host of packets rather than addresses).

4.2.2.2 Data Ownership

According to [4], data replication service must enforce one or more of the following three models for data ownership: single, dynamic (i.e., transferable) and shared. 'Ownership' of a given data element refers to the right to modify the value of that data element. In a single-ownership model, the originator of the data retains ownership throughout the lifetime of the data element. In a dynamic-ownership model, any given data element has a single owner at any point in time, but the ownership is transferable; the present owner of the data element may or may not be the originator of that data element. In a shared-ownership model, there is no single owner for the data element. Any participating database may initiate a change in the value of the data element. A shared-ownership model should be avoided because it maximizes the possibility of data conflicts, and resolution of data conflicts can generate an undesirable increase in network traffic. Furthermore, it is important that the system track the identity of the owner of the data element and the ownership model that applies to that data element (when more than one ownership model is employed).

4.2.2.3 Data Recovery

It is also important [4] for data retransmission and recovery to be intelligently managed. At the packet level, this means tying the number of attempted retransmissions to the assigned transmission priority for the packet and to the known current performance of the network. At the level of bulk recovery (e.g., by a node that has been disconnected from the network for an extended period), this means minimizing the size of the recovery package by being selective as to content (for example, by ignoring time-sensitive data whose operational value ages quickly). With the latter approach, by definition the content of the recovering database will never be fully consistent with that of the other databases. However, a bulk recovery may tie up the radio sub-net for minutes or tens of minutes. By limiting the content of the recovery package, one is trading off consistency of stale data in one database in favour of consistency of fresh data in all databases.

4.2.2.4 Functional Requirements – Data Replication Service

A data replication service can be considered to consist of two parts. The Replication Mechanism (RM) is responsible to determine when replication of data from a local node to other network nodes is to occur and what information is to be replicated. It also packages and unpackages data to be replicated into/from a unit called a Replication Protocol Data Unit (PDU)[2]. A PDU is information that is delivered as a unit among peer entities of a network that may contain control information, address information, or data. The Replication Transport Layer (RTL) sits beneath the Replication Mechanism and provides a transport mechanism for the purposes of passing replication data, in the form of Replication PDUs, over the communications bearer.

The combat net radio environment, when compared to wired networks, is characterized by:

- Low bandwidth (e.g., 1 Kbps – 80 Kbps);
- High and variable latency (e.g., 0.5 sec); and
- Intermittent connectivity.

For optimal performance in the tactical wireless domain, the communication protocol in the Replication Transport Layer needs to be matched to both the characteristics of the communications bearer and the information exchange requirements of the Replication Mechanism.

One of the discussion groups at the data replication workshop identified a list of key functional requirements that should be satisfied by the Replication Mechanism and Replication Transport Layer. The list was organized into three parts, namely:

1) Functions to be implemented in the Replication Mechanism;
2) Functions to be implemented through some combination of the Replication Mechanism, the Replication Transport Layer, and, perhaps, an additional application; and
3) Functions to be implemented in the Replication Transport Layer.

4.2.2.4.1 Functional Requirements

A) Replication Mechanism:

- Intelligence to determine when replication is to occur. This function must be context sensitive.
- Intelligence to determine what data is to be replicated once a decision to replicate has been taken.
- Assembly of the replication protocol data unit (PDU). If information is available on what replications a sender or senders already possess, it is possible to determine if the PDU is semantically complete.

B) Combination of Replication Mechanism and Replication Transport Layer or other application:

- Functionality must be provided to enforce and mediate dependencies of RM and RTL on other system architecture components. Examples: dependence of RM on characteristics of a particular DBMS or other application, dependence of RTL on characteristics of a particular Data Model, RM, etc.
- Data Ownership: RM and RTL must be able to identify the owner of any given piece of data at any given time. This is tied to the policy for database key management and requires decisions about the

[2] In some analyses, the Replication Mechanism is further decomposed into a Replication Agent (the component that determines when replication is to occur and what information is to be replicated) and a Replication Server (the component that packages and unpackages data to be replicated into/from Replication PDUs).

implementation level and authority structure for such management. These decisions can depend upon how the data ownership information is being used (enforcing transactional integrity, resolving data conflicts, providing traceability of ownership over time, etc.).

- Intelligence to determine the level of effort devoted to transmitting a Replication PDU across the network based on the importance of the information contained in the PDU. For example, (1) how many retransmission attempts should be made by the RTL before it stops trying to transmit the PDU and, (2) if the RTL has more than one class of transport service (e.g., 'guaranteed' and 'best effort' transport services), which one should be used for each Replication PDU? Assuming the RTL has several mechanisms to distinguish classes of service, one of the significant issues is defining the criteria for assigning a level of 'importance' to each PDU.

C) Replication Transport Layer:

- Must support an Acknowledgement Scheme (negative acknowledgement preferred; positive acknowledgement scheme is very bandwidth-intensive).
- Must supply a retransmission protocol that takes account of time-varying communications bearer performance.
- Priority Scheme: RTL must support a prioritization scheme at the PDU level that takes account of time-varying bearer performance.
- Must enforce a degree of fault tolerance and be sensitive to time-varying communications bearer performance.
- Must support fragmentation and defragmentation of PDUs, as appropriate, determined by the characteristics of the network.
- Addressing Scheme: RTL should support an addressing scheme. Choices are a broadcast communications protocol or an addressing scheme that would allow point-to-point or point-to-multipoint communications to occur between RTL peers.
- Cognition of Network Structure: RTL should be cognizant of the network structure (who is on the radio net and the status of communications transmitted from each node). This information, if available, could be used to determine what information transmitted by a remote node has not been received by the local node and vice-versa. This capability would, however, introduce overhead to the RTL communications protocol.

4.2.2.4.2 Non-Functional Requirements

- The RM and RTL must operate effectively in both the non-disadvantaged and disadvantaged communication environments.

4.2.2.5 ATCCIS Replication Mechanism (ARM)

One of the most important examples of selective data distribution in the NATO context is the ATCCIS Replication Mechanism [5]. The present section contains a brief summary of the ATCCIS replication mechanism (a fuller description can be found in Annex D). Some advantages and disadvantages of the selective data distribution model versus all-informed data distribution model for the tactical wireless domain are discussed in Section 4.2.2.6.

ATCCIS (Army Tactical Command and Control Information System) was an international programme consisting of NATO nations (although not a formal NATO effort) aimed at identifying the minimum set of specifications to be included within C2ISs to allow the automatic transfer of selected command and control (C2) data. Their objective was to develop a specification for a hardware/software/vendor-independent interoperability solution. The ATCCIS programme ran from 1982 to 2002.

The Multinational Interoperability Programme is an international programme consisting of NATO nations (also not a NATO effort) whose focus is the fielding of an interoperability solution for multinational C2ISs. In 2002, ATCCIS merged with MIP. MIP adopted the products of the ATCCIS work as the basis for direct database-to-database exchange. However, MIP also maintains a structured message exchange mechanism.

The ATCCIS concept of interoperability is based upon the automatic transfer of standardized data elements that utilize agreed and common data identifiers based upon a common data interchange model. The common data model is called the Land C2 Information Exchange Data Model.

Additionally, the ATCCIS programme developed the specification for a mechanism that will permit interoperability of automated C2ISs through the partial replication of database content. The ATCCIS Replication mechanism (ARM) is selective in:

a) The data to be exchanged;

b) The recipients of the data; and

c) The transfer facility to be used.

Under the ATCCIS concept, nations use the common data model to preserve the meaning and relationships of the information exchanged between C2ISs across national boundaries. Nations are free to develop differently structured C2 databases for national use. The ARM is used to manage the exchange of information between databases of C2ISs across national boundaries based on the common data model.

The ARM employs the concept of contracts and filters. A *replication contract* is the means for controlling (selective) replication of database changes. A contract is established between a pair of replication nodes, designated as Data Provider (DP) and Data Receiver (DR). In the contract, the DP and DR agree that the DP will provide the DR with all data that satisfies the conditions of the contract. A contract specifies a *filter* and parameter values used to set filter conditions, as well as a DP and a DR. A filter is a set of criteria applied to the instances of a database in order to reduce the total set of data selected to a subset. Examples of filter types include geographical area, time, and order of battle (organizational). The contracts enforce a 'push' model in which the only data pushed to recipient nodes are those negotiated with the recipient node under the pre-agreed contract. To modify the set of data pushed to a particular data recipient by a data provider, a filter must be applied or the contract must be modified.

4.2.2.6 'All-Informed' Data Distribution Model versus 'Selective' Data Distribution Model

Two forms of data distribution model are important for the tactical wireless domain. An 'all-informed' data distribution model is based on the assumption that there is value in maintaining synchronized data content across all participating nodes on a subnetwork, at least for an agreed subset of entities within the data model. A 'selective' data distribution model makes the opposite assumption, i.e., that only data of interest to a particular node should be sent to that node. The underlying assumption in such a data distribution strategy is that a node should only receive data from external sources that are important to its assigned role (or that are permitted by a security policy). By the very nature of this distribution model, database content will not be

consistent across nodes on a subnet. An important example of this approach, the ATCCIS Replication Mechanism, was described in the preceding section.

A selective data distribution model ensures that a user of data is not forced to deal with data that are of no interest or importance to his role, and that, in the event of an unsuccessful retransmission, no effort is expended retransmitting the data to nodes that have no identified need for it. As well, it limits the quantity of received data that must be stored at the receiver node. Finally, if a point-to-point addressing scheme is used, it can support a security policy based on selective dissemination since transmission of a data payload can be limited to specific receiver nodes. However, in the tactical wireless domain, there can be serious operational disadvantages associated with selective data exchange due to the very limited bandwidth and variable connectivity of the wireless network. The disadvantages are:

a) Does not take advantage of shared medium (radio)

Most of the time, nodes on a wireless data subnet overhear data transmissions that are addressed to other nodes on the subnet. When data replication is used as the exchange mechanism, the database changes contained in those transmissions can be applied to the database in the overhearing node, even though they are not addressed to, or directly pertinent to the role of, the overhearing node. An all-informed data exchange model exploits these overheard data transmissions to the maximum extent possible while a selective data exchange model routinely discards this 'free' information.

b) May introduce single point(s) of failure

Under selective data distribution, each node is the sole provider of data which it 'owns', and subsets of its owned dataset are shared with a set of data receiver nodes. Each receiver node is subscribed, in principle, to a unique subset of the provider's dataset (the subsets for different receivers may overlap, but are, in general, not identical). Under this scheme, a given receiver node will only have knowledge of data being shared with it, not with the other receiver nodes, since the exchange agreements are between node-pairs. If the provider node is disabled, due to enemy action or other reason, for all practical purposes each receiver node will be limited to the narrow slice of the provider's dataset received from the provider node, and the set of receiver nodes will have no simple means of recovering or reconstructing all or part of the total dataset on the provider node, if that proves necessary (certainly no mechanism efficient enough to work acceptably over a tactical radio system). Under an all-informed data distribution model, nodes on the subnet overhear and capture data transmissions, even those for which they were not explicitly identified as a data receiver. If a provider node is disabled, each node on the subnet will have a copy of the content of all (or almost all) data transmissions emanating from the provider node to that point in time recorded in its database, thereby largely avoiding the problems noted above.

c) Requires data recovery from multiple nodes

In the event that a node leaves the subnetwork for a period of time, rejoins the network, and wants to recover data that it has missed, it will have no choice but to request missing data from each data provider node with which it has an exchange contract. If there are n such provider nodes, that will require n separate data transmissions. If three nodes need to recover data at the same time, and they have exchange contracts with n_1, n_2 and n_3 provider nodes respectively, that will require $n_1 + n_2 + n_3$ distinct data transmissions. This recovery process increases the traffic level on the radio channel and wastes bandwidth. By contrast, in an all-informed data exchange model, the node(s) can recover missing information in a single data transmission from any neighbour node, since the database content is, in principle, the same across all nodes.

d) Node cannot readily assume role of neighbour node without substantial one-time data transfer from that neighbour node

If a node is required to assume the role of a neighbour node, under a selective data exchange model it would be required to download all, or a substantial portion of, the information specific to the neighbour's role from the database of the neighbour node, since it would not have information specific to the neighbour's role in its own database. If the neighbour node database is not available, due to enemy action or other reason, it would have to request recovery of those role-specific data from the *n* data provider nodes with which the original node had an exchange contract. This would require n distinct (potentially large) data transmissions. In an all-informed data exchange model, no such large data transmissions are required because, in principle, the replacing node has the same data in its database as the replaced node.

An all-informed distribution model can still enforce a security policy of selective dissemination. Tactical radio transmissions are routinely encrypted for transmission and then decrypted at the receiving end. Thus the received decrypted transmission can be assumed to have originated from an authentic source. Selective dissemination can be enforced at the software level by implementing control software in the receiving nodes to ignore a (decrypted) data transmission if a security caveat in the received transmission instructs it to do so. With this model, data will be shared on an all-informed basis across all nodes on a subnet except for instances when the security policy specifically excludes this option.

If a point-to-multipoint addressing scheme is employed, it is possible to combine an all-informed distribution model for certain types of data (e.g., reports of friendly vehicle positions) with a selective distribution model for other types of data (e.g., passage of a fragmentary order from battalion HQ to company HQs). This hybrid distribution model will permit full synchronization of database content for certain types of data, and partial synchronization of database content for other types of data. Such a hybrid model may well provide the best match between operational requirements and bandwidth utilization in the tactical wireless domain.

4.3 DATA EXCHANGE USING XML

The Extensible Markup Language (XML) is a data interchange format developed by the World Wide Web Consortium and released in 1998. Inspired by the success of HTML (Hyper Text Markup Language), the established language of the Web, it is a simplified (and therefore more useable) subset of the older, more sophisticated SGML (Standard Generalized Markup Language). As such, XML is actually a meta-language describing a family of languages, namely languages defining interoperable, text based data formats. Although the historical roots of XML are in publishing, XML is also suited to the task of unambiguously identifying complex data structures that may never be viewed or printed. The ability to represent complex data structures and the data portability derived from the interoperable nature designed into XML, have made XML the data interchange format of choice for many applications where generality and portability are important.

An XML-based document has both a logical and a physical structure. The logical structure allows a document to be divided into named units and sub-units, called 'elements'. The physical structure allows components of the document, called 'entities', to be named and stored separately, sometimes in other data files (permitting information re-use and use of non-XML data (e.g., image data) by reference). An XML processor module is used to manage entities and combine them into a single data stream for validation by a parser and for accessing by the main application.

Because XML is a meta-language, there is no pre-defined list of elements. The user may name and use elements as desired. To allow automatic syntax checking, the user may use a Document Type Definition or an

XML Schema to define the elements allowed in a particular type of document. XML enables a high degree of control over the logical document structure. Unlike HTML, XML is targeted at the definition of data structures rather than text formatting. It therefore encourages the use of names for the elements that describe the nature of the object, rather than how it should be displayed or printed. This generalized markup approach means that the information is self-describing, and so can be located, extracted and manipulated as desired.

As a general rule, if one implements a data exchange mechanism and data exchange format flexible enough to handle the general case, the price one pays is that one is required to transmit more metadata (most of which describes data or document structure) to preserve that flexibility and generality. The 'self-describing' property of XML derives from the structure metadata (tags et al) accompanying the data content. This trend runs contrary to what is required in a low bandwidth communications environment, where the objective is to maximize data content and to minimize associated metadata in a transmitted payload. One option to reduce the size of XML data is compression. A description and comparison of different XML compression mechanisms can be found in [6].

4.4 SUMMARY AND CONCLUSIONS

Operational benefits of digital information technology employed in support of military command and control are derived mainly from the speed with which computers can access, retrieve, process and display structured data entered into them. Benefits are also highly dependent upon the ability to exchange structured data of operational importance between computers with the speed necessary to support operational tempo. In an operational military environment, two forms of data exchange are employed: structured message exchange or exchange of database transactions. The latter form can be more bandwidth-efficient because it propagates data changes only. It is implemented via a data replication mechanism.

The term 'replication' refers to the systematic propagation and maintenance of copies of data between datastores within a distributed computing environment. In the tactical domain, the datastores in question are contained in computers located in stationary or mobile command headquarters, in military vehicles, or on dismounted soldiers, all operating forward of battalion and communicating via a combat net radio. The combat net radio environment, when compared to wired networks, is characterized by very low bandwidth, high and variable latency, and intermittent connectivity.

A data replication service consists of two basic components, a Replication Mechanism and a Replication Transport Layer. The Replication Mechanism is responsible to determine when replication is to occur and what information is to be replicated. The Replication Transport Layer provides a transport mechanism for the purposes of passing replication data over the communications network.

To be effective in the low-bandwidth tactical wireless environment, the Replication Mechanism and Replication Transport Layer must include intelligent information management protocols capable of adapting to the variable throughput of the radio network and to changing battlefield priorities. This adaptivity requires that the mechanisms must have dynamic awareness of the network state and battlefield state.

Chapter 5 – MIDDLEWARE ISSUES

For the purpose of this report[1], middleware can be thought of as software providing a set of enabling services that reside between applications and the underlying operating systems, network protocol stacks and hardware. Middleware allows multiple processes running on one or more hosts to interact transparently across a network and can also enable and simplify integration of heterogeneous software and hardware components.

In the descriptions which follow, the term 'server' always refers to a piece of software offering a service to a client. A 'client' is consequently a software application or component using that service. Usually, a computer has both client and server components. Often a server uses other services and thus acts itself as a client of the server(s) providing those services. An example of such software is a file viewer for, e.g., a word processor format. For embedded graphics this viewer might use the service of a graphic rendering server while at the same time offering rendering services to the network. For programmes that have to display the word processor format it would act as a server while at the same time behaving as a client of the graphics viewer.

The term "application" (without any qualifier) is used in this chapter for any piece of software that accesses the services offered by the middleware.

5.1 MIDDLEWARE CATEGORIES

According to [7], middleware can be divided into different categories based upon the intended domain of application. The most important categories are:

- Transactional middleware;
- Message-oriented middleware;
- Procedural middleware; and
- (Distributed) Object or component oriented middleware.

These categories will be briefly described in the sections which follow.

5.1.1 Transactional Middleware

Transactional middleware offers the functionality associated with replication of database transactions described in the previous chapter. For synchronous data replication, a replicated transaction is either executed completely at all participating nodes or not at all. This requires a two-phase commit protocol where all nodes confirm to the initiating (master) node that the transactional information has been successfully received and that the transactions can be executed. Once these confirmations have been received by the master node, it tells the other nodes to commit the transaction and make it final. If at least one node indicates that it did not completely receive the transaction or that it cannot execute the transaction, the master node tells all nodes to do a rollback and not to apply *any* changes indicated in the transaction.

In a bandwidth-constrained environment, as noted in the previous chapter use of such a protocol generates such a significant communication overhead that it can seriously impede throughput of operational data. In this case, asynchronous replication is employed. Under asynchronous replication the transaction(s) are applied to

[1] It is important to note that no generally agreed-upon definition of the term "middleware" exists.

the source database before being replicated. Replicated transactions are applied to the replicate databases sometime after they are applied to the source database.

A prominent example of transactional middleware in the military domain is the ATCCIS Replication Mechanism (ARM). The ARM is a specification for an asynchronous replication mechanism that will permit interoperability of automated C2ISs through the partial replication of database content. The ARM is selective in:

a) The data to be exchanged;

b) The recipients of the data; and

c) The transfer facility to be used.

For more information on ATCCIS, ARM and selective vs. all-informed replication, see Section 4.2.2.5.

5.1.2 Message-Oriented Middleware

In message-oriented middleware, a client sends an asynchronous message containing a request and all meta-data such as authentication information to a server. The server processes the request and sends the results (or errors) in an (equally asynchronous) message back to the client. While it is possible to access local services using this type of middleware, it is mainly targeted at the (asynchronous) access to remote resources.

Message-oriented middleware is thus a form of asynchronous communication as described in Section 4.2.1.1, while the other classes described here are synchronous. Nevertheless, it is important to note, this does not necessarily require that the applications using the middleware be of the same nature. Due to the layered architecture (business application, middleware, network, …), it is possible to build an application communicating synchronously on top of a message-oriented (and thus asynchronous) middleware and vice versa.[2]

The disadvantage of message-oriented middleware is that, in general, it is implemented using a centralized server. This introduces a single point of failure. In domains featuring high bandwidth and reliable communication links, the risks associated with a single point of failure can be reduced through use of redundancy (e.g., a backup server). In disadvantaged environments, the very limited capacity of the communication links may not allow for fully redundant backup making this issue even more critical.

Decentralized versions of message-oriented middleware do exist, but the scalability is undetermined. All message oriented systems known to the authors also support multicast.

Examples of message-oriented middleware include JMS (Java Message Service), TIBCO Rendezvous™[3], publish/subscribe or synchronous message queues (or both), CORBA event notifications and other event notification systems, e.g., JXTA (Juxtapose) and Jini.

5.1.3 Procedural Middleware

Procedural middleware implements the concepts of remote procedure calls: the client initiates a procedure or function call on the server and receives the results in a similar manner. This is a synchronous system and

[2] This is not to say that such a combination would make sense. As a rule a synchronous application should use synchronous middleware and an asynchronous application should use asynchronous middleware.

[3] Messaging software product supplied by TIBCO Software Inc.

designed for the access to remote resources. The primary example of procedural middleware is DCE (Distributed Computing Environment) [8].

Procedural middleware requires that the calling function receive an answer before control is released back to the calling programme. This limitation means that procedural middleware is not favoured for disadvantaged computing environments especially when timeliness is an important aspect of performance. An important example for such an environment is the military tactical wireless domain.

5.1.4 Object and Distributed Object (Component) Middleware

Distributed object (or component) oriented middleware is based on the concepts of the object-oriented software development paradigm. One of the objectives is to allow transparent (and synchronous) communication to local and remote components while always using the most efficient transport mechanism.

Objects are the primitive elements of object-oriented programming. Objects are entities that encapsulate both the data describing the object and the instructions operating on those data. Distributed objects are packaged as independent pieces of code that can reside anywhere on a network and can be accessed by remote clients via method invocations. Components are standalone entities that can interact and interoperate across networks, applications, languages, tools and operating systems. Distributed objects are components, but not all components need be objects.

There are many examples of object and distributed-object middleware. These include: Object Management Group's CORBA (Common Object Request Broker Architecture) and CCM (CORBA Component Model), Microsoft's DCOM (Distributed Component Object Model), SUN's Enterprise Java Beans, Jini, JXTA (Juxtapose), Web Services, Agent technologies, .NET, OGSA (Open Grid Services Architecture), RMI (Remote Method Invocation) and P2PS (Peer-to-Peer System).

The advantages that distributed-object middleware can offer to the disadvantaged environment come at a price. Distributed-object middleware has significant network communication overhead required to support features such as the sharing of context data, object and service discovery, and the brokering of inter-object calls within multiple processes running across networks. In the tactical wireless domain where data throughput can be as low as 1 kilobit per second, such overhead can leave little or no network bandwidth available for transmission of real operational data. Thus, performance considerations may continue to inhibit widespread use of distributed-object middleware in disadvantaged environments and real-time systems. Efforts such as Real-Time CORBA (RTCORBA) [9] and minimumCORBA [10] are attempting to address this issue. However, until distributed-object middleware adequately addresses the issue, the limitation that it places on performance (effective throughput) due to its demands on the communication network may limit its use in the tactical wireless domain.

5.2 TRADITIONAL MIDDLEWARE REQUIREMENTS

There are five types of requirements addressed by traditional middleware solutions [7]. These are: network communication, coordination, reliability, scalability and heterogeneity. Each of these traditional requirements is discussed in turn.

5.2.1 Network Communication

All types of middleware described in the previous section allow access to resources on remote systems. To achieve this, they have to employ network communication. Depending on the type of middleware the

nature of the communication is hidden from the application to a greater or lesser extent. While component middleware completely hides the communication technologies from the application, the others only abstract to a certain degree or not at all.

5.2.2 Coordination

Beside the basic (network or local) communication between client and server, some degree of coordination is required. All types of middleware discussed offer this functionality to the extent required by the implemented paradigm. Message oriented middleware has the lowest requirements for coordination, because it works completely asynchronously. Transactional middleware on the other hand requires strict coordination to ensure the transactional character of its operation.

5.2.3 Reliability

Middleware is often deployed in mission-critical processes. This requires a high degree of reliability, which has always been an objective in the design and implementation of middleware systems.

5.2.4 Scalability

One central objective especially of component-oriented middleware is the abstraction from the actual location where a service resides. For the client it should not make a difference (at least not more than absolutely necessary) whether the service is located on the local machine, the local area network (LAN) or on a different continent. To achieve this, the middleware must scale arbitrarily in both directions: The smallest situation in which it must work is an isolated node without network connection, the largest a collaboration of nodes all over the globe.

5.2.5 Heterogeneity

While some middleware systems such as DCOM are limited to a single hardware or software platform, most systems have been designed to function in heterogeneous environments. One of their objectives has been to integrate different hardware and software platforms into a single environment. To a lesser extent this is as well true for different communication technologies such as network protocols or local communication mechanisms such as pipes or signals.

5.3 NEXT GENERATION MIDDLEWARE REQUIREMENTS

To overcome some of the limitations of existing middleware solutions and to increase the range of applicability of middleware, next-generation middleware will have to satisfy one or more of the following requirements: dynamic reconfiguration, adaptivity, context-awareness, asynchronous communication and lightweight design. Each of these requirements is discussed in turn.

5.3.1 Dynamic Reconfiguration

Next-generation middleware should be able to detect changes in available resources and to reallocate remaining resources, or to notify the application to change its behaviour. For example, interruptions that occur when servers are disconnected (e.g., because they are powered down or because they get out of range in a wireless environment) should be minimized and should not require manual intervention of the user. The middleware should search for an alternative server or combination of servers and continue to operate transparently.

5.3.2 Context Awareness

Next-generation middleware should serve as a mediator for collecting, organizing, and disseminating *relevant* context information to the upper layers (application) and lower layers (transport mechanisms). The context may include device or network characteristics, user activities and services.

To fulfil this role, the middleware should maintain a shared perception of network state [11]. In order to do so, it needs regular feedback from the underlying transport mechanisms on the network state, for example: link quality (speed and BER), transmit power (at local and remote nodes), and residual energy at the nodes. If the network does not provide this kind of information, the middleware must try to obtain it by doing its own measurements.

Each node on a radio net is autonomous and has its own unique perception of the performance of the radio net based upon its own experience interacting with other nodes on the network. A node hidden from the other nodes behind a hill may encounter communication difficulties that provide it with a totally different perception of network performance than the perception held by the remaining nodes, which are within line of sight of each other. In order to develop a shared perception of network state, it is necessary that each node on a radio net agree to share its local perception of network performance with the other nodes on the subnet at regular intervals (exploiting the shared radio medium). Development of local perception of network performance, sharing of that perception, and synthesis of a common view of network performance is best handled as a service provided through middleware residing on local nodes but collaborating with middleware on other nodes. The middleware should make this context information continuously or periodically available to the local application.

As well, next-generation middleware should maintain a shared perception of its own state. For example, for replication middleware, a data provider needs to know which nodes are available at any time as data receivers. This common information should be shared among all data provider nodes.

5.3.3 Adaptivity

Next generation middleware should have the ability to:

1) Recognize changes to its execution context and to adapt its behaviour to the changes in execution context; and
2) Recognize unmet needs within its execution context and to adapt itself to meet those needs.

An example of the first type of adaptivity would be an adjustment of middleware services provided to an application (e.g., reduction in frame rates for a real-time video streaming application) based on middleware awareness of reduced network throughput. An example of the second type of adaptivity would be automatic server reconfiguration described in Section 5.3.1.

5.3.4 Lightweight Design

Next generation middleware that will be effective in the tactical wireless domain should feature a lightweight design that minimizes demand on the network and implements a minimum range of functionality. The requirement for minimum functionality is most important for nodes with limited memory, storage capacity and processing power. Such nodes are usually man-portable nodes or sensor nodes powered by battery, where the minimum functionality can also serve the important goal of energy conservation.

5.3.5 Asynchronous Communication

As noted in Section 4.2.1.1, communication is described as 'asynchronous' when no (immediate) acknowledgement is sent back to the sender and/or the sender and receiver engage in the communication at different times.

Next generation middleware employing a client-server architecture should decouple the client and server components and use multicast communications where appropriate. The decoupling of client request and server response is particularly important in the tactical wireless domain where nodes may connect and disconnect from the network at unpredictable times and network access and latency are issues.

The disadvantage of introducing a single point of failure through use of a central server is discussed in Section 5.1.2.

5.4 MIDDLEWARE REQUIREMENTS FOR WIRED VS. WIRELESS DOMAINS

Military operations require use of more than one type of communications network. On the strategic and operational level, networks using reliable high-bandwidth communication links are normal, while in the tactical domain, disadvantaged networks and network nodes are common. Armies have a need to share information, services, functionality and, in some cases applications, across both domains. Middleware to support such sharing must be capable of functioning and providing services in both communications environments. This section examines design considerations for middleware to function effectively in both environments.

5.4.1 Differences between Wired Networks and Wireless Ad Hoc Networks

This topic is discussed in detail in Section 6.2. The primary differences result from the lack of an infrastructure for ad hoc networks. In fully connected wireless LANs, since there is single-hop connectivity among all the nodes, routing is not an issue. However, in ad hoc wireless networks it is possible to establish a link between any pair of nodes, provided that the signal-to-noise ratio at the receiving node is sufficiently high. Unlike the case of wired networks, the set of network links and their capacities are not determined a priori, but depend on factors such as distance between nodes, transmitted power, error-control schemes, other-user interference, and background noise.

Current civilian wireless technologies offer rather reliable links with relatively high data rates (>1 Mbit/s). Therefore, similar if not identical middleware technologies can be applied in both the wired and wireless domain. In the military tactical domain on the other hand, unreliable links with very low data rates are normal (16 – 64 Kbit/s at the physical layer; can be as low as 1 Kbit/s at the application layer). Additionally, these links may be subject to enemy interference. Middleware that is not designed specifically to adapt to the varying network topology and link capacity in the disadvantaged tactical wireless domain is not likely to function effectively in this domain.

5.4.2 Resource Limitations

In addition to the network limitations, there may be resource limitations at the nodes in the tactical domain. In a battalion or company command post, nodes may have sufficient computational resources and more than adequate energy, memory and storage space. Dismounted units may have only Personal Digital Assistants where all of these factors are strictly limited. Middleware designed for the disadvantaged environment must therefore scale to both kinds of nodes. This means that it must be possible to deploy only a subset of the middleware functionality in the smaller devices without losing interoperability.

5.4.3 Important Middleware Design Considerations

Each of the five requirements for next generation middleware identified in Section 5.3 are important for middleware to function well in both the non-disadvantaged and disadvantaged communications environment. In addition, the following design considerations are considered important.

5.4.3.1 Upperware and Lowerware

An important function of most middleware is to provide services to the application. These services can include meta-services such as a discovery or lookup service or services targeted at a special task. Examples of the latter are a VoIP service, a database replication service or a service providing track information such as sonar or radar. A discovery service is a service that gathers information about the available services on a network; such information is usually made available through a lookup service.

It can be useful to consider middleware as being divided into layers dubbed 'upperware' and 'lowerware'. The 'upperware' contains the services directly accessible by the application, as described in the preceding paragraph. The primary purpose of the 'lowerware' is to provide a connection for these services to the underlying system and the network. With such a scheme it is possible in principle to hide in the lowerware many of the adaptations required by a transition from a non-disadvantaged to a disadvantaged environment. The problem of designing upperware (and consequently the applications using it) to function in both communication environments can be considerably simplified.

An additional advantage of this internal layering of the middleware is that it allows the application and the network technologies to evolve at different rates. In particular, it may be possible to adapt the lowerware to take advantage of new network and communication technologies while maintaining a relatively stable interface between the upperware and the application.

5.4.3.2 Abstraction vs. Transparency

A decision must be made as to how much of the underlying network to abstract and how much to pass through to the application. An important example where this may prove critical is the multicast[4] network service. Usually non-disadvantaged networks do not offer multicast as it would not provide much advantage over a unicast service. The disadvantaged environment on the other hand usually employs a shared medium where bandwidth is at a premium and a multicast service offers a distinct advantage. It must thus be decided how to design middleware that can provide the service appropriate to each environment. Given the separation of middleware into upperware and lowerware a possible solution could be that the upperware offer a multicast service to the application (broadcast being a special case of multicast). The lowerware could implement this functionality in different ways for different environments: In a broadcast domain it could send multicast transmissions as multicast or broadcast over the shared medium, whereas in wired environments unicast transmissions could be used.

5.5 SUMMARY AND CONCLUSIONS

This chapter has examined the implications for the design of middleware targeted at the deployment in disadvantaged grids. Middleware for the disadvantaged environment has special requirements not present in

[4] Multicast refers to a data transmission that is addressed to several recipients at once while unicast transmissions are addressed only to one specific recipient. A multicast transmission in a shared medium like radio can be picked up by all intended recipients within range of the sender with only one transmission instead of one transmission per recipient. In a wired environment on the other hand a transmission can only be received by the network node at the other end of the wire. This reduces the advantages of multicast somewhat.

the wired domain. These special requirements derive primarily from limitations of available energy at the nodes as well as very limited and unpredictable bandwidth and unreliable connectivity to the network. To design middleware that can operate effectively in this environment, these constraints must be respected throughout the design process. At the same time it is imperative to remember that the same middleware may be deployed in non-disadvantaged environments such as command posts. Middleware designed only for the non-disadvantaged environment may not operate efficiently in a disadvantaged environment, and vice versa. It is necessary to design middleware from the beginning for both environments where good performance in both environments is required.

In dedicated C2 networks where all participating nodes run only the C2 applications the middleware is often implemented as an integral part of the overall C2 system sitting on top of the network. Nevertheless it can be assumed that the functionality is present and the conclusions from this chapter are applicable.

Chapter 6 – NETWORK ISSUES

In the traditional approach to networking, user applications view the network as a service provider, and are not concerned with the characteristics of the network, as long as it can support the desired user traffic. This approach led to the development of layered network architectures, which facilitate modular network design and interoperability. Such approaches have distinct advantages for wired networks such as the Internet. However, they may be less appropriate for ad hoc wireless networks such as tactical networks. In this chapter the conventional approach of layered network design that is used in wired networks such as the Internet is analyzed. This is followed by a discussion of some characteristics of wireless networks that are markedly different from those of wired networks, characteristics which suggest that novel approaches are needed to provide good performance in wireless networks. Finally, it is shown how cross-layer techniques can be used to approach the design and control of ad hoc wireless networks.

A discussion of many of the major issues relating to cross-layer design in ad hoc wireless networks can be found in Goldsmith and Wicker [12].

6.1 LAYERED NETWORK DESIGN

A hierarchical, or layered, structure has traditionally been used to reduce a network's complexity. Figure 6-1 shows one commonly used abstraction of the layered protocol stack; it is similar to the seven-layer OSI reference model, except that it does not include the session layer (OSI Layer 5) or presentation layer (OSI layer 6).[1] In such an architectural design, each layer offers services to the layer above it. Thus, for example, the application layer, which represents the functionality the user would like to obtain from the network, interfaces directly only with the transport layer. It is shielded from the lower layer functions, and consequently it does not have to know how they are implemented. These lower layers are the network layer (which performs routing in multihop networks), the data link (or simply link) layer (which performs media access control, error control, etc.), and the physical layer (which supports communication over the specific medium, which may be wire, fibre, wireless, etc.). An excellent description of layered network structures is provided in the textbook by Tanenbaum [13].

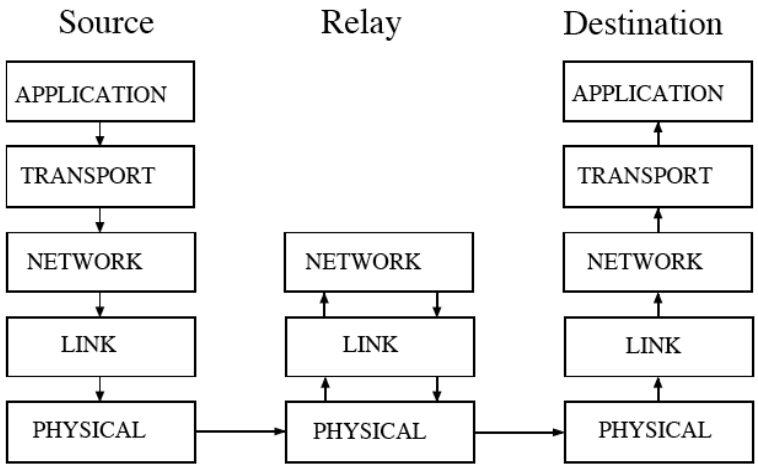


Figure 6-1: The Conventional Layered Protocol Stack.

[1] These two layers are not included in the TCP/IP protocol stack either.

Such a modular design permits the development of "open standards" that facilitate both hardware and software development. Updates to individual layers are possible without disturbing the overall network structure. Such updates may be introduced as a consequence of improved network design (e.g., improved equipment or improved algorithm design) or updated user requirements.

In the example of Figure 6-1, the network provides a service in which the application layer at the Source node communicates with the application layer at the Destination node. It does so by using the functions of the transport layer, which in turn employs the network layer, and so on down the stack. Note that the functions of relay nodes involve only the network layer and those below it.

6.2 CHARACTERISTICS OF AD HOC NETWORKS

This section identifies some key characteristics of wireless ad hoc networks that differentiate them fundamentally from wired networks, and describes how novel approaches that deviate from the rigid layered structure discussed in Section 6.1 may provide improved performance.

In addition to their obvious differences from wired networks, wireless ad hoc networks are also fundamentally different from the cellular systems and wireless local area networks (LANs) that have been developed in the commercial domain. The primary differences result from the lack of an infrastructure. For example, cellular systems have fixed base stations, which communicate among themselves using dedicated non-wireless lines; thus, the primary issues to be addressed in cellular systems involve tracking the mobile users. Otherwise, wireless cellular communication is limited to that between mobile users and base stations. In fully connected wireless LANs, since there is single-hop connectivity among all the nodes, routing is not an issue.

In ad hoc wireless networks it is possible to establish a link between any pair of nodes, provided that the signal-to-noise ratio at the receiving node is sufficiently high. Thus, unlike the case of wired networks, the set of network links and their capacities are not determined *a priori*. Factors relating to the existence of a link include:

- Distance between nodes;
- Transmitted RF power;
- Background noise;
- Data rate;
- Error-control code rate;
- Modulation scheme;
- Other-user interference; and
- Quality of service (QoS) requirements.

Thus, even when the physical locations of the nodes are fixed, many of the factors that affect network topology can be (at least partially) influenced by the actions of the network nodes. While some of the issues listed above are obvious, others may not be. For example, the interference level at a node depends not only on background noise (and possibly jamming) levels, but also on the interference caused by other nodes; thus, the mechanism used to schedule transmissions affects the interference level at nearby nodes. The specification of data rate and error-control code rate (along with the modulation scheme) affect the BER, and hence impact on whether or not the desired QoS requirement is satisfied. Perhaps more subtle is the fact that the specified QoS level

determines whether or not a link is present; reduction in the acceptable level of QoS permits the use of a link, but would be appropriate only if the user application can tolerate such a reduced QoS.

Furthermore, in ad hoc networks no distinction can be made between uplink and downlink traffic[2], thus greatly complicating the interference environment. Therefore, the wireless networking environment poses many new challenges not encountered in either wired or cellular networks.

6.2.1 Examples of Potential Cross-Layer Relationships in Tactical Ad Hoc Networks

Figure 6-2, taken from Prof. Michael Pursley's keynote presentation at the Cross-Layer workshop [14], illustrates many of the potential interactions among communication and networking functions at various layers of the protocol stack. The most obvious interactions are among the lowest three layers (physical, data link, and network). However, the impact of the higher layers is apparent as well, once the applications that must be supported are addressed.

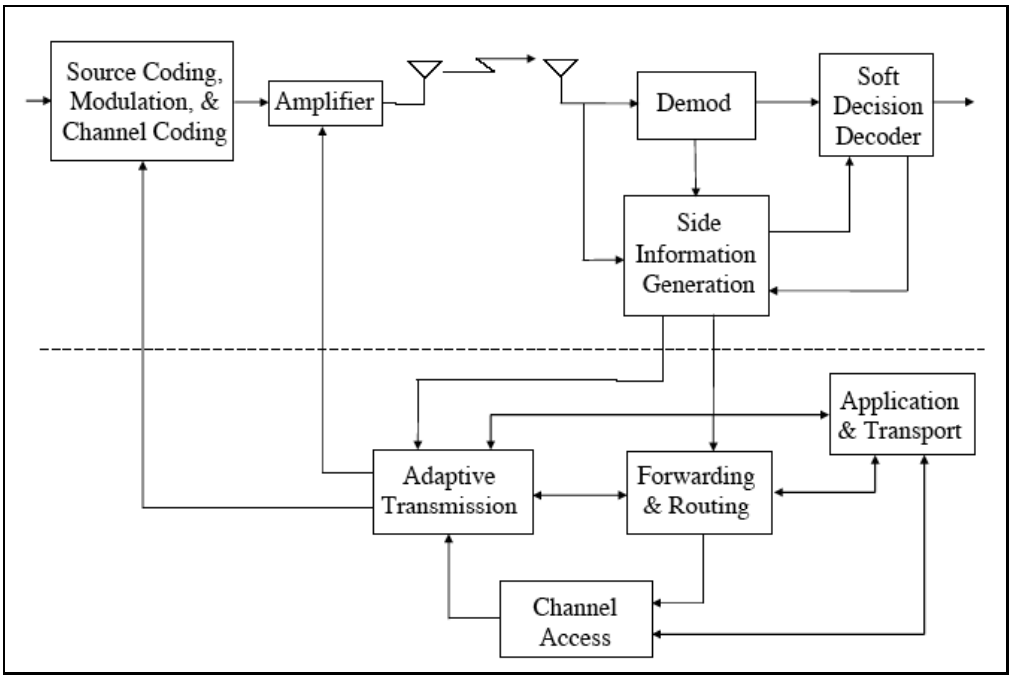


Figure 6-2: Some Protocol Interactions in Wireless Networks.

To illustrate the need for cross-layer protocols, Prof. Pursley presented the following example based on voice messaging. Voice traffic typically requires low delay to preserve its intelligibility and timeliness, but can tolerate significantly higher error rates (frame erasures) than data traffic. Moreover, voice traffic typically consists of relatively long sessions, rather than individual packets. The impact of these requirements at the various layers is summarized as follows:

[2] Cellular networks are characterized by a hierarchical structure, in which, there is an uplink between mobile users and the base station and a downlink between the base station and the mobile users. The uplink and downlink communication generally use distinct (orthogonal) channels. However, in peer-to-peer architectures such as those of ad hoc networks, a single channel is typically shared by all users, which may use multihop routes from the source node to the destination node.

- Application layer: Speech compression must match available routes and links and satisfy QoS needs (e.g., intelligible speech vs. speaker recognition).
- Network layer: Routing should emphasize the need for low delay; high-quality (low BER) routes are not needed.
- Data link layer: To accommodate a voice session, it is necessary to reserve multiple time slots on each link via the channel access (MAC) protocol. Detected packet errors may not require retransmission (because some errors can be tolerated).
- Physical layer: Code rate should adapt to link quality. Low-rate codes should be used on poor links (to avoid need for retransmissions), and high-rate codes on good links (to reduce delay). Energy conservation is secondary to need for timely delivery when considering voice traffic.

In addition to illustrating the relationships among functions at various layers, this example highlights the particular dependence on an application that involves session-based voice messaging. For example, if packet-oriented data communication were considered, it might be possible to consider the imposition of less-stringent delay requirements; however, more-stringent bit-error-rate requirements would be appropriate. Therefore, the cross-layer dependencies depend strongly on the particular application that is being supported.

When considering multimedia traffic, the various classes of traffic (e.g., voice and data) should be handled differently through the use of adaptive transmission and routing protocols, based on trade-offs such as those discussed here.

The potential benefits of cross layering are greatest in ad hoc wireless networks because of the strong interrelationships among the physical, data link, and network layers. However, some benefits may be possible in wired networks as well.

6.3 CROSS-LAYER ISSUES IN TACTICAL MILITARY NETWORKS

Some characteristics of wireless networks that distinguish them from wired networks and that suggest the potential benefits of the use of cross-layer techniques were discussed in Section 6.2.1. Several unique considerations imposed by the tactical military communications environment serve to further distinguish military ad hoc networks from both wired networks and typical commercial wireless networks. These include the following:[3]

- Tactical military equipment can support only low data rates.
- Heterogeneous equipment with different capabilities must function in the same network.
- Hostile environment (e.g., jammers, node destruction).
- Applications with very different requirements and priorities must be supported.
- Widely varying communication conditions and network topologies must be supported.
- Legacy systems must be supported while transitioning to future systems.

The authors believe that cross-layer approaches can, in fact, provide improved performance relating to at least some of these problems. For example:

[3] The two lists in this section are based largely, although not exclusively, on material from the panel discussion presentation by Prof. Andrea Goldsmith at the Cross-Layer workshop [15].

- Adaptation and diversity can provide robustness to jamming and destruction or compromise of nodes.
- Cross layering can support different requirements (e.g., voice, data) and priorities across all layers of the network protocol stack.
- Cross layering can adjust higher layer protocols to the capabilities of underlying equipment.
- Cross layering can adapt to and provide robustness against variations in the communication capabilities and network topology.
- Cross layering can allow nodes to use information obtained by one layer at a higher or lower layer as well (particularly important to permit exploitation of network status information by the application/ middleware layers).

However, it will be difficult to overcome some obstacles, such as the need to support communication with legacy systems that cannot provide the necessary degree of adaptivity. For example, new military radio systems such as the USA's Joint Tactical Radio System (JTRS) will be able to use adaptive cross-layer protocols to save energy and to improve QoS performance. However, legacy SINCGARS radios do not have the necessary degree of adaptivity to support cross-layer operation. For example, SINCGARS radios have only manual power settings, and SINCGARS ACKs do not provide the necessary physical-layer information to permit JTRS equipment to adapt intelligently in joint JTRS-SINCGARS networks. Additionally, security issues may impose significant obstacles on the exchange of some information up and down the protocol stack.

6.4 THE IMPACT OF ENERGY-RELATED CONSIDERATIONS

Energy-awareness is a crucial aspect for those ad hoc or sensor networks where the nodes are powered by batteries. For example, the batteries carried by a soldier may constitute a significant fraction of the overall load that must be carried. Therefore, means to reduce energy consumption are extremely desirable. The traditional approaches to energy reduction involve the development of energy-efficient electronics and energy-efficient modulation schemes, as well as the use of directional antennas (which focus the beam in the desired direction, thereby eliminating the transmission of energy in unnecessary directions). Additionally, energy awareness can be viewed from the networking perspective, which involves multiple layers in the protocol stack.

Two basic forms of energy-aware network operation can be considered. Under "energy-efficient" operation, the goal is to maximize the number of bits that are delivered per unit energy over a period of time. In this mode of operation, energy use may be considered as a cost (e.g., the cost of replacing the batteries). However, in some applications batteries cannot be replaced during the course of a mission. Such a situation can arise when soldiers are unable to return to base, or alternatively in the case of sensor networks. This case is known as "energy-constrained" operation.

It is important to note that energy-efficient operation does not ensure good performance in energy-constrained applications. For example, use of the most energy-efficient routes may result in premature depletion of energy at some nodes.

The issue of energy awareness (in both its energy-efficient and energy-constrained forms) crosses several layers of the protocol stack. One obvious trade-off is that of energy used for signal processing versus that used for communications. For example, signal-processing algorithms that significantly compress the data can provide benefits to overall network operation by reducing the number of bits that must be transmitted, thereby saving RF energy and reducing bandwidth requirements. However, the reduction in energy may be outweighed by the increased energy needed for data compression and decompression operations. Energy

consumed by the nodes' hardware is an especially significant component of overall energy consumption in short-range networks, where RF transmission energy is relatively low. Furthermore, the complicated nature of energy-related trade-offs is illustrated by the fact that the amount of energy consumed by hardware can be reduced by reducing the bit duration of the transmitted symbol, whereas RF energy can be reduced by doing the opposite. Therefore, trade-offs between energy and delay must be considered.

Energy may be saved by the use of sleep modes, because nodes consume energy even when they are not transmitting or receiving. However, use of such sleep modes complicates many aspects of networking, including synchronization, routing, channel access, sensing functionality, etc. Another approach to energy-aware operation involves controlling transmitted power levels, which are a key factor in determining both connectivity and interference levels. Additionally, modulation, coding, and data rate are key factors in determining whether or not a reliable link is present. Therefore, there is a need to coordinate functions at several layers of the protocol stack.

The bottom line is that the introduction of energy considerations, especially in energy-constrained operations, results in fundamental changes to the considerations that need to be addressed in ad hoc network design. Most importantly, it introduces trade-offs among performance measures such as delay, throughput, and node/ network lifetime, and necessitates tight coupling among the layers if near-optimal performance is to be obtained.

6.5 CROSS-LAYERING VS THE CONVENTIONAL LAYERED MODEL

The use of cross-layering techniques does not mean that the layered architecture should be abandoned. To the contrary, the layered architecture has worked well in the Internet, and its modular structure (see Section 6.1) provides an efficient and scalable framework for network design. Nevertheless, significant performance improvement can be expected if cross-layer techniques are implemented in ad hoc networks.

Thus, cross-layer design is not about eliminating layers, but is rather about designing across them. Wireless networks can benefit most from cross-layer design, but benefits are possible for wired networks as well. The degree of improvement that can be achieved depends strongly on the type of network. For example, sensor networks are expected to benefit more from cross layering than general mobile ad hoc networks (see Section 6.6). The research community is only beginning to understand the nature of cross-layer design, and still needs to determine where significant cross-layer gains are possible.

Despite the potential benefits of cross layering, caution is needed to avoid the possible unintended consequences of some cross-layer interactions.[4] For example, the tight coupling of layers may lead toward a tendency to develop proprietary protocols, and hence the need to redesign a new system for every application, thereby eliminating many of the benefits originally obtained by layering. In addition, tight coupling of layers may lead to "spaghetti code" in which patches are continually added to improve performance, resulting in a system that is difficult to understand and hence difficult to update to accommodate changing requirements. Furthermore, the performance of an optimally designed system may be highly sensitive to the operating point, and minor errors in system parameters or minor environmental or topological changes may result in significant performance degradation from that at the optimal operating point. Finally, there is no consensus in the research community as to whether use of cross-layer optimization would increase vulnerability to attacks by intelligent adversaries, or, to the contrary, improve network robustness against such threats. It is expected

[4] See a recent paper by Kawadia and Kumar [16] for a discussion of such issues.

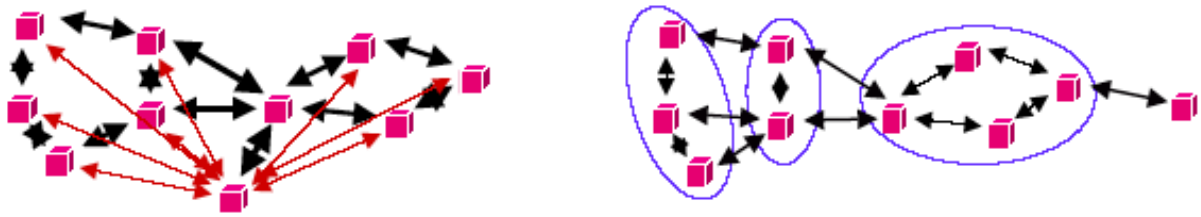
that the answer to this question will depend on the specific application and on the layers involved in the optimization.

Notwithstanding the above cautions, the authors concluded that further research and development of cross-layer techniques is important because of potential benefits that may be difficult or impossible to achieve by other means. The following list, which was adapted from [15], addresses key issues in cross-layer research and development:

- Development of the right framework for cross-layer design;
- Determination of information to be exchanged across layers, and how to use it;
- Balancing of adaptivity, diversity, and scheduling;
- Identification of the key cross-layer synergies, and which layers should be involved;
- Avoidance of unexpected interactions across layers;
- Management of cross-layer complexity; and
- Accommodation of legacy systems and protocols.

6.6 SIMILARITIES AND DIFFERENCES BETWEEN MOBILE AD HOC NETWORKS AND SENSOR NETWORKS

Wireless ad hoc networks and sensor networks are similar, in that both classes of networks are infrastructureless and use the wireless channel, as shown in Figure 6-3. In fact, in many ways sensor networks may be viewed as a special case of ad hoc networks. However, there are some significant differences between these types of networks, typical characteristics of which are summarized in Table 6-1, which was adapted from [15]. Although the Task Group's focus was ad hoc networks, rather than sensor networks, a comparison of ad hoc and sensor networks is useful in understanding the role of cross layering in ad hoc network design and control.



(a) Wireless ad hoc network (b) Wireless sensor network

Figure 6-3: Examples of Ad Hoc and Sensor Networks.

Table 6-1: Typical Characteristics of Ad Hoc Networks and Sensor Networks

Ad Hoc Networks	Sensor Networks
• Peer-to-peer with no backbone infrastructure	• Data flows to a centralized location
• Nodes typically mobile	• Nodes typically stationary
• Network size: up to tens of nodes	• Network size: hundreds or thousands of nodes
• Nodes can be well-equipped	• Nodes typically have limited capability
• Nodes generate independent information	• Node information correlated in time and space
• Can require high data rates	• Low per-node rates, but large number of nodes
• Typically support multiple applications	• Typically support a single application
• Batteries can usually be recharged or replaced	• Nodes typically energy-constrained

Although both types of networks can benefit from cross-layer design, it is felt that sensor networks stand to benefit more. One of the primary concerns in sensor networks is that nodes are typically energy-constrained, as discussed in Section 6.4. (Some examples of mobile ad hoc networks, such as those in which the nodes consist of individual soldiers, are also energy constrained because it may not be possible to replenish batteries during the course of a mission.) Therefore, it is essential that energy use be optimized across the protocol stack. Additionally, the fact that sensor networks are generally designed for one dedicated purpose (target detection, surveillance, etc.) permits a tightly coupled design in which only those functions that are necessary need to be supported. By contrast, ad hoc networks may be more constrained by existing standards and the need for interoperability, constraints which make effective cross-layer design more challenging.

6.7 SUMMARY AND CONCLUSIONS ON NETWORKING ISSUES

To fully exploit the wireless channel, some use of cross-layer techniques will likely be necessary in future wireless network applications, both in commercial and military environments. Appropriate use of cross-layer techniques would not involve the abandonment of the layered protocol structure; rather, cross layering would be used to augment the network's capability by sharing information among the layers and by jointly optimizing their performance. There appears to be considerable potential for performance improvement. Nevertheless, this field is still in its early stages of development, and the research community does not yet have sufficient insight to understand the big picture. One fundamental question, which has not yet been answered, is that of which layer interactions provide the best opportunities for performance improvement. The research community is just starting to ask the right questions, and there is now a basis for fruitful research and development.

The wireless networking environment, particularly in the case of ad hoc networks, is quite different from that of wired networks. Consequently, the properties of the physical layer play a large part in ad hoc network design and performance, and cross-layer design techniques are especially well suited for them. Although cross-layer techniques may provide some benefits in wired networks, their greatest benefits are expected to be obtained in ad hoc wireless networks and sensor networks.

Energy concerns are extremely important in ad hoc networks, and especially in sensor networks (which may be viewed as a special case of ad hoc networks). The fact that a finite quantity of energy must be shared

among all of a node's functions (e.g., transmission, reception, and signal processing) strongly links virtually all layers of the protocol stack. In fact, it may now be appropriate to consider a "hardware layer," which functions under the physical layer. The fact that sensor networks have severe constraints on energy makes them especially good candidates to benefit from cross-layer design.

Despite the potential benefits of cross-layer approaches, a degree of caution is needed in their application. For example, an obstacle to the use of cross-layer approaches is the need to accommodate legacy systems, which may not have the capabilities to implement such control functions. Additionally, excessive coordination among the layers may result in the design of special purpose systems, thereby eliminating some of the advantages of layered design. The result could be unwieldy systems that are difficult to understand, and hence difficult to update to accommodate changing requirements. Furthermore, attempts to optimize performance may result in a system that is overly sensitive to parameter values, thereby running the risk of poor performance unless all parameters are perfectly tuned (which is virtually impossible in practice).

In conclusion, further research and development of cross-layer techniques for tactical ad hoc networks should be pursued to better exploit the characteristics of the wireless communication medium. However, in doing so it is necessary to take into consideration the impact of military requirements (such as the need to support legacy systems) as well as to understand and avoid the unintended consequences that may result from attempts to optimize network functionality at several layers simultaneously.

Chapter 7 – CANADIAN EXPERIMENTS USING LOW BANDWIDTH TEST BED

This chapter presents some results of simulations conducted under a technology demonstration project of Defence R&D Canada entitled 'High Capacity Tactical Communications Networks' using a simulation environment called the Low Bandwidth Test Bed (LBTB) [17]. The simulations were conducted to examine the impact of application-layer information management techniques on the throughput of operationally important information over a congested tactical radio subnet. The results illustrate the positive impact that some techniques discussed in Chapters 4 and 5 of this report can have on information flow over disadvantaged tactical communication grids.

7.1 DESCRIPTION OF LOW BANDWIDTH TEST BED

The Low Bandwidth Test Bed is a discrete event simulator that couples a performance model of a tactical combat net radio system to simulated tactical nodes (STNs) containing real databases. It supports two modes of operation: distributed and integrated mode.

In distributed mode, up to 32 computers can be employed. One of the computers acts as the controller for the simulation while a number of STNs can be defined on each of the remaining computers. There is no limit to the number of STNs per machine. In integrated mode, the entire simulation environment (control element and all defined STNs) resides on, and is executed on, a single machine. This option permits execution of multiple simulations simultaneously, but each simulation executes more slowly.

To provide the capability to predefine the identity of the participating nodes, as well as the timing and nature of the communication events emanating from each node in a master communication event script, a Scenario Script Management Tool was integrated into the test bed. At the time of simulation initialization, this script is propagated to each simulated node. When the simulation begins, the scripts are executed simultaneously on each node under the direction of the control node, which advances scenario time in a synchronized fashion at one-second intervals. Each communication event in the script mimics the creation of a tactical message report on one of the local STNs and its replication to the other STNs. When a script event is executed, the template maps the pre-defined data fields for that event to the local database tables. Templates for ten tactical message types have been implemented, but for the experiments described in this report, the only message type employed was the Own Station Position Report. The mechanisms for exchanging this information with other nodes are discussed below. At any time, the simulation can be paused, the databases locked, and the database tables on each node interrogated to populate the measurement logs. Once execution of the script has been completed, the data in the measurement log tables are analyzed to correlate the quality of the shared tactical picture (as represented by the database content on each STN) with the time-varying behaviour of the simulated tactical network.

Information exchange in the Low Bandwidth Test Bed is accomplished via a custom data replication mechanism that replicates data between databases on participating network nodes. Replicated payload content for a particular script event is based upon the information content of the tactical message type associated with that event in the script. Two custom replication mechanisms are available:

a) Rule-Based Replication Implemented using Triggers and Stored Procedures

   When a particular event instance is encountered in the event script, a stored procedure is called to prepare the replication payload. At the same time, a separate stored procedure is called containing the

control logic that decides whether to permit or suppress replication of the replication payload associated with that event instance. This mechanism is event-driven and permits the use of information management rules that provide fine-grained context-sensitive control over the replication process based upon knowledge of battlefield state and network state.

b) Scheduled Replication

In contrast to the first mechanism, the second mechanism is time-driven. This mechanism approximates closely the incremental update employed by the ATCCIS Replication Mechanism (see Annex D). The mechanism employs a selective negotiated 'push' of information on a periodic basis from a provider node to one or more receiver nodes. The information content to be pushed is specified in a 'contract'. The contract specification consists of the set of database tables from which data will be extracted on the provider node, one or more filter criteria to specify which rows from those tables will be selected, the identity of the receiver nodes, and the wait period between successive contract fulfilments. This mechanism does not provide the context-sensitive control of replication available with rule-based replication.

For the experiments reported in this chapter, only the rule-based replication mechanism was employed.

Underpinning each of these mechanisms is a Replication Transport Layer (RTL) with features designed for the bandwidth-constrained wireless domain (Figure 7-1). These features include a system of priority output queues and advanced load-balancing mechanisms. As well, it uses a multicast protocol, implements a packet recovery mechanism based on negative acknowledgement, and can manage transmissions, requests for retransmission, and retransmissions. The RTL interfaces to the transport layer through a User Datagram Protocol (UDP), which acts as a pass-through to the network layer. UDP is used in place of TCP because, in an environment with bit errors and long latency, TCP's congestion controls and timeouts seriously degrade throughput. The RTL provides some functionality such as packet sequencing normally provided by TCP.

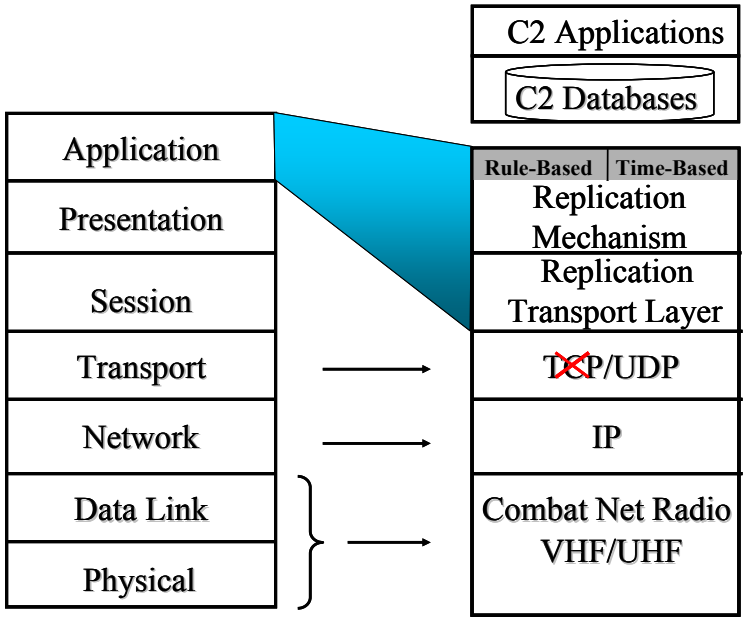


Figure 7-1: Position of Custom Replication Mechanisms in Network Protocol Stack.

The Low Bandwidth Test Bed simulates in software a single-hop subnet of combat net radios (i.e., all radios on a common frequency; multi-hop routing between nodes does not occur). The simulator implements a performance model of the network layer and the data link layer. The term 'performance model' refers to the fact that, for certain functionality, the functionality itself is not fully implemented, but the effect that that functionality would have if it were present is accurately modelled. The Media Access Control functionality for the data link layer is based on Carrier Sense Multiple Access with Collision Avoidance (CSMA/CA).

Forward error correction capability modelled in the data link layer includes a majority vote detection (MVD) scheme[1] and Golay coding[2]. Simulations can be executed with either or both error correction mechanisms activated or deactivated. Application of the majority vote detect algorithm effectively multiplies the size of the transmitted protocol data unit by a factor *n*, while the application of Golay coding effectively doubles the size of the transmitted protocol data unit.

At the radio physical layer, the radio channel model uses a probability function to generate bit errors on the channel during transmission. One of several channel profiles can be selected. One profile permits perfect channel conditions ('all-pass') or extremely bad channel conditions ('all-fail') to be specified. A second profile permits the overall percentage of bit errors or packet errors to be specified. The model does not have the ability to model transmission on a link-by-link basis. Rather, the model assumes that, for a given receiving node, the channel profile that applies to received transmissions is independent of the identity of the sending node. The radio model also supports two modes of operation – 'data only' or 'mixed voice and data'. In mixed mode, voice transmissions can pre-empt data transmissions.

All experiments described in this report used the following characteristics:

- 'Data-only' mode of operation;
- Overall percentage of packet errors (0%, 10%, 20% and 30% packet loss);
- Majority Vote Detection with $n = 5$; and
- Golay FEC deactivated.

7.2 OVERVIEW OF EXPERIMENTS

The experiments described in this chapter are based on a scenario in which 15 tactical nodes participating on a single tactical subnet with a common assigned frequency exchange updates of their own position at an interval of 60 seconds, over a period of 60 minutes. Eleven of the nodes are moving at a constant speed of 40 km/hr while four of the nodes are moving very slowly (2 km/hr). The differential in speed was selected to serve the requirement of an information management rule, described below, which selectively suppressed replications of position updates based on vehicle displacement. One of the important measures described below, position error, is directly proportional to vehicle displacement, and so is sensitive to both vehicle speed and trajectory. To eliminate distortions due to varying vehicle speed and trajectory, a given node was assigned the same speed in each experiment, the vehicle trajectories were defined as parallel straight lines, and the vehicles were all moving in the same direction. The rule-based replication mechanism was employed for all experiments.

[1] In a majority vote detection scheme, each bit of real data is duplicated an odd number of times *n* in the transmission. At the receiving node, the values of the duplicate bits are compared. If, due to transmission error, not all *n* duplicate bit values agree, the value that is in the majority is assumed to be the correct value.

[2] Golay coding is a form of forward error correction coding in which, for a number *m* of data bits, *k* additional bits are added to the original *m* bits. These *k* bits are used to detect and correct errors in the original *m* bits on a receiving node.

The objective of the experiments was to demonstrate how information management protocols operating at the application layer can improve the throughput of operationally important information over a tactical radio subnet under conditions of network congestion (i.e., when the offered load to the network exceeds the network's capacity, resulting in reduced transmission success rate and high latency). Two approaches are demonstrated:

1) Dynamic reduction of the offered load through use of an information management rule to selectively suppress transmissions; and
2) Reduction in payload size to reduce transmission time on the radio channel.

7.2.1 Dynamic Reduction of Offered Load through Use of an Information Management Rule

The rule-based replication mechanism described in Section 7.1 causes a stored procedure to be called on a local node whenever a replication message is generated on that node. The stored procedure contains the information management rule that determines whether or not to suppress replication of that replication message. The particular information management rule employed for these experiments is predicated on the supposition that it may be advisable for stopped or slowly moving nodes to broadcast their position updates less frequently than the faster-moving nodes when the network is congested. Specifically, the rule implements the following logic:

> "If the cumulative distance travelled by the local node since the last permitted replication of a position update is less than *x* meters, suppress transmission of the present position update."

For the experiments, a value of $x = 600$ meters and a replication interval of one minute were selected. This means, that for the eleven nodes moving at 40 km/hr, the vehicle displacement satisfied the rule each time and replication of the position update was allowed. On the other hand, for each of the remaining four nodes that were moving at only 2 km/hr the cumulative displacement exceeded 600 meters only once every 18 minutes. Replication of position updates from these nodes was thus suppressed 17 out of 18 times.

The LBTB permits exactly the same scenario to be repeated under exactly the same simulated radio channel conditions first with, then without, the IM rule active. Analysis showed that the overall effect of using this IM rule in this particular scenario was to reduce the offered load to the network by 25%. The results of simulations with and without the IM rule active are reported in Section 7.3.2.2.

7.2.2 Reduction in Payload Size through Choice of Payload Format

The second set of experiments examined the impact of reduced payload size on the throughput of operationally important information. Two types of payload were examined:

a) ODB PDU

   The format for this payload type conforms to the format of a Replication Manager Data PDU defined in the specification for the ATCCIS Replication Mechanism (now known as a 'MIP Data PDU' and used by the MIP Data Exchange Mechanism). The Canadian version of this Data PDU is termed an 'Operational Database (ODB) PDU'. Data exchange between ODBs exploits the 'ODB Data Model', which is based on the MIP C2IEDM data model with some Canadian extensions. The version of C2IEDM known as 'Generic Hub 4' is used in the LBTB.

b) Simple Format

   This format was developed to support the LBTB design by the LBTB contractor. It is more compact than the ODB PDU format. The Scenario Script within the LBTB consists of a series of communication

events; each event corresponds to the generation of a tactical message report on a specific node, and the replication of that report to other nodes, at a specific scenario time. When a script event is encountered during a simulation, a template corresponding to the report type for that event takes the pre-defined data for the event stored in a file and inserts the data into the appropriate tables in the database on the local Simulated Tactical Node. The same template is then used to extract the just-inserted data from the database to compose the replication message. The Simple Format exploits the fact that, within a template for a particular report type, the relevant set of database tables, the referential relationships between those tables, and the order of data insertion, are well defined. If the same template also exists on the receiving nodes, this information can be exploited to parse and insert the received data into the database tables on the receiving node as well. Information about table identities, order of insertion, etc. does not have to form part of the replicated information because it is already part of the template. What need to be transmitted over the network are the data values in appropriate order and format, and a call to the stored procedure on the receiving node that contains the appropriate template.

Elimination of metadata about database structure permits creation of a much smaller transmission payload. The same position update information that requires 1134 bytes to transmit in ODB PDU format requires only 161 bytes in Simple Format. It should be noted that, in using the Simple Format, one is trading off generality to achieve reduced payload size. The contracting mechanism that is the basis of the ARM is, by design, general in nature. There are few restrictions on the identity of the nodes that establish a contract, and on contract content. This generality is deliberate. The ATCCIS/MIP efforts are focused on achieving interoperability in the context of a NATO coalition. However, the generality comes at a price. Table identities that cannot be predefined, or defined by reference, must be transmitted each time, resulting in larger payloads. In using the Simple Format, one accepts that data exchange will occur only through pre-defined templates with a well-defined structure (the same principle behind the use of structured messages). The loss of generality associated with use of the Simple Format may not be acceptable in all situations. However, in cases where the trade-off is acceptable, the results presented in Section 7.3.2.1 suggest that considerable performance gains may be achieved over tactical communication nets.

7.2.3 Reduction in Payload Size through Use of Data Compression

The effect of reducing payload size through use of a classical lossless data compression algorithm was also examined. Such algorithms achieve compression through detection and removal of redundancies in the data. zlib (compression level 6)[3], the lossless compression algorithm mandated for use by participants in MIP, was implemented in the LBTB and simulations were carried out employing ODB PDU or Simple Format, with and without zlib compression. The average compression factors achieved with this algorithm are listed in Table 7-1:

Table 7-1: Data Compression Achieved with zlib (Compression Level 6)

Payload Format	Payload Size (bytes)		
	Normal	Compressed	Compression Ratio
ODB PDU	1134	250	4.7
Simple	161	117	1.4

[3] Open-source lossless data compression algorithm (http://www.zlib.net). Nine levels of compression are available, offering different tradeoffs between compression level and execution time. Default is level 6.

7.3 ANALYSIS/INTERPRETATION OF RESULTS

This section describes the measures of performance employed and presents the experimental results.

7.3.1 Measures of Performance

7.3.1.1 Location Fidelity

The location fidelity measure is based upon the measurement of the difference between the last reported position of a neighbour node and its actual position within a particular node's database at a particular time. For analysis purposes, the 60-minute scenario duration was divided into one-minute time slices. For each receiving node the error in the position of each of the other (sending) nodes was calculated at one-minute intervals. These errors were then averaged over all sending nodes. For each time slice, an average over all receiving nodes was then performed to produce a network-averaged position error for that time slice. For statistical validity, each simulation was repeated ten times, each time with a different random number seed. The same analysis was carried out for each simulation. For each time-slice, the network-averaged position error from each of the ten simulations was statistically averaged and an expected error in the average value based on the standard deviation was calculated. Plots of network-averaged position error vs. scenario time revealed that position error value achieved a plateau. A time interval lying within this plateau region was selected, and the average value of network-averaged position error over this time interval was calculated. It is this time-averaged and statistically averaged value for network-averaged position error that is quoted in the charts that follow.

7.3.1.2 Currency

The currency measure is based upon the time elapsed since the last successful update of a given piece of information. From data logged during the simulation, an average currency value for position updates, averaged over all receiving nodes and all scenario time, was calculated for each sending node. A network-wide average for currency was then calculated by averaging over the sending nodes[4]. For statistical validity, each simulation was repeated ten times with a different random number seed. A network-averaged currency value was calculated for each simulation. These ten values were statistically averaged and an expected error in the average value based on the standard deviation was calculated. It is this time-averaged and statistically averaged value for network-averaged currency that is quoted in the charts that follow.

7.3.1.3 Latency

The latency measure is based upon a measurement of the actual transit time of a replication message through the radio network from a sending node to a receiving node. The measurement is defined as the difference between the time that a replication message is submitted to the network layer of the protocol stack on the sending node to the time that that same replication message is received, intact, by the network layer of the protocol stack on a receiving node, ready to be passed to the application layer on the receiving node. Latency measurements are recorded only for transmissions that are successfully received. From data logged during the simulation, an average latency value averaged over all receiving nodes and all scenario time was calculated for each sending node. A network-wide average for latency was then calculated by averaging over all sending

[4] For the experiments reported in this chapter, the Information Management rule employed had the effect of suppressing replication of position updates from four of the 15 nodes, thereby increasing the currency for position updates from these four nodes. The decrease in currency for the fast-moving nodes (due to improved network conditions) was masked by the rule-induced increase in currency for the slow-moving nodes. To eliminate this distortion, the currency averages were performed only over the 11 fast-moving sending nodes (all 15 nodes were still counted as receiving nodes).

nodes. For statistical validity, each simulation was repeated ten times with a different random number seed. A network-averaged latency value was calculated for each simulation. These ten values are statistically averaged and an expected error in the average value based on the standard deviation is calculated. This statistically averaged value for network-averaged latency is quoted in the charts that follow.

7.3.2 Experimental Results

For all of the simulations reported in this chapter, strict control of experimental variables was exercised. To compare the impact of a particular variable (e.g., payload format), only that variable was varied from one experiment to the next. The results with and without use of that variable were then compared. Each experiment was performed for four different sets of channel conditions – all pass (no lost packets) and three different levels of packet loss (10%, 20%, and 30%). The variable in question was considered to have a positive influence if a reduction in any or all of the measured quantities, position error, currency, and latency, was observed when that variable was employed. To minimize variability arising from the stochastic nature of the channel access protocol and the transmission process a given simulation was repeated ten times with a different random number seed, and the results averaged as described in Section 7.3.1.

7.3.2.1 Effect of Payload Format

Figure 7-2 compares the network-averaged values of position error, currency and latency obtained for four different packet loss rates with the ODB PDU (size 1134 bytes) and Simple Format (161 bytes). It is clear that, for all packet loss rates, the smaller payload size offered by the Simple Format results in a reduction of all three quantities. Figure 7-3 presents the reductions in position error, currency and latency due to use of the Simple Format, expressed as a percentage of the value obtained with the ODB PDU format. Percentage reductions are significant in all cases, lying in the range of 40% to 70%.

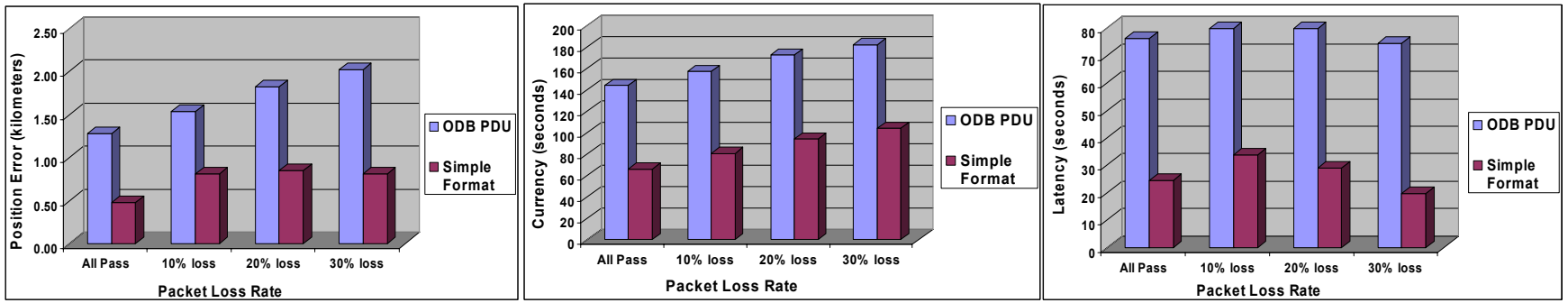


Figure 7-2: Network-Averaged Position Error, Currency and Latency for Different Payload Formats.

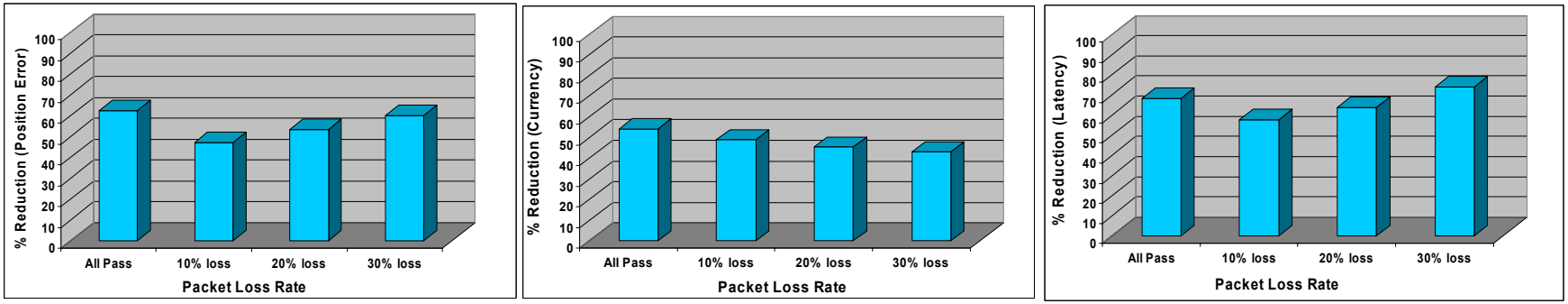


Figure 7-3: Percent Reduction in Network-Averaged Position Error, Currency and Latency Due to Use of Simple Payload Format.

7.3.2.2 Effect of Information Management Rule

Figure 7-4 compares the network-averaged values of position error, currency and latency obtained for four different packet-loss rates without, and with, the IM Rule described in Section 7.2.1 (use of the IM Rule has the overall effect of reducing network traffic by 25%). The first tier of charts corresponds to the ODB PDU payload format. The second tier of charts corresponds to the Simple Format. Figure 7-5 presents the reductions in position error, currency and latency due to use of the IM Rule, expressed as a percentage of the value obtained without use of the IM Rule. For the ODB PDU format, reductions in the range of 30 – 40% are obtained except for the case of 'All Pass', where the reductions are in the range of 60 – 95%. For the Simple Format, the reductions due to the smaller payload size (Section 7.3.2.1) are apparent (Figure 7-4). For non-zero packet loss rates, the IM Rule results in further reductions of 20 – 45% for position error and 25 – 60% for latency (Figure 7-5), with the size of reduction being smallest for the poorest channel conditions. For the 'All Pass' condition, marked reductions of 80% and 90% are observed for position error and latency respectively. Currency shows a more modest reduction of 10 – 20%, because its value is already close to the minimum (60 seconds).

7.3.2.3 Effect of Data Compression

Figure 7-6 compares the network-averaged values of position error, currency and latency obtained for four different packet loss rates without, and with, use of the zlib lossless data compression algorithm. The first tier of charts corresponds to the ODB PDU payload format. The second tier of charts corresponds to the Simple Format. Figure 7-7 presents the reductions in position error, currency and latency due to use of data compression, expressed as a percentage of the value obtained without use of data compression. For the ODB PDU format, reductions of 50 – 70% are observed for position error and latency, while reductions of 30 – 40% are observed for currency. For the Simple Format, the reductions due to the smaller payload size (Section 7.3.2.1) are apparent (Figure 7-6). When data compression is employed, small reductions, in the range of 0 – 15% for position error and latency, and 0% for currency, are observed (Figure 7-7). The small reductions can be attributed to the fact that data compression improves throughput by reducing payload size, but most of this advantage has already been achieved by the choice of the smaller payload format.

7.3.2.4 Combined Effect of Information Management Rule and Data Compression

Figure 7-8 compares the network-averaged values of position error, currency and latency obtained for four different packet loss rates without, and with, combined use of both the IM rule and zlib lossless data compression algorithm. The first tier of charts corresponds to the ODB PDU payload format. The second tier of charts corresponds to the Simple Format. Figure 7-9 presents the reductions in position error, currency and latency due to combined use of IM rule and data compression, expressed as a percentage of the value obtained when no IM rule or data compression are employed. For the ODB PDU format, reductions of 65% to 95% are observed for position error and latency, and 50% to 60% for currency. For the Simple Format, reductions due to smaller payload size (Section 7.3.2.1) are apparent (Figure 7-8). For non-zero packet loss rates, combined use of IM rule plus data compression results in further reductions of 20 – 50% for position error and 20 – 60% for latency (Figure 7-9), with the size of reduction being smallest for the poorest channel conditions. For the 'All Pass' condition, reductions of 80% and 90% are observed for position error and latency respectively. Currency shows a more modest reduction of 10 – 20%, because its value is already close to the minimum (60 seconds).

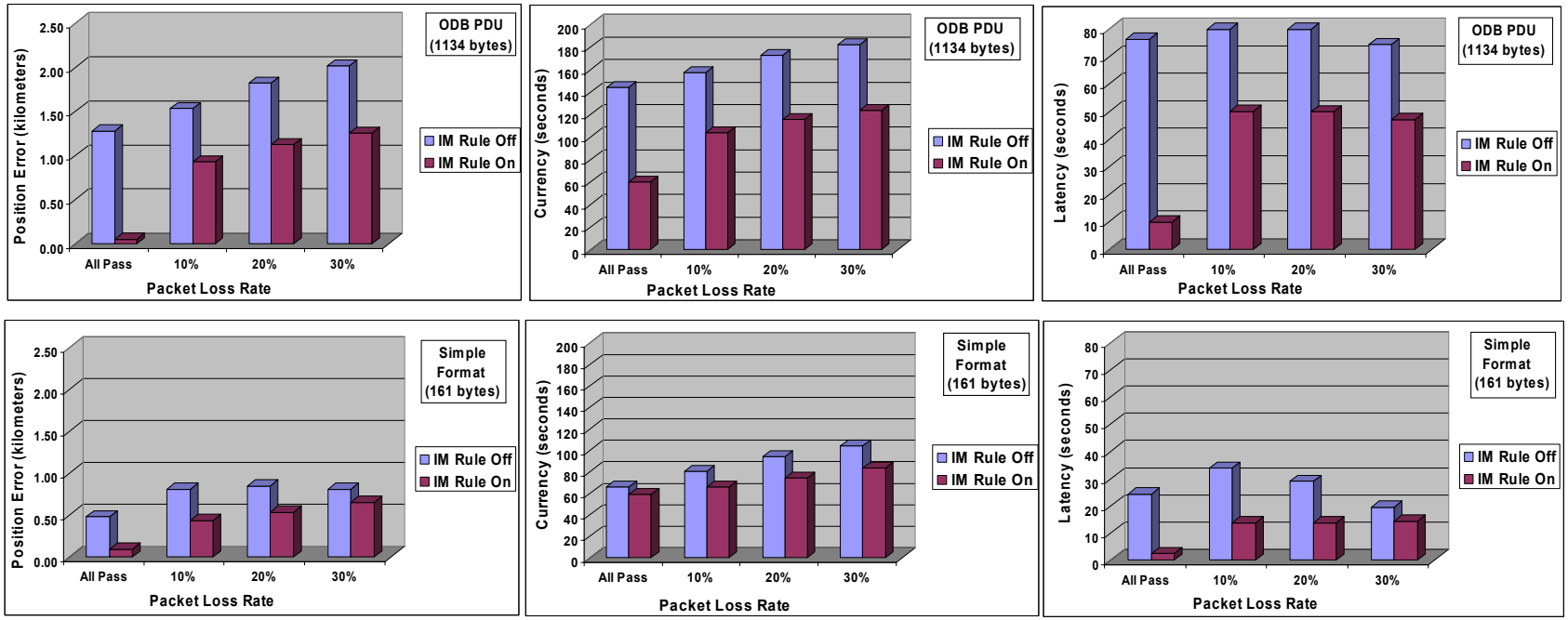


Figure 7-4: Effect of Information Management Rule on Network-Averaged Position Error, Currency and Latency.

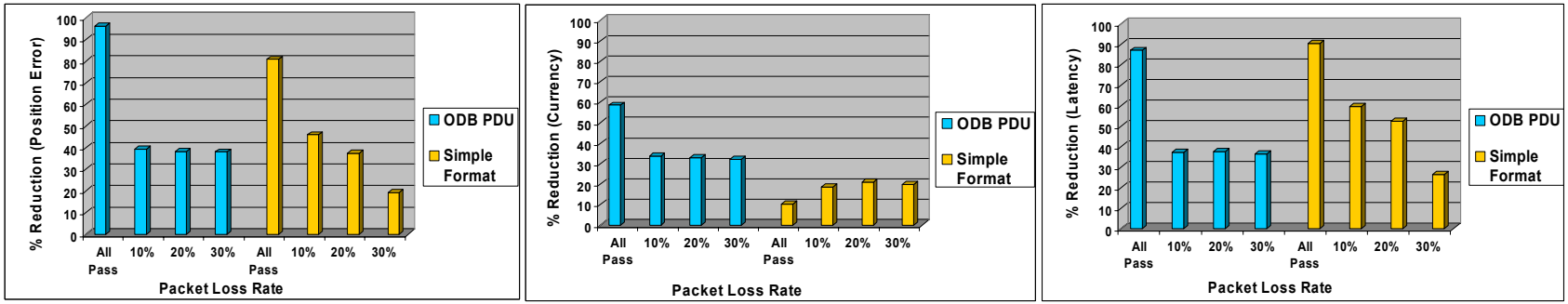


Figure 7-5: Percent Reduction in Position Error, Currency and Latency Due to Use of Information Management Rule.

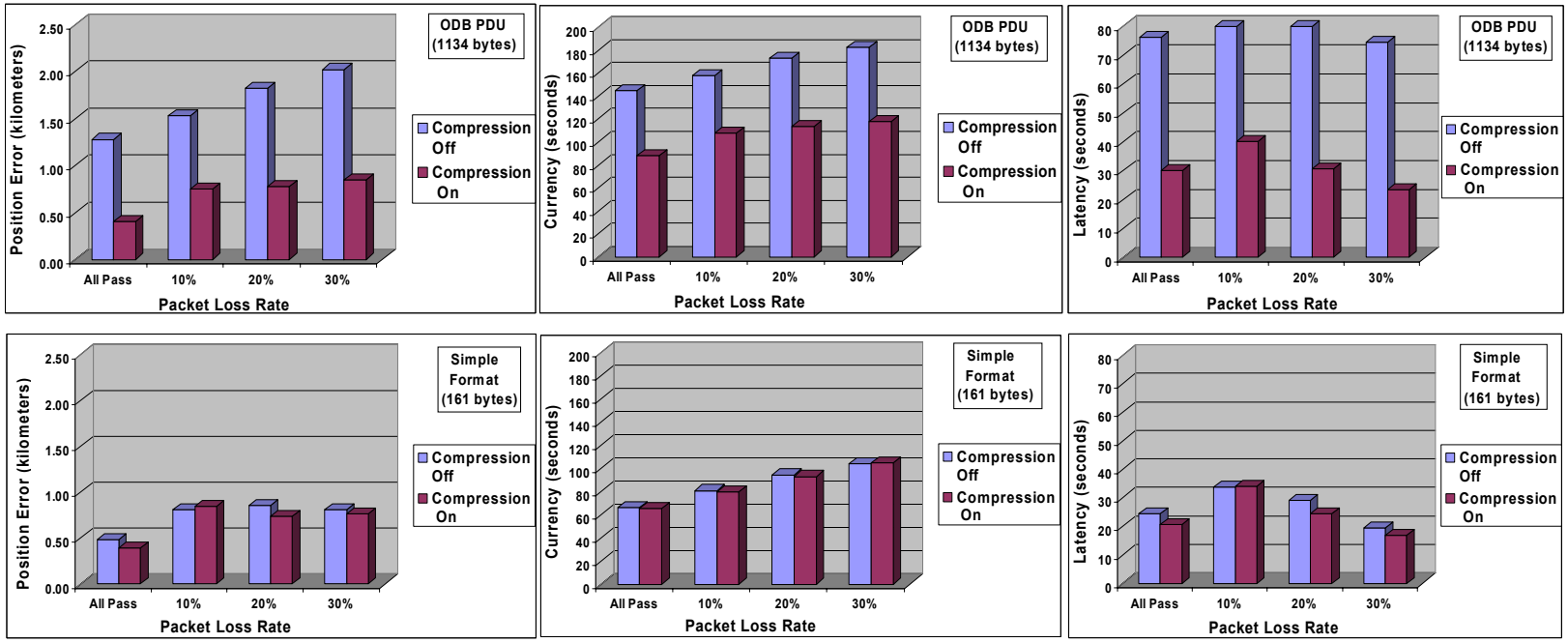


Figure 7-6: Effect of Data Compression on Network-Averaged Position Error, Currency and Latency.

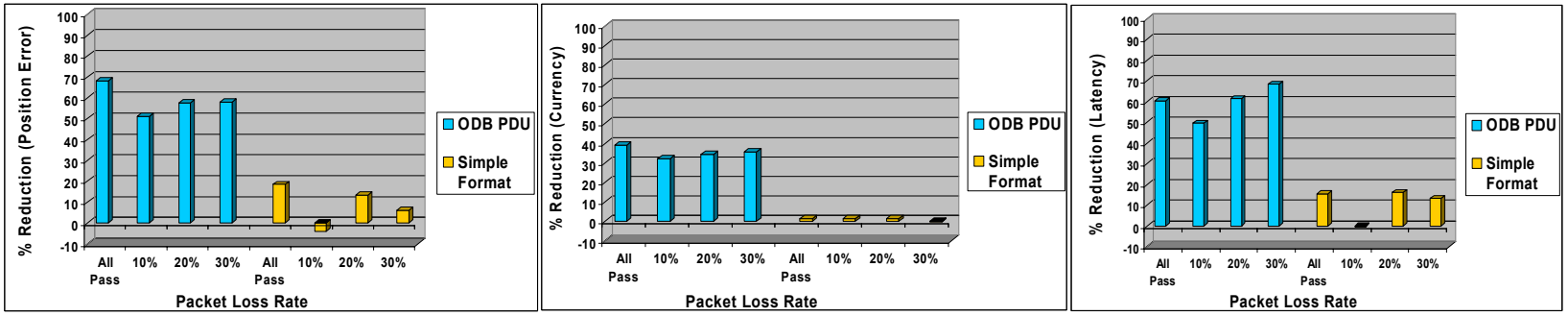


Figure 7-7: Percent Reduction in Position Error, Currency and Latency Due to Use of Data Compression.

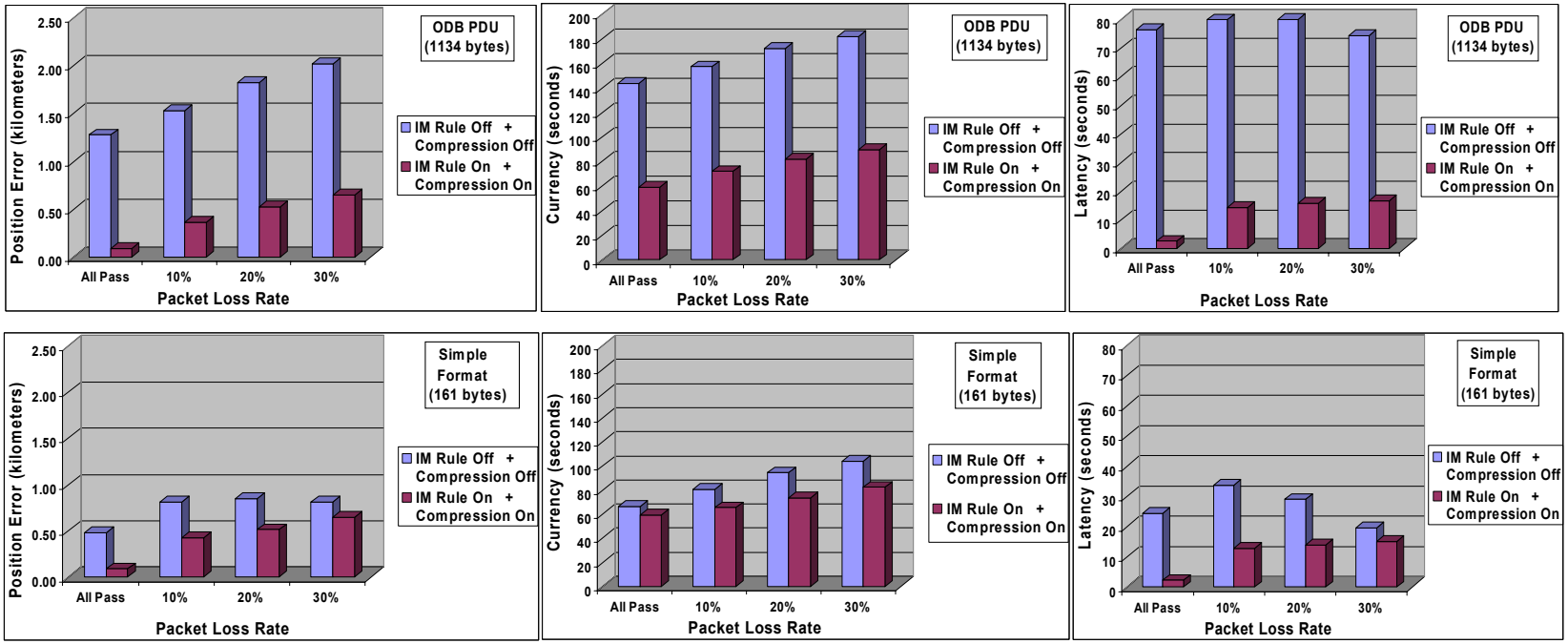


Figure 7-8: Combined Effect of IM Rule and Data Compression on Network-Averaged Position Error, Currency and Latency.

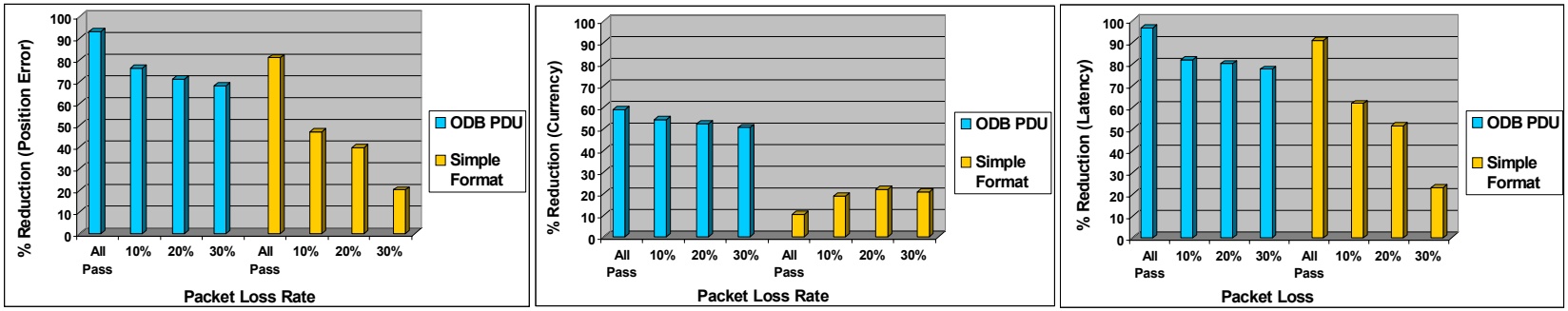


Figure 7-9: Percent Reduction in Position Error, Currency and Latency Due to Combined Use of IM Rule and Data Compression.

7.3.2.5 Summary of Results

For the ODB PDU format, both the IM rule and the data compression, used alone, result in significant reductions in network-averaged position error, latency and currency. The reductions achieved when employed together are greater than when they are employed alone. Data compression reduces payload size, while the IM rule reduces the offered load to the network. For the Simple Format, significant reductions in network-averaged position error, latency and currency are achieved when the IM rule alone is used. However, little reduction is achieved when data compression alone is used since most of the gain achievable through reducing payload size has already been achieved through the more compact payload format. For the Simple Format, virtually all of the reductions achieved when the techniques are employed together are attributable to use of the IM rule.

7.4 SUMMARY AND CONCLUSIONS

This chapter has presented results of simulations conducted to examine the impact of application-layer information management techniques on the throughput of operationally important information over a congested tactical radio subnet. In the simulations, information exchange is accomplished via data replication (Section 4.2) based on an all-informed distribution model (Section 4.2.2.6). An experimental Replication Transport Layer satisfying many of the characteristics identified in Section 4.2.2.4 was employed. The results are specific to the particular network and load conditions used in the simulations, but serve to illustrate the positive impact that application-layer information management techniques that reduce payload size or offered load can have on information flow over disadvantaged tactical communication grids. In particular, the value of using a context-sensitive information management rule (Section 4.2.2 and 4.2.2.4.1) to adjust, without user intervention, the offered load to the network in response to network conditions is demonstrated.

Chapter 8 – SUMMARY AND CONCLUSIONS

This report summarizes a four-year study carried out by NATO RTO/IST-030 Research Task Group 012 on the problem of "Information Management over Disadvantaged Grids". Such disadvantaged grids (e.g., tactical ad hoc networks) are characterized by low bandwidth, variable throughput, unreliable connectivity, and energy constraints imposed by the wireless communications grid that links the nodes. The Task Group studied managed information exchange from three different perspectives or levels within a system architecture: the application level, the middleware level and the network level.

The Task Group limited its scope to land-based data exchange on the tactical battlefield (i.e., below brigade level) where all nodes are mobile and the exchange medium is combat net radio. The objective of managed information exchange is to support the commander's ability to execute command and control by providing a timely flow of accurate, relevant information. At the tactical level, periodic updates of 'blue' situational awareness information (position of friendly units) every few minutes is the most important component of information exchange. Superimposed on this regular, periodic traffic is battle management traffic sent on an as-required situation-specific basis (e.g., report of enemy contact, patrol report, call for fire, or a fragmentary order).

Currently, there are two alternative approaches to tactical information exchange: data replication and formal messaging, both of which were addressed in this report. Of these options, data replication offers the most potential for minimizing bandwidth demands (by propagating data changes only, at the database transaction level), and for maximizing interoperability by exchanging data based on an agreed formal data schema. The Task Group focused primarily on this option. It was assumed that each mobile node exchanging data possessed a database that enforced a common data structure based on an agreed data schema, and that data exchange occurred by replicating database changes (transactions) on a local node to the databases on other network nodes using the shared radio medium.

The highly mobile tactical military environment creates several challenges not endemic to either strategic or civilian environments. The most crucial challenges include: low and varying data rate on a shared medium, unreliable links, and possibly severely limited resources such as energy or computational power as well as radio-silence situations. This report concludes that asynchronous replication mechanisms are best for this type of communications environment.

It is argued that an "all-informed" data distribution scheme offers several advantages in the disadvantaged, tactical domain. It takes maximum advantage of the shared radio medium, allows an easy hand-over of responsibility with minimum synchronization requirements, and avoids a single point of failure for important data. However, in the disadvantaged tactical communications environment, it will not be possible to maintain complete consistency of database content across all nodes in the network. Under these conditions, the replication mechanism needs to coordinate its efforts with the network to ensure the consistency of more important data at the cost of inconsistency of less important data.

The authors conclude that, with an appropriate addressing scheme, it is possible to combine an all-informed distribution model for certain types of information (e.g., position of friendly units) with a selective distribution model for other types of information (e.g., a fragmentary order). This hybrid distribution model, which combines full synchronization of database content for certain types of data with selective synchronization for other types of data, may provide the best match between operational requirements and bandwidth utilization in the tactical wireless domain.

Due to the highly variable quality of the tactical communications channels and the unpredictable nature of the tactical battlefield, it is argued that dynamic adaptation to rapid changes in either the communications or battlefield environment is required to achieve optimal information exchange. This adaptation is possible only if some information about the current status of the network is available to the middleware and/or application layer in each participating node.

Concerning middleware design, the report concludes that, in addition to the traditional requirements such as scalability, reliability and support for heterogeneity, next-generation middleware must meet several new requirements to satisfy the operating challenges in the tactical domain. The most important of these are context awareness, adaptivity, and the ability to function with acceptable levels of performance in both non-disadvantaged and disadvantaged communication environments. Another desirable quality for tactical middleware is an architecture that separates, to the extent possible, middleware services directly accessible to the application from the underlying connections to the communications network that support these services (so-called 'upperware' and 'lowerware'). Such an architecture eases the problem of adapting slowly evolving applications to more rapidly evolving communications technologies. It also permits applications to use the services of a single exchange mechanism that will exploit the best available transport mechanism based on context. The authors conclude that it is feasible to design middleware having all of the above characteristics, provided that both non-disadvantaged and disadvantaged environments are kept fully in mind from the outset.

This report has identified special characteristics of ad hoc wireless networks that differentiate them fundamentally from wired or cellular networks. The primary differences result from the lack of an infrastructure and from the fact that the set of network links and their capacities are not determined *a priori*. The tactical military domain imposes additional challenges such as low data rates, variable communications conditions and susceptibility to hostile environments. The report has discussed how designing tactical ad hoc wireless networks using cross-layer techniques rather than traditional layered design principles can provide performance benefits throughout the whole system, i.e., from the radio physical layer through the application layer.

Although this report recommends the use of cross-layer techniques, it discourages the complete abandonment of layers. It is felt that appropriate use of cross-layer techniques would respect the layered structure, but would optimize the network's overall performance in specific contexts by sharing selected information across layers. This approach would also facilitate sharing of key information about network state with higher layers (middleware and/or application).

The report has also discussed the difference between 'energy-efficient' and 'energy-constrained' operation. Both of these 'energy-aware' modes of operation are crucial to tactical networks involving dismounted soldiers because of the need to limit the weight of batteries that soldiers carry. Techniques that are based on minimization of total energy expenditure (summed over all network nodes) do not necessarily perform well when the batteries at the nodes cannot be replaced. Moreover, the total energy available to a node must be shared among the functions it must support. Energy-constrained operation leads to a strong coupling among functions at several layers of the protocol stack, and consequently is a strong candidate to benefit from the use of cross-layer network protocols.

Results have been presented from simulations driven by a tactical scenario in which exchange of position updates over a single tactical radio subnet is accomplished via data replication based on an all-informed distribution model. The results illustrate the positive impact that application-layer information management techniques that reduce payload size or limit offered load can have on information flow over disadvantaged tactical communication grids. In particular, the potential value of using context-sensitive information

management rules at the application layer to adjust the offered load to the network in response to network conditions without user intervention is demonstrated.

The authors' overall conclusion is that, for optimal information exchange performance in the tactical domain, systems need to be designed from a holistic perspective. All levels of a system architecture (application/ database, middleware and network) must be designed to work cooperatively to manage the information flow. This report has attempted to identify required attributes that must be present at each level to enable this cooperative behaviour.

Chapter 9 – REFERENCES

(Task Group workshop presentations referred to below can be accessed through workshop programmes in Annexes A, B and C.)

[1] Multilateral Interoperability Programme Tactical C2IS Interoperability Requirement, MIP Operational Working Group document MTIR-OWG, draft for Edition 2.0, iteration 3, 27 September 2003, p. 17.

[2] Johnson, T., "The Harsh Reality of a Harsh Communications Environment", NATO IST-030/RTG-012 Workshop on 'Data Replication over Disadvantaged Tactical Communication Links', Québec City, Canada, 11-12 September 2002.

[3] Chamberlain, S., "Design Concepts for Resilient Database Replication in Tenuous Communication Environments", keynote presentation at NATO IST-030/RTG-012 Workshop on 'Data Replication over Disadvantaged Tactical Communication Links', Québec City, Canada, 11-12 September 2002.

[4] St-Jacques, J.-C., "Artillery Regimental Data System Advanced Development Model – Replication Issues", NATO IST-030/RTG-012 Workshop on 'Data Replication over Disadvantaged Tactical Communication Links', Québec City, Canada, 11-12 September 2002.

[5] Angel, P., "ATCCIS Replication Mechanism (ARM) Fundamental Concepts", NATO IST-030/RTG-012 Workshop on 'Data Replication over Disadvantaged Tactical Communication Links', Québec City, Canada, 11-12 September 2002.

[6] Corcus, M. and Winkowski, D., "XML Sizing and Compression Study For Military Wireless Data", XML Conference & Exposition 2002, Baltimore, USA, December 2002.

[7] Kunz, T., "Why Current Middleware Fails for Mobile Peer-to-Peer Computing", NATO IST-030/ RTG-012 Workshop on 'Role of Middleware in Systems Functioning Over Mobile Wireless Networks', Wachtberg, Germany, 26-27 August 2003.

[8] Johnson, B.C., "A Distributed Computing Environment Framework: An OSF Perspective", The Open Group, Inc., Woburn, MA, http://www.opengroup.org/dce/info/papers/dev-dce-tp4-1.ps, June 1991.

[9] Real-Time CORBA Specification, Version 1.1, formal/02-08-02, http://www.omg.org/docs/formal/02-08-02.pdf, August 2002.

[10] minimumCORBA Joint Revised Submission to OMG, OMG TC Document orbos/98-08-04, http://www.omg.org/docs/orbos/98-08-04.pdf, 17 August 1998.

[11] Eikerlink, H.-J., "Context-Awareness in Middleware for Mobile Networks", NATO IST-030/RTG-012 Workshop on 'Role of Middleware in Systems Functioning Over Mobile Wireless Networks', Wachtberg, Germany, 26-27 August 2003.

[12] Goldsmith, A.J. and Wicker, S.B., "Design Challenges for Energy-Constrained Ad Hoc Wireless Networks," IEEE Wireless Communications Magazine, 9, pp. 2-22, August 2002.

[13] Tanenbaum, A.S., Computer Networks, Fourth Edition, Upper Saddle River, NJ: Prentice Hall PTR 2003.

[14] Pursley, M.B., "Integrated Cross-Layer Protocols for Adaptive Transmission and Routing of Multimedia Traffic in Tactical Spread-Spectrum Networks", keynote presentation at NATO IST-030/RTG-012 Workshop on 'Cross-Layer Issues in the Design of Tactical Mobile Ad Hoc Wireless Networks', Naval Research Laboratory, Washington, DC, USA, 2-3 June 2004.

[15] Goldsmith, A.J., "Can Cross-Layer Techniques Enhance the Performance of Tactical Military Networks?," panel discussion presentation at NATO IST-030/RTG-012 Workshop on 'Cross-Layer Issues in the Design of Tactical Mobile Ad Hoc Wireless Networks', Naval Research Laboratory, Washington, DC, USA, 2-3 June 2004.

[16] Kawadia, V. and Kumar, P.R., "A Cautionary Perspective on Cross-Layer Design", IEEE Wireless Communications Magazine, Vol. 12, pp. 3-11, January 2005.

[17] Gibb, A.W., St-Jacques, J.-C., Plante, P., Caron, J.-D., Stemate, L. and Nadeau, F., "Information Management Studies Conducted for High Capacity Tactical Communications Network TD Project: Final Report", DRDC Valcartier Technical Report, 2007 (to be published). UNCLASSIFIED.

Annex A – DATA REPLICATION WORKSHOP TECHNICAL PROGRAMME

NATO IST-030/RTG-012 INFORMAL WORKSHOP

DATA REPLICATION OVER DISADVANTAGED TACTICAL COMMUNICATION LINKS

Defence R&D Canada – Valcartier
Québec City, Canada
11-12 September 2002

(Click on the links below to view the presentations)

Wednesday, 11 September

0910 *Overview of DRDC Valcartier R&D Programme in Command and Control*
– Allan Gibb, Defence R&D Canada – Valcartier, Canada

0940 Keynote Presentation – *Resilient Database Replication in Tenuous Communication Environments*
– Sam Chamberlain, U.S. Army Research Laboratory

1100 *Data Replication in a Combat Net Radio Environment – The Harsh Reality of a Harsh Communications Environment*
– Tim Johnson, IP Unwired

1300 *Overview of ATCCIS Replication Mechanism*
– Peter Angel, Advanced Systems Management Group

1330 *Dynamic Contracting – Saving Bandwidth and Controlling the Data Flow in Danish Army Command and Control Information System*
– Erling Rasmussen, Operational User Group for Danish Army CCIS

1450 *Analysis of Limiting Information Flow and Information Storage*
– Freek Driesenaar, TNO Physics and Electronics Laboratory

1520 *Artillery Regimental Data System Advanced Development Model – Replication Issues*
– Jean-Claude St-Jacques, DRDC Valcartier

1550 Wrap-up and Close

Thursday, 12 September

0920 *Data Replication over Disadvantaged Links – A Navy Perspective*
– John Bycroft, Canadian Navy

1000 *Polish Tactical Data Exchange System*
– Jaroslaw Michalak, Military University of Technology

1050 *Measuring Performance of Replication Mechanisms in Tactical Mobile Environments*
– Allan Gibb, Defence R&D Canada – Valcartier, Canada

1120 *Replication in Mobile Environments*
– Heinz Fassbender, FGAN/FKIE

1300 Group Discussions

The Plenary session broke up into two syndicates for discussion of the following questions:

1) What key concepts and design principles should drive the design of data replication/transport mechanisms to function optimally over disadvantaged tactical communication links?
2) What key operational requirements should drive the design of data replication/transport mechanisms to function optimally over disadvantaged tactical communication links?

1530 Report of Discussion Group 1 and Discussion by Plenary

1555 Report of Discussion Group 2 and Discussion by Plenary

1620 Wrap-Up

1630 Workshop Close

Annex B – MIDDLEWARE WORKSHOP TECHNICAL PROGRAMME

NATO IST-030/RTG-012 INFORMAL WORKSHOP

ROLE OF MIDDLEWARE IN SYSTEMS FUNCTIONING OVER MOBILE WIRELESS NETWORKS

FGAN/FKIE
Wachtberg, Germany
26-27 August 2003

(Click on the links below to view the presentations)

Tuesday, 26 August

0905 Overview of FGAN/FKIE
– Habil J. Grosche, Director FGAN/FKIE

0930 *Challenges for Middleware Imposed by the Tactical Army Communications Environment*
– Allan Gibb, Defence R&D Canada – Valcartier, Canada

1010 *Replication Middleware for a Tactical Mobile Wireless Environment*
– Allan Gibb, Defence R&D Canada – Valcartier, Canada

1100 *An Operational View of the Problems in Data Exchange in the Army Mobile Environment*
– Erling Rasmussen, Operational User Group for Danish Army Command, Control and Information System (DACCIS)

1300 *Why Current Middleware Fails for Mobile Peer-to-Peer Computing*
– Abdulbaset Gaddah and Thomas Kunz, Department of Systems and Computer Engineering, University of Ottawa, Canada

1330 *Architectures for Mobile Wireless Publish/Subscribe Networks*
–David S. Rosenblum, Chief Technology Officer, PreCache Inc.

1400 *Context-Awareness in Middleware for Mobile Networks*
– Heinz-Josef Eikerling, Siemens SBS C-LAB, Paderborn, Germany

1450 *Challenges for a Distributed Collaborative Environment Functioning over Mobile Wireless Networks*
– Jean-Claude St-Jacques, Defence R&D Canada – Valcartier

1520 *Secure Middleware for Robust and Efficient Interoperability over Disadvantaged Grids*
– Ramesh Bharadwaj, Centre for High Assurance Computer Systems, Naval Research Laboratory, Washington, DC

1600 Wrap-up

Wednesday, 27 August

0920 *Flexible CORBA Components for Mission-Critical Distributed Applications*
– U. Lang, University of Cambridge Computer Lab, T. Ritter, Fraunhofer FOKUS, R. Schreiner, ObjectSecurity Ltd.

0950 *Network Simulation Tools for Prototyping Scalable P2P Applications*
– I.J. Taylor, Dept of Computer Science, Cardiff University, Brian Adamson, Naval Research Laboratory, Washington DC

1040 Group Discussions

The Plenary session broke up into two groups for discussion of the following specific questions:

1) Discussion Group 1 – Can you design middleware to be equally effective in the wired and wireless domains? If not, in what ways must they be different, and why?
2) Discussion Group 2 – What types of middleware offer the most benefit in the tactical wireless domain, and why?

1530 Report of Discussion Group 1 and Discussion by Plenary (report unavailable)

1555 Report of Discussion Group 2 and Discussion by Plenary

1620 Wrap-Up

1630 Workshop Close

Annex C – CROSS-LAYER WORKSHOP TECHNICAL PROGRAMME

NATO IST-030/RTG-012 INFORMAL WORKSHOP

CROSS-LAYER ISSUES IN THE DESIGN OF TACTICAL MOBILE AD HOC WIRELESS NETWORKS

Integration of Communication and Networking Functions to Support Optimal Information Management

Naval Research Laboratory
Washington, DC, USA
2-3 June 2004

(Click on the links below to view the presentations)

Wednesday, 2 June

0900 Welcoming Remarks
– Jeffrey E. Wieselthier, Workshop Chair, and John McLean, Superintendent, Information Technology Division, Naval Research Laboratory

0930 *Information Management in a Tactical Mobile Wireless Communications Environment*
– Allan Gibb, Defence R&D Canada – Valcartier, Canada

1045 Keynote Presentation – *Integrated Cross-Layer Protocols for Adaptive Transmission and Routing of Multimedia Traffic in Tactical Spread-Spectrum Networks*
– Michael B. Pursley, Holcomb Professor of Electrical Engineering, Clemson University

1300 Session 1

Cross-Layer Design of Wireless Networks with Resource-Constrained Nodes
– Andrea Goldsmith, Stanford University

Delay-Energy Analysis of Wireless Networks
– Shih Yu Chang, Achilleas Anastasopoulos, and Wayne Stark, University of Michigan

Scheduling on a Channel with Time-Varying Capacity
– Ashay Dhamdhere and Ramesh R. Rao University of California, San Diego

Topology Management from Bottom to Top
– Martha Steenstrup, Stow Research LLC and Clemson University

1445 Session 2

Medium-Access Control Protocols for Heterogeneous Mobile Ad Hoc Networks with Directional Antennas
– Daniel L. Noneaker and Harlan B. Russell, Clemson University

Cross-Layer Simulation and Aggregation Techniques for Wireless Ad Hoc Networks
– Vincent Gauthier, Monique Becker, André Luc Beylot, and Riadh Dhaou, Institut National des Télécommunications, Evry, France

Cross-Layer Optimization and Adaptation in Wireless Mobile Ad Hoc Networks
– Ashutosh Dutta, Raquel Morera, Tony McAuley, Nim Cheung, and Ken Young, Telcordia Technologies, Inc.

Interlayer Routing Issues for Wireless Networks
– Thomas R. Henderson, Marcelo Albuquerque, Phillip A. Spagnolo, and Jae H. Kim, Boeing

1615 Poster Session and Reception

Thursday, 3 June

0830 Session 3

Cross-Layer Design Issues for MANET Autoconfiguration
– Joseph P. Macker, Naval Research Laboratory

Attacks and Defenses Utilizing Cross-Layer Interactions in MANET
– John S. Baras and Svetlana Radosavac, University of Maryland, College Park

A Cross-Layer Diversity Technique for Multi-Carrier OFDM Multimedia Networks
– Yee Sin Chan, Pamela C. Cosman, Larry Milstein, University of California, San Diego

Balancing Transport and Physical Layers in Wireless Ad Hoc Networks: Jointly Optimal TCP Congestion Control and Power Control
– Mung Chiang, Princeton University

1015 Session 4

Cross-Layer Design for Data Accessibility in Mobile Ad Hoc Networks
– Klara Nahrstedt, University of Illinois at Urbana-Champaign

Adaptive Middleware for Challenged Networks
– Tom Ritter (FhG FOKUS, Germany) and Rudolf Schreiner (ObjectSecurity Ltd., UK)

Energy, Routing and Decentralized Detection in a Sensor Network
– Steven A. Borbash, Department of Defense/NSA; and Anthony Ephremides, University of Maryland, College Park

Cross Layer Design for Large Scale Sensor Networks
– Ananthram Swami, Army Research Laboratory

A Line in the Sand: A Wireless Sensor Network for Target Detection, Classification, and Tracking
– Anish Arora et al., The Ohio State University

1315 Panel Discussion – *Can Cross-Layer Techniques Enhance the Performance of Tactical Military Networks?*

Moderator: Jeffrey E. Wieselthier, Naval Research Laboratory

Panellists:

Andrea Goldsmith, Stanford University

Larry Milstein, University of California, San Diego

Klara Nahrstedt, University of Illinois at Urbana-Champaign

Raymond Pickholtz, George Washington University

Michael B. Pursley, Clemson University

Martha Steenstrup, Stow Research LLC and Clemson University

1600 Final Remarks and Discussion

1630 Workshop Close

POSTER SESSION

Efficient Message Authentication for Spread Spectrum Wireless Communications
– Charles Boncelet, University of Delaware; and David Carman, McAfee Research

Future Combat System Scalable Mobile Network Demonstration -- Tactical Mobile Ad-Hoc Networking Performance and Validation Results
– Wayne W. Brown, Vincent Marano IV, William MacCorkell, The Boeing Company; and Timothy Krout, Cengen, Inc.

Joint Iterative Decoding and Authentication
– David Carman, McAfee Research, Network Associates; Michael Jordan, The Johns Hopkins University Applied Physics Laboratory; and Charles Boncelet, University of Delaware

On the Use of Path Diversity with Bursty Channels
– Roch Guerin, University of Pennsylvania

Multipath Routing – A Cross-Layer Design Tool for QoS Provisioning in MANETs
– Zygmunt J. Haas, Cornell University

Support Multimedia SIP Applications over MANET Using Cross Layer Design
– Li Li and Louise Lamont, Communications Research Centre of Canada

Cross-Layering Approach for GPS-based Routing and Network Topology Construction
– Yibei Ling, Wai Chen, and Russell Hsing, Telcordia Technologies

Cooperative Diversity in Tactical Networks

– John M. Shea, Tan F. Wong, Yuguang Fang, Arun Avudainayagam, Wing Hin Wong, and Xin Li, University of Florida

Cross-Layer Approach to Low Energy Wireless Ad Hoc Networks

– Geethapriya Thamilarasu, State University of New York at Buffalo; Sumita Mishra, CompSys Technologies, Inc.; and Ramalingam Sridhar, State University of New York at Buffalo

Scaling Cooperative Diversity to Large Networks

– Matthew C. Valenti, West Virginia University

Energy-Aware Broadcasting and Multicasting in Wireless Ad Hoc Networks: A Cross-Layering Approach

– Jeffrey E. Wieselthier and Gam D. Nguyen, Naval Research Laboratory; and Anthony Ephremides, University of Maryland

Annex D – ATCCIS REPLICATION MECHANISM

ATCCIS (Army Tactical Command and Control Information System) was an international programme consisting of NATO nations (although not formally a NATO effort) aimed at identifying the minimum set of specifications to be included within C2ISs to allow the automatic transfer of selected command and control (C2) data. Their objective was to develop a specification for a hardware/software/vendor-independent interoperability solution. The ATCCIS programme ran from 1982 to 2002.

The Multinational Interoperability Programme is an international programme consisting of NATO nations (also not a NATO effort) whose focus is the fielding of an interoperability solution for multinational C2ISs. In 2002, ATCCIS merged with MIP. MIP adopted the products of the ATCCIS work as the basis for direct database-to-database exchange. However, MIP also maintains a structured message exchange mechanism.

The ATCCIS concept of interoperability is based upon the automatic transfer of standardized data elements based upon a common data interchange model called the Land C2 Information Exchange Data Model.

The ATCCIS programme also developed the specification for a mechanism that will permit interoperability of automated C2ISs through partial replication of database content. The ATCCIS Replication Mechanism (ARM) is selective in: (a) data to be exchanged; (b) recipients of the data; and (c) transfer facility to be used.

Under the ATCCIS concept, nations use the common data model to preserve the meaning and relationships of the information exchanged between C2ISs across national boundaries. The ARM manages the exchange of information between databases of C2ISs across national boundaries based on the common data model.

The major component areas of an ATCCIS compliant system are shown in Figure D-1.

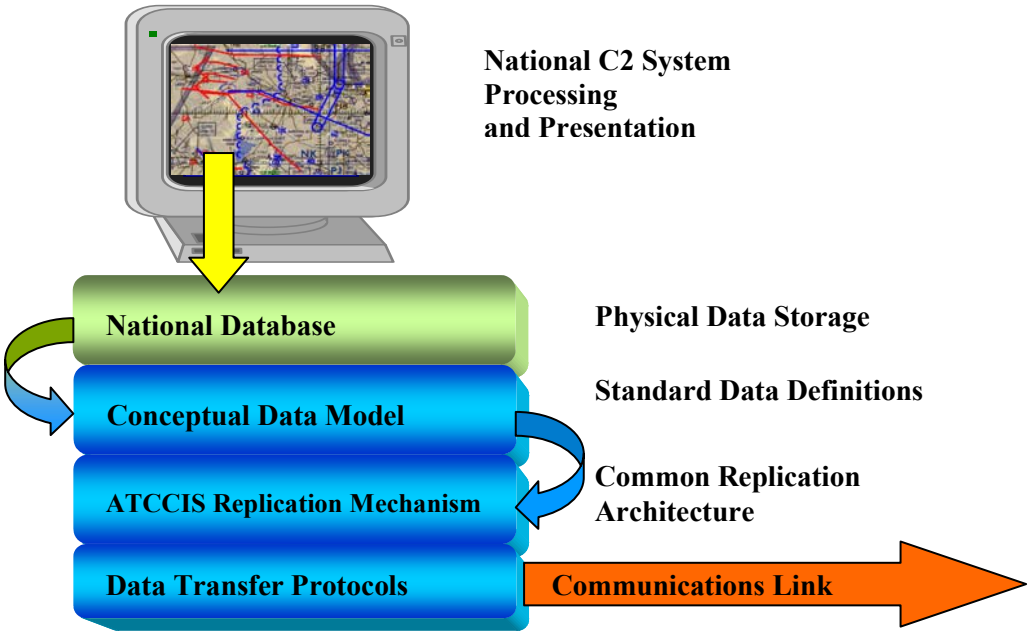


Figure D-1: ATCCIS Concept of Operations.

The top two levels, that is the national application and national database, represent that portion of the implementation that is a national responsibility. Nations develop independent applications that support national operating procedures and language. They also maintain national information within their own databases.

When an update to C2 information occurs in the National Database, and the update needs to be shared, the data to be exchanged are referred to a conceptual data model, the Command and Control Information Exchange Data Model (C2IEDM)[1].

If the data are identified as part of replication contracts pre-established with other systems for Command and Control purposes, they are packaged as part of a replication Protocol Data Unit (PDU) and the PDU is formatted within the selected commercial transfer protocol (e.g., TCP) for transmission. The ATCCIS Replication Mechanism provides this functionality. The ARM can be considered to consist of three layers as shown in Figure D-2. The principal responsibilities of each layer are shown in the figure.

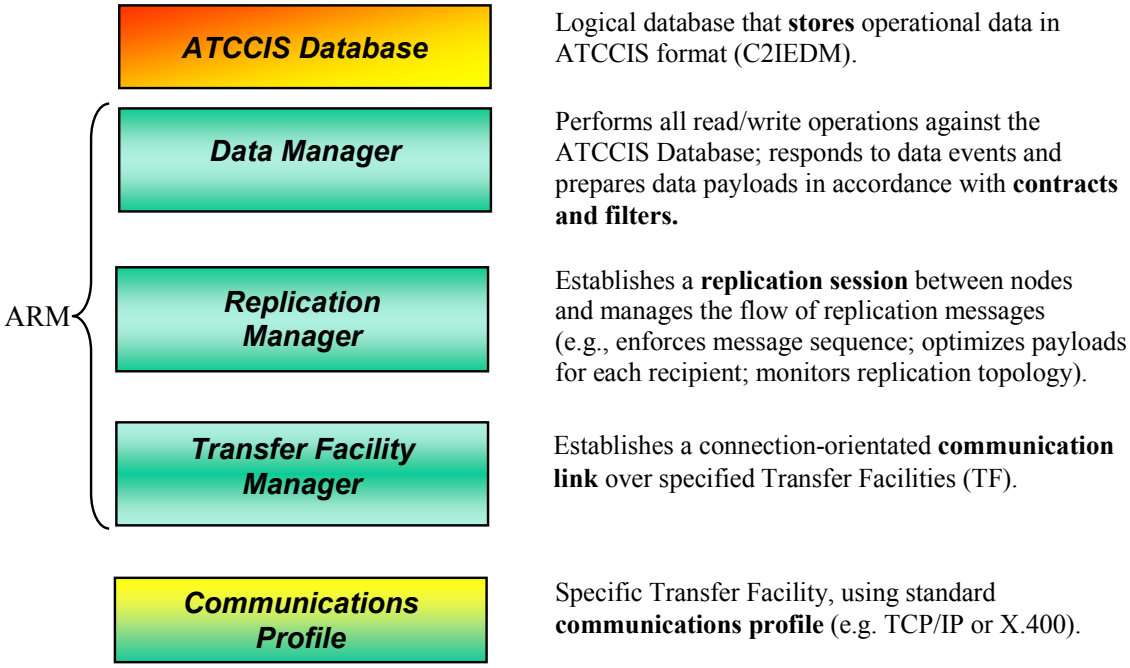


Figure D-2: ARM Layers.

The ARM employs the concepts of contracts and filters.

A *replication contract* is the means for controlling (selective) replication of database changes. A contract is established between a pair of replication nodes, designated as Data Provider (DP) and Data Receiver (DR). In the contract, the DP and DR agree that the DP will provide the DR with all data that satisfies the conditions of the contract. A contract specifies a *filter* and parameter values used to set filter conditions, as well as a DP and a DR. A filter is a set of criteria applied to the instances of a database in order to reduce the total set of data selected to a subset. Examples of filter types include geographical area, time, and order of battle (organizational). The contracts enforce a 'push' model for information exchange in which the only data

[1] The Land Command and Control Information Exchange Data Model (LC2IEDM) developed under ATCCIS has been extended under MIP to serve Joint Force requirements. The extended data model is referred to as 'C2IEDM', without the qualifier 'Land'.

pushed to recipient nodes are those negotiated with the recipient node under the pre-agreed contract. To modify the set of data pushed to a particular data recipient by a data provider, a filter must be applied or the contract must be modified.

The ARM supports three types of exchanges. A *bulk* update is the total set of data elements required to satisfy all valid contracts between a DP and a DR at a given point in time. A bulk update is used for database synchronization when a connection between nodes is established for the first time (e.g., initializing database content prior to deployment of C2ISs). A *partial bulk* update is used for database synchronization when one or more existing contracts between a DP and a DR are modified or a new contract activated. An *incremental* update is a copy of a set of one or more database transactions that have occurred since the most recent bulk or partial bulk update.

The ARM implements a selective data distribution model. Advantages and disadvantages in the tactical wireless domain of this distribution model versus an 'all-informed' distribution model are discussed in Section 4.2.2.6 of the main report.

The ARM enforces a single data ownership model (see Section 4.2.2.2 of main report). A data element is associated with one, and only one, data owner throughout its lifetime.

The C2IEDM is a relational data model. A relation is a mathematical term for a table. In a relational model, data are perceived as being organized in tables (and only tables). Database operations performed on these tables always result in new tables. The rows in a table must be unique within the table. To ensure uniqueness, one or more columns in the table are designated as a primary key. A primary key is a set of columns selected so that the set of values associated with those columns uniquely specify a table row (i.e., do not repeat within the table). Relationships between tables are established by including the primary key from one table in another table (the primary key is said to *migrate* to the second table). In this case, the primary key from the first table becomes a *foreign key* in the second table. A foreign key is defined as a set of columns that is a primary key in another table. The foreign key is said to *refer back* to the table for which it is the primary key. The table with the foreign key is the *referencing* table, and the table with the primary key is the *referenced* table. The structure of primary and foreign keys is the means by which relationships between different tables are established and maintained. If this structure is corrupted, the traceability of relationships between entries in different tables is compromised. Preservation of *referential integrity* refers to the act of ensuring that no invalid foreign key values exist in the database (i.e., that every foreign key value refers to an existing primary key value in another table). When data are replicated between databases, preservation of referential integrity is an important consideration.

The C2IEDM model design encompasses two categories of objects: those that can be identified individually (OBJECT-ITEMs) and those that represent grouped or class properties, e.g., a tank, a ship (OBJECT-TYPES). The two categories are used in parallel as basic structural elements of the model. Every instance of OBJECT-ITEM must be associated with an OBJECT-TYPE at its time of creation (for example, a particular tank belonging to an organization might be identified as type 'M1A1 Abrams'). Since the values of attributes of OBJECT-TYPES tend to be relatively static or persistent, OBJECT-TYPE information in the model is regarded as referential information that is *inherited* by each instance of OBJECT-ITEM associated with it (the above-mentioned instance of an M1A1 Abrams tank would inherit all the characteristics such as calibre of main gun, track width and load class associated with that type of tank). This referential information about OBJECT-TYPES is stored in the database of each replication node and is not replicated across the network, although references to OBJECT-TYPES can be replicated. This parallel structure of the C2IEDM model that obviates the need to share TYPE information across the network permits important savings in terms of bandwidth.

Annex E – TERMS OF REFERENCE

Task Group on
Information Management over Disadvantaged Grids
IST-030/RTG-012
15 December, 2000

I. ORIGIN

A) Background

The Research and Technology Organization (RTO) Information Systems Technology (IST) Panel recognized the challenge inherent in distributing timely and relevant tactical information as digital data using a mobile wireless communication system characterized by low and variable throughput and unreliable connectivity. In order to address that problem, the Panel authorized in October 1999 the formation of an Exploratory Team on Information Management over Disadvantaged Grids. . The Exploratory Team met at DGA HQ in Paris in May 2000 and concluded that the problem of Information Management over Disadvantaged Grids should be addressed through formation of a Task Group under the IST Panel.

B) Military Benefits

Mobile communication is an important military requirement. Voice communications still occupy a pre-eminent place in Army operations. Present-generation digital data communications at the tactical level (below Brigade) are accomplished using radio systems designed primarily with voice in mind. Data throughput tends to be very limited (less than one kbit/second is not uncommon) and highly variable. Digital C2 systems offer the promise of increased battlefield awareness. To deliver on this promise, the communication backbone must be capable of distributing relevant sets of digital data among participating C2IS nodes accurately and with a timeliness that permits friendly commanders to act within the decision cycle of the enemy commanders. Satisfying data distribution requirements of completeness, accuracy and timeliness when the communication system is characterized by low and variable throughput and highly unreliable connectivity represents a considerable challenge. Realistically, the limitations of the mobile wireless communications network will make it impossible to satisfy fully all of these requirements all of the time. Dynamic trade-offs between these factors will be required. A key factor in managing these tradeoffs is a set of adaptive protocols within each C2IS node which exploit current information about the constantly-evolving situation picture contained in the node's database, and information about the current state of the communications network, to optimize the timeliness and relevance of information passed between nodes. Commercial data replication products do not provide protocols with the sophistication required for the demanding wireless military environment. In general, the products assume the presence of reliable high bandwidth links between databases and/or an environment in which as much time as necessary can be taken to synchronize database content. Neither of these assumptions are valid on the tactical battlefield.

II. OBJECTIVES

1) Area of research and scope of activity – investigation of adaptive information management schemes, implemented in the nodes of tactical command and control systems, to mitigate the effects of low bandwidth, variable throughput, unreliable connectivity and energy-constrained nodes imposed by the mobile wireless communications grid that links the command and control nodes.

2) Specific goals:

   a) Identify the characteristics of mobile wireless communications grids which pose a challenge to the timely and accurate distribution of tactical information over the grids;

   b) Investigate how the application layer can acquire and exploit information about the state of the communications grid;

   c) Investigate and identify information management protocols specific to the application layer which can respond to changing network and battlefield conditions to optimize the timely flow of relevant information over such grids;

   d) Investigate techniques for implementing the protocols in the application layer, such as the use of database triggers and exploitation of COTS or MOTS (e.g., NATO ATCCIS) data replication mechanisms;

   e) Identify measures of effectiveness (MoE) that can be used to evaluate the operational impact of these techniques; and

   f) Investigate how advances in mobile wireless communications and database technology may influence the problem.

3) Expected deliverables:

   a) A prescription for adaptive information management schemes, and methods for implementing the schemes, in tactical command and control nodes, to counteract the communication grid characteristics of low and variable throughput, unreliable connectivity, and energy-constrained nodes.

   b) An analysis of the potential gain, in throughput of relevant information, to be achieved by use of each technique, or by combinations of techniques.

   c) Reports, technical reports, conference papers, publications documenting the analysis of information management techniques and methods for their implementation in tactical command and control nodes.

4) Overall duration of Task Group should be not more than three years.

III. RESOURCES

A) Membership

Representatives from government (civilian and military) and industry with expertise on the topics of data replication in low bandwidth military environments, tactical communication systems, or mobile wireless communications. Knowledge and expertise on the following topics is also pertinent: communication protocols, Army common data models (e.g. ATCCIS Generic Hub), tactical messaging (Army organization, procedures, communication patterns, and message types), data compression schemes, and measures of performance or measures of effectiveness for the information distribution component of military command and control systems.

Canada, United States, Germany, and Poland have agreed to participate in the Task Group. The Team Leader and Lead Nation will be chosen at the first meeting of the Task Group.

B) National and/or NATO Resources Needed

Participating nations agree to fund travel for their national representatives to attend two meetings of the Task Group per year over the three year lifetime of the Task Group. It is expected that national representatives will be able to devote sufficient time between meetings to complete successfully the mutually-agreed Programme of Work.

Nations may be required to furnish information concerning the performance characteristics and/or architecture of their tactical communication systems.

C) RTA Resources Needed

Nil.

IV. SECURITY LEVEL

Most work will be unclassified. However, if details of the performance and/or architecture of national tactical communication systems are divulged, the classification of this part of the activity could be up to NATO SECRET.

V. PARTICIPATION BY PARTNER NATIONS

Partner nations will not be invited to participate in the Technical Team.

VI. LIAISON (WITH OTHER NATO BODIES)

The TGonIMDG should liaise with the following NATO bodies:

- IST Panel:
  - Information Management Challenges in Achieving Coalition Interoperability (IST-022/RSY-007);
  - Military Communications (IST-023/RSY-008); and
  - Awareness of Emerging Wireless Technologies (IST-ET-020).
- NATO Tri-Service Group on Communications and Electronics, Project Group 6, developers of STANAG for 'Tactical Communications Systems for the Land Combat Zone – Post 2000' (TACOMS Post-2000).
- NATO Permanent Working Group on Army Tactical Command and Control Information System (ATCCIS).

REPORT DOCUMENTATION PAGE

1. Recipient's Reference	2. Originator's References	3. Further Reference	4. Security Classification of Document
	RTO-TR-IST-030 AC/323(IST-030)TP/33	ISBN 978-92-837-0082-1	UNCLASSIFIED/ UNLIMITED

5. Originator

Research and Technology Organisation
North Atlantic Treaty Organisation
BP 25, F-92201 Neuilly-sur-Seine Cedex, France

6. Title

Information Management over Disadvantaged Grids

7. Presented at/Sponsored by

Final Report of the RTO Information Systems Technology Panel
Task Group IST-030/RTG-012.

8. Author(s)/Editor(s)	9. Date
Multiple	December 2007
10. Author's/Editor's Address	**11. Pages**
Multiple	94

12. Distribution Statement

There are no restrictions on the distribution of this document.
Information about the availability of this and other RTO
unclassified publications is given on the back cover.

13. Keywords/Descriptors

Ad hoc network	Disadvantaged network	Mobile network
Adaptive middleware	Dynamic adaptation	Network analysis (management)
Combat net radio	Energy-constrained network	Protocols
Command and control	Information management	Radio communication
Computer networks	Information systems	Radio links
Cross-layering	Integrated systems	Requirements
Data links	Low bandwidth	Tactical communications
Data replication	Managed information exchange	Tactical radio
Data transmission	Military communication	Wireless network
Disadvantaged grid		

14. Abstract

This report summarizes a four-year study carried out by NATO RTG-012/IST-030 Research Task Group on the problem of "Information Management over Disadvantaged Grids". Such disadvantaged grids (e.g., tactical ad hoc military radio networks) are characterized by low bandwidth, variable throughput, unreliable connectivity, and energy constraints imposed by the wireless communications grid that links the nodes. The scope of this study was limited to land-based digital data exchange below brigade level where all nodes are mobile and the exchange medium is combat net radio. Managed information exchange in this communications environment was analyzed from three different perspectives within a system architecture: the application level, the middleware level and the network level. Due to the highly variable quality of the tactical communications channels and the unpredictable nature of the tactical battlefield, it was concluded that dynamic adaptation to rapid changes in either the communications or battlefield environment, without user intervention, was key to achieving optimum information exchange. This report identifies functional and performance objectives for the application, middleware and network levels that enable the levels to cooperate to detect and adapt to those changing environments in a way that will enhance delivery of data of highest operational importance.

NORTH ATLANTIC TREATY ORGANISATION

RESEARCH AND TECHNOLOGY ORGANISATION

BP 25
F-92201 NEUILLY-SUR-SEINE CEDEX • FRANCE
Télécopie 0(1)55.61.22.99 • E-mail mailbox@rta.nato.int

DIFFUSION DES PUBLICATIONS
RTO NON CLASSIFIEES

Les publications de l'AGARD et de la RTO peuvent parfois être obtenues auprès des centres nationaux de distribution indiqués ci-dessous. Si vous souhaitez recevoir toutes les publications de la RTO, ou simplement celles qui concernent certains Panels, vous pouvez demander d'être inclus soit à titre personnel, soit au nom de votre organisation, sur la liste d'envoi.

Les publications de la RTO et de l'AGARD sont également en vente auprès des agences de vente indiquées ci-dessous.

Les demandes de documents RTO ou AGARD doivent comporter la dénomination « RTO » ou « AGARD » selon le cas, suivi du numéro de série. Des informations analogues, telles que le titre est la date de publication sont souhaitables.

Si vous souhaitez recevoir une notification électronique de la disponibilité des rapports de la RTO au fur et à mesure de leur publication, vous pouvez consulter notre site Web (www.rto.nato.int) et vous abonner à ce service.

CENTRES DE DIFFUSION NATIONAUX

ALLEMAGNE
Streitkräfteamt / Abteilung III
Fachinformationszentrum der Bundeswehr (FIZBw)
Gorch-Fock-Straße 7, D-53229 Bonn

BELGIQUE
Royal High Institute for Defence – KHID/IRSD/RHID
Management of Scientific & Technological Research for Defence, National RTO Coordinator
Royal Military Academy – Campus Renaissance
Renaissancelaan 30, 1000 Bruxelles

CANADA
DSIGRD2 – Bibliothécaire des ressources du savoir
R et D pour la défense Canada
Ministère de la Défense nationale
305, rue Rideau, 9e étage
Ottawa, Ontario K1A 0K2

DANEMARK
Danish Acquisition and Logistics Organization (DALO)
Lautrupbjerg 1-5, 2750 Ballerup

ESPAGNE
SDG TECEN / DGAM
C/ Arturo Soria 289
Madrid 28033

ETATS-UNIS
NASA Center for AeroSpace Information (CASI)
7115 Standard Drive
Hanover, MD 21076-1320

FRANCE
O.N.E.R.A. (ISP)
29, Avenue de la Division Leclerc
BP 72, 92322 Châtillon Cedex

GRECE (Correspondant)
Defence Industry & Research General Directorate, Research Directorate
Fakinos Base Camp, S.T.G. 1020
Holargos, Athens

HONGRIE
Department for Scientific Analysis
Institute of Military Technology
Ministry of Defence
P O Box 26
H-1525 Budapest

ISLANDE
Director of Aviation
c/o Flugrad
Reykjavik

ITALIE
General Secretariat of Defence and National Armaments Directorate
5th Department – Technological Research
Via XX Settembre 123
00187 Roma

LUXEMBOURG
Voir Belgique

NORVEGE
Norwegian Defence Research Establishment
Attn: Biblioteket
P.O. Box 25
NO-2007 Kjeller

PAYS-BAS
Royal Netherlands Military Academy Library
P.O. Box 90.002
4800 PA Breda

POLOGNE
Centralny Ośrodek Naukowej Informacji Wojskowej
Al. Jerozolimskie 97
00-909 Warszawa

PORTUGAL
Estado Maior da Força Aérea
SDFA – Centro de Documentação
Alfragide
P-2720 Amadora

REPUBLIQUE TCHEQUE
LOM PRAHA s. p.
o. z. VTÚLaPVO
Mladoboleslavská 944
PO Box 18
197 21 Praha 9

ROUMANIE
Romanian National Distribution Centre
Armaments Department
9-11, Drumul Taberei Street
Sector 6
061353, Bucharest

ROYAUME-UNI
Dstl Knowledge Services
Information Centre
Building 247
Dstl Porton Down
Salisbury
Wiltshire SP4 0JQ

SLOVENIE
Ministry of Defence
Central Registry for EU and NATO
Vojkova 55
1000 Ljubljana

TURQUIE
Milli Savunma Bakanlığı (MSB)
ARGE ve Teknoloji Dairesi Başkanlığı
06650 Bakanliklar
Ankara

AGENCES DE VENTE

NASA Center for AeroSpace Information (CASI)
7115 Standard Drive
Hanover, MD 21076-1320
ETATS-UNIS

The British Library Document Supply Centre
Boston Spa, Wetherby
West Yorkshire LS23 7BQ
ROYAUME-UNI

Canada Institute for Scientific and Technical Information (CISTI)
National Research Council Acquisitions
Montreal Road, Building M-55
Ottawa K1A 0S2, CANADA

Les demandes de documents RTO ou AGARD doivent comporter la dénomination « RTO » ou « AGARD » selon le cas, suivie du numéro de série (par exemple AGARD-AG-315). Des informations analogues, telles que le titre et la date de publication sont souhaitables. Des références bibliographiques complètes ainsi que des résumés des publications RTO et AGARD figurent dans les journaux suivants :

Scientific and Technical Aerospace Reports (STAR)
STAR peut être consulté en ligne au localisateur de ressources uniformes (URL) suivant: http://www.sti.nasa.gov/Pubs/star/Star.html
STAR est édité par CASI dans le cadre du programme NASA d'information scientifique et technique (STI)
STI Program Office, MS 157A
NASA Langley Research Center
Hampton, Virginia 23681-0001
ETATS-UNIS

Government Reports Announcements & Index (GRA&I)
publié par le National Technical Information Service
Springfield
Virginia 2216
ETATS-UNIS
(accessible également en mode interactif dans la base de données bibliographiques en ligne du NTIS, et sur CD-ROM)

NORTH ATLANTIC TREATY ORGANISATION

RESEARCH AND TECHNOLOGY ORGANISATION

BP 25

F-92201 NEUILLY-SUR-SEINE CEDEX • FRANCE

Télécopie 0(1)55.61.22.99 • E-mail mailbox@rta.nato.int

DISTRIBUTION OF UNCLASSIFIED RTO PUBLICATIONS

AGARD & RTO publications are sometimes available from the National Distribution Centres listed below. If you wish to receive all RTO reports, or just those relating to one or more specific RTO Panels, they may be willing to include you (or your Organisation) in their distribution.

RTO and AGARD reports may also be purchased from the Sales Agencies listed below.

Requests for RTO or AGARD documents should include the word 'RTO' or 'AGARD', as appropriate, followed by the serial number. Collateral information such as title and publication date is desirable.

If you wish to receive electronic notification of RTO reports as they are published, please visit our website (www.rto.nato.int) from where you can register for this service.

NATIONAL DISTRIBUTION CENTRES

BELGIUM
Royal High Institute for Defence – KHID/IRSD/RHID
Management of Scientific & Technological Research
for Defence, National RTO Coordinator
Royal Military Academy – Campus Renaissance
Renaissancelaan 30
1000 Brussels

CANADA
DRDKIM2 – Knowledge Resources Librarian
Defence R&D Canada
Department of National Defence
305 Rideau Street, 9th Floor
Ottawa, Ontario K1A 0K2

CZECH REPUBLIC
LOM PRAHA s. p.
o. z. VTÚLaPVO
Mladoboleslavská 944
PO Box 18
197 21 Praha 9

DENMARK
Danish Acquisition and Logistics Organization (DALO)
Lautrupbjerg 1-5
2750 Ballerup

FRANCE
O.N.E.R.A. (ISP)
29, Avenue de la Division Leclerc
BP 72, 92322 Châtillon Cedex

GERMANY
Streitkräfteamt / Abteilung III
Fachinformationszentrum der Bundeswehr (FIZBw)
Gorch-Fock-Straße 7
D-53229 Bonn

GREECE (Point of Contact)
Defence Industry & Research General Directorate
Research Directorate, Fakinos Base Camp
S.T.G. 1020
Holargos, Athens

HUNGARY
Department for Scientific Analysis
Institute of Military Technology
Ministry of Defence
P O Box 26
H-1525 Budapest

ICELAND
Director of Aviation
c/o Flugrad, Reykjavik

ITALY
General Secretariat of Defence and
National Armaments Directorate
5th Department – Technological
Research
Via XX Settembre 123
00187 Roma

LUXEMBOURG
See Belgium

NETHERLANDS
Royal Netherlands Military
Academy Library
P.O. Box 90.002
4800 PA Breda

NORWAY
Norwegian Defence Research
Establishment
Attn: Biblioteket
P.O. Box 25
NO-2007 Kjeller

POLAND
Centralny Ośrodek Naukowej
Informacji Wojskowej
Al. Jerozolimskie 97
00-909 Warszawa

PORTUGAL
Estado Maior da Força Aérea
SDFA – Centro de Documentação
Alfragide
P-2720 Amadora

ROMANIA
Romanian National Distribution
Centre
Armaments Department
9-11, Drumul Taberei Street
Sector 6, 061353, Bucharest

SLOVENIA
Ministry of Defence
Central Registry for EU and
NATO
Vojkova 55
1000 Ljubljana

SPAIN
SDG TECEN / DGAM
C/ Arturo Soria 289
Madrid 28033

TURKEY
Milli Savunma Bakanlığı (MSB)
ARGE ve Teknoloji Dairesi
Başkanlığı
06650 Bakanliklar – Ankara

UNITED KINGDOM
Dstl Knowledge Services
Information Centre
Building 247
Dstl Porton Down
Salisbury, Wiltshire SP4 0JQ

UNITED STATES
NASA Center for AeroSpace
Information (CASI)
7115 Standard Drive
Hanover, MD 21076-1320

SALES AGENCIES

NASA Center for AeroSpace Information (CASI)
7115 Standard Drive
Hanover, MD 21076-1320
UNITED STATES

The British Library Document Supply Centre
Boston Spa, Wetherby
West Yorkshire LS23 7BQ
UNITED KINGDOM

Canada Institute for Scientific and Technical Information (CISTI)
National Research Council Acquisitions
Montreal Road, Building M-55
Ottawa K1A 0S2, CANADA

Requests for RTO or AGARD documents should include the word 'RTO' or 'AGARD', as appropriate, followed by the serial number (for example AGARD-AG-315). Collateral information such as title and publication date is desirable. Full bibliographical references and abstracts of RTO and AGARD publications are given in the following journals:

Scientific and Technical Aerospace Reports (STAR)
STAR is available on-line at the following uniform resource locator: http://www.sti.nasa.gov/Pubs/star/Star.html
STAR is published by CASI for the NASA Scientific and Technical Information (STI) Program
STI Program Office, MS 157A
NASA Langley Research Center
Hampton, Virginia 23681-0001
UNITED STATES

Government Reports Announcements & Index (GRA&I)
published by the National Technical Information Service
Springfield
Virginia 2216
UNITED STATES
(also available online in the NTIS Bibliographic Database or on CD-ROM)

ISBN 978-92-837-0082-1