UNCLASSIFIED

AD NUMBER
ADB046175
LIMITATION CHANGES

TO:

Approved for public release; distribution is unlimited. Document partially illegible.

FROM:

Distribution authorized to U.S. Gov't. agencies only; Test and Evaluation; JAN 1980. Other requests shall be referred to Rome Air Development Center, AFSC, Griffiss AFB, NY 13441. Document partially illegible.

AUTHORITY

RADC USAF ltr 8 Nov 1982

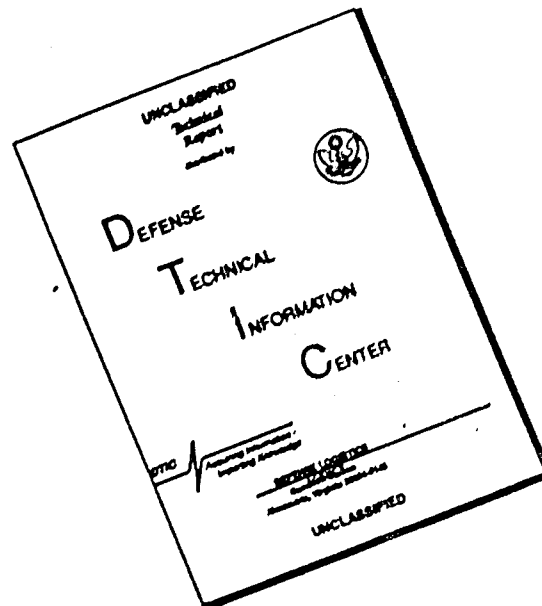THIS PAGE IS UNCLASSIFIED

AD $B046175$

THIS REPORT HAS BEEN DELIMITED
AND CLEARED FOR PUBLIC RELEASE
UNDER DOD DIRECTIVE 5200.20 AND
NO RESTRICTIONS ARE IMPOSED UPON
ITS USE AND DISCLOSURE.

DISTRIBUTION STATEMENT A

APPROVED FOR PUBLIC RELEASE;
DISTRIBUTION UNLIMITED.

DISCLAIMER NOTICE

THIS DOCUMENT IS BEST QUALITY AVAILABLE. THE COPY FURNISHED TO DTIC CONTAINED A SIGNIFICANT NUMBER OF PAGES WHICH DO NOT REPRODUCE LEGIBLY.

AD B046175

LEVEL II

TAC C³ DISTRIBUTED OPERATING SYSTEM STUDY

Operating Systems, Inc.

John R. Thompson
Enrique H. Ruspini
Christine A. Montgomery

ROME AIR DEVELOPMENT CENTER
Air Force Systems Command
Griffiss Air Force Base, New York 13441

80 4 3

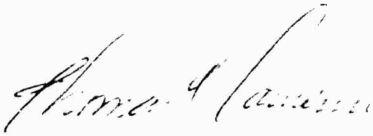This report contains raferences to the following classified document:

Tactical Air Forces Integrated Information System (TAFIIS)
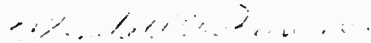Master Plan (U), dated September 1977, SECRET.

RADC-TR-79-360 has been reviewed and is approved for publication.

APPROVED:

THOMAS F. LAWRENCE
Project Engineer

APPROVED:

WENDALL C. BAUMAN, Col, USAF
Chief, Information Sciences Division

FOR THE COMMANDER:

JOHN P. HUSS
Acting Chief, Plans Office

631294

REPORT DOCUMENTATION PAGE	READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
RADC-TR-79-360	AD-B046 175	

4. TITLE (and Subtitle)	5. TYPE OF REPORT & PERIOD COVERED
TAC C^3 DISTRIBUTED OPERATING SYSTEM STUDY	Final Technical Report. 4 Dec 78 — 4 Sep 79
	6. PERFORMING ORG. REPORT NUMBER OSI-R79-045

7. AUTHOR(s)	8. CONTRACT OR GRANT NUMBER(s)
John R. Thompson Enrique H. Ruspini Christine A. Montgomery	F30602-79-C-0016

9. PERFORMING ORGANIZATION NAME AND ADDRESS	10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
Operating Systems, Inc. 21031 Ventura Boulevard Woodland Hills CA 91364	62702F 55812107

11. CONTROLLING OFFICE NAME AND ADDRESS	12. REPORT DATE
Rome Air Development Center (ISCP) Griffiss AFB NY 13441	January 1980
	13. NUMBER OF PAGES 120

14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office)	15. SECURITY CLASS. (of this report)
Same	UNCLASSIFIED
	15a. DECLASSIFICATION DOWNGRADING SCHEDULE N/A

16. DISTRIBUTION STATEMENT (of this Report)

Distribution limited to U.S. Government agencies only; test and evaluation; January 1980. Other requests for this document must be referred to RADC (ISCP), Griffiss AFB NY 13441.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

Same

18. SUPPLEMENTARY NOTES

RADC Project Engineer: Thomas Lawrence (ISCP)

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Distributed Operating Systems
TAC C^3 Distributed Data Processing
Computer Networking

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

This study addresses the interconnection of data processing facilities into a Tactical Air Force Integrated Information System (TAFIIS). A strawman distributed data processing architecture is presented to show system sizing and concept of operation. The study establishes a rationale for two types of computer networking: (1) a "mininet" for high-bandwidth interconnection of locally clustered computers for resource sharing and operation as a single processing "cell", and (2) a "maxinet" for interconnection of processing cells
(Cont'd)

DD FORM 1473

311774

Item 20 (Cont'd)

through long-haul communication links which will have variable bandwidths, low reliability, and susceptibility to saturation. A three-level operating system design is proposed by this study. Technology requirements to support the functionality of each of the three levels are identified.

CONTENTS

i

EVALUATION

The TAC C^3 Distributed Operating System Study Final Report
presents a generalized distributed system architecture for applica-
tion to TAC operations and also provides an assessment of the
technology needed to support future realization of this system
concept.

This effort applies to TPO-R3, Thrust D, "C^2 Information
Processing"; Subthrust 1, "C^2 Information System Structures;
specifically, Project 5581, Task 21, "Network and Distributed
Processing Studies" and Project 2530, Task 01, "Distributed Data
Processing".

This effort will be used as the basis for a follow-on effort
to implement an experimental distributed operating system as a
means of demonstrating the feasibility of the technology. In
addition, the technology assessment contained in the study will be
used in structuring future R&D efforts within the above referenced
TPO.

THOMAS F. LAWRENCE
Project Engineer

TAC C³ DISTRIBUTED OPERATING SYSTEM STUDY

Technical Summary

Operating Systems, Inc.

1. 1. TECHNICAL PROBLEM

The general problem being addressed by this study is the interconnection of data processing facilities into a Tactical Air Force Integrated Information System (TAFIIS). Specifically, this study deals with the issue of the operating system which will tie the individual data processing components into an effective information system in a tactical environment.

The goal of the distributed operating system is to exploit available technology for achieving an operational configuration for the mid 1980s which will have attributes survivability, flexibility, interoperability, and responsiveness to user information needs.

2. GENERAL METHODOLOGY

The study was conducted by examination of technology in areas of:

[1]   Computer networking

[2]   Operating systems

[3]   Communications

[4]   Large system development methodology

[5]   Distributed data base management

The criteria used to evaluate the applicability of these technology areas were: (1) operability in the tactical environment, (2) evolution of system modules, and (3) packaging for Tactical Air Force contingencies.

A strawman data processing architecture was used to illustrate system sizing and concept of operation.

3. TECHNICAL RESULTS

The study establishes a rationale for a distributed data processing architecture which recognizes two types of computer networking in its implementation. The first is the networking of computers in a local area where reliable high bandwidth communications channels are available (the mininet). The second is the interconnection of data processing elements over long-haul communications networks which will tend to have frequent outages, low bandwidth channels, and susceptibility to saturation. The operating system design concept which is advanced in the strawman architecture, utilizes three levels of operating system services. These three levels have been designated:

[1]   Constituent operating systems (COS) - set of heterogeneous manufacturer-supplied operating systems which support the internal functions of individual computers.

[2]   Mininet Distributed Operating System (MINIDOS) - supports the high-bandwidth interconnection of computers into a single data processing cell; primary functions are capacity extension and reconfiguration.

[3]   Maxinet Distributed Operating System (MAXIDOS) - supports the interconnection of processing cells through a communication network which has variable bandwidth, low reliability, and a dynamically changing connectivity; primary functions are external interfaces, handling of

S-1

critical system data and ensuring the survivability of the information system.

The study develops the rationale for the maxinet/mininet architecture and identifies technology that is required to support this concept of operation.

The system development process is identified as x an area that requires a non-classical approach in order to achieve the goals of modular evolutionary development and flexibility. The study proposes a concept of a Centralized Development and Staging Facility (CDSF) to overcome many of the pitfalls in large tactical system development.

3. IMPLICATIONS FOR FURTHER RESEARCH

The study identifies technology needs in the following areas:

[1] Tactical information model development

[2] Automated tactical requirement collection and analysis

[3] Distributed system simulation/emulation

[4] MAXIDOS/MAXINET interface study

[5] Criticality oriented data manipulation languages

[6] Critical data transmission algorithms/protocols

[7] Heterogeneous processor interfacing

[8] MAXIDOS/MINIDOS interfacing

[9] Probabilistic resource management

[10] Adaptive resilient algorithms

[11] MAXINET encryption techniques

[12] MAXIDOS authentication algorithms

[13] Multilevel/multiuser distributed security

[14] Global configuration management

[15] Cell reconfiguration procedures

[16] Extensible distributed data base management

[17] Probabilistic data management algorithms

[18] Heuristic procedure development

Many of these technology requirements are being addressed by current or planned Air Force programs. A testbed is required to focus these areas of technology development on the implementation of a working distributed data processing architecture for TAFIIS. This testbed should provide an environment where system developers have access to real data and the where the user can be an integral part of the system development and staging process. The orientation of the testbed program should emphasize the rapid conversion of technology into fieldable tactical information systems which are good but not necessarily optimized for all uses.

4. SPECIAL COMMENTS

One of the observations made during the course of this study was the rapid convergence of technology needs in different program areas toward a common set of functions which could be provided by a common distributed operating system. These common requirements are seen in the development of automated communication network control systems, distributed data base systems, and in the development of special applications systems involving multiple processors. Air Force and contractor groups involved in these types of programs should have a direct participation in the review of distributed operating system technology or testbed experiments produced by any follow-on work in this area.

TAC C³ DISTRIBUTED OPERATING SYSTEM STUDY FINAL REPORT

Operating Systems, Inc.

1. INTRODUCTION

The objectives of the TAC C³ Distributed Operating System Study are to:

- Define a viable data processing architecture and operating concept for the tactical command, control, and communications (C³) environment.

- Determine the role of the Distributed Operating System (DOS) in the operating concept.

- Identify and evaluate issues involved in the feasibility of the data processing architecture.

The Tactical Air Force Master Plan defines the general goals and operating constraints for a Tactical Air Force Integrated Information System (TAFIIS). The TAFIIS Master Plan document [Reference 1] defines the mission of the data processing facilities and identifies the types of users and functions to be supported by TAFIIS. For the purpose of this study, data processing will be defined to exclude the communications devices although the services and performance requirements of communications networks will be addressed as part of the architecture analysis.

The requirement for an integrated information system is based on the need to most effectively utilize the tactical units, intelligence, surveillance and identification, personnel, and logistics resources available in a Tactical Air Force deployment. Sharing of information is the key to surveillance and intelligence effectiveness. Coordination of air resources for defense or tactical air operations can be done effectively only through close communication between supporting and supported units.

Continual flow of unit status and material expenditure data to logistics support units is required to ensure the effectiveness of the resupply chain. Carefully coordinated and positive management of the electromagnetic spectrum will be vital in the Electronic Warfare threat environment. The ability to share weather and topographic analysis information will mean that more accurate and sophisticated systems can be used to support various mission requirements. The seemingly simple function of identifying friendly units requires an array of information sharing on operations, procedures, and visibility areas.

The data processing architecture of TAFIIS is intended to support the information handling requirements and improve performance over that achievable with manual means. Data processing itself is a high-cost resource that must be allocated effectively in the tactical environment.

1.1 Approach to Defining a Viable Data Processing Architecture

The approach that OSI has pursued in the definition of a viable dp architecture has been guided by three TAFIIS Master Plan concepts:

- The rolling force package

- Modular element evolution

- Survivability

1.1.1 The Rolling Force Package.

The rolling force package concept is aimed at assembling an information system that is appropriate for accomplishing a specific mission with a given force level in a specific geographic area.

This concept implies that data processing components will also be configured in accordance with the number and type of users and will support the missions and loading requirements imposed by the particular operating environment. The time required to mobilize the rolling force package should be minimal and the package should be extensible as the force level, mission requirements, or external interfaces change with phases of engagement.

1.1.2 *Modular Element Evolution.*

The TAFIIS Master Plan recognizes the realities of on-going system development in its stated objective of supporting continuing evolution of system elements. The goal is to modularize the functions of TAFIIS elements such that hardware substitutions, software upgrades, capacity expansion, and addition of new elements can be achieved as painlessly as possible. The life cycle of the TAFIIS concept will extend beyond the 1980s and computer hardware, software, weapons, communications, and intelligence systems will evolve through many generations within this period.

The modular evolution concept if implemented successfully should help to avoid the many pitfalls of large system development for the tactical environment which have so frequently been experienced. These pitfalls include:

- Lengthy development cycles which result in obsolete hardware reaching the field.

- Isolation of end-user from the development cycle resulting in systems which don't address the real problems.

- Ignoring the system operating and maintenance skills required to adapt the system to the environment, resulting in unworkable systems or excessive staff costs to maintain the system.

1.1.3 *Survivability.*

The TAFIIS Master Plan goal of high survivability tends to have an overriding effect on many types of design decisions such as in data protection, redundancy, and alternate operating modes. Interpretation of survivability requirements cannot be applied directly to data processing requirements for capacity redundancy, data transmission and storage protocols, and program mobility without consideration of the human element involved. Assumptions about elements backing each other up must take into account cross-training requirements and background information required to accomplish the mission. Data transmission and storage protocols must take into account the alternate human-readable media used to back up the computer and digital transmission media. The increased mobility of smaller command elements should not sacrifice the group decision-making capabilities of current larger, less-mobile (and less survivable) command centers.

1.2 Architectural Approach

The approach which has been taken to the selection of a data processing architecture is based heavily on a humanistic model of information handling. Key features in this architecture concept include:

- Decentralized control of TAFIIS data processing elements

- Non-classical information structures

- Data base concurrency control through loose coherence algorithms

- Adaptive resource management algorithms

The development approach is as important as the data processing architecture because an attempt is being made to achieve an open-ended system configuration. The key concepts in the

development approach include:

- Centralized system development and deployment staging with consolidated use of technology, data, communications, software development tools, and testing resources

- A "make it work" development philosophy with direct involvement of users

- Heavy emphasis on user and data base preparation prior to deployment

- Sliding readiness time windows within which the components of the information system must be fieldable

- Continuous and overlapping system upgrades being performed by a combination of contractor, government project staff, and blue-suit personnel.

1.3 Summary of Document

The concept of operation for the TAFIIS data processing system is discussed in Section 2.0 of this document. The concept of operation describes the user population, preparation of the information system prior to deployment, and adaptive mechanisms for achieving survivability, flexibility, and gradual evolution. The functions of the data processing architecture are described from a generic standpoint rather than specific applications. TAFIIS functions were defined in this way in order to arrive at the requirements for a distributed operating system which is capable of supporting this type of data processing architecture.

A strawman architecture is described in Section 3.0 as a means of presenting the design issues of the data processing configuration and the associated operating system. The strawman architecture is presented as a distributed data processing configuration which uses two levels of computer networking, the maxinet and the mininet. The operating system is logically structured into three levels in order to address the functional and performance differences of operating system services performed within one processor, the mininet, or across the maxinet. Software structuring is also included in the description of the strawman.

The rationale for the strawman architecture is presented in Section 4.0. This section identifies open architectural issues and assesses the risk involved in the selected architecture approach. The key areas addressed include:

- interprocessor communications

- resource management

- security

- configuration management

- data base management

- system development

Section 5.0 provides a matrix which summarizes the technological requirements required to proceed with the development of a data processing architecture to support the TAFIIS operating concept. Technology requirements are evaluated from the perspective of availability of technology to satisfy requirements, development time, and risk.

2. SYSTEM OBJECTIVES

This section discusses the primary purposes for which a data processing configuration would be deployed into a tactical environment to support the information requirements of a Tactical Air Force command and control organization.

2.1 Operating Environment

The potential operating environments for TAFIIS have extreme variability in terms of:

- Force size and composition
- Type of engagement
- Length of engagement
- Types of missions
- Mission activity level
- Opposing force threat
- Knowledge of opposing force threat
- Terrain effects on communications, disposition, and operations
- We ither effects on communications and operations
- Electronic warfare

These environmental factors all contribute to the requirements for the data processing architecture which is deployed to support the TAF.

2.2 Concept of Operation

The TAFIIS is an information system that will be composed of human elements, automated data processing, communications services, and hardcopy data. The scope of operation has been purposely broadened to include the development and garrison states as well as tactical deployment in order to address the goal which has been so elusive in past systems -- that of shortening the system development cycle. Inclusion of these two phases of system operation is integral to the proposed approach which places more emphasis on the system developm... ocess than on producing

an all-encompassing system design. This approach foregoes the assumption that requirements are definable and that a system can be developed to meet those requirements.

Instead, the operating concept assumes that requirements are continuously changing with the world situation, the threat, and the state of technology. The operating concept presented in the following paragraphs delineates the nature of the user population, the preparation of the system for deployment, and adaptation of the system to the operating environment. There is no definition of a normal operating mode or typical environment. Instead, the system designer will provide an architecture which will utilize whatever information and resources that are available to prepare for contingencies and adaptive mechanisms for achieving adequate performance in the operating environment.

Optimization is not addressed -- more important is the concept of making the system work under whatever operating constraints exist. The adaptive mechanisms which are integral to the operating concept are important to achieving evolution of the system as well as effective management of resources in the tactical environment.

As a key feature in each of the following discussions is the role that the operating system plays in the use of data processing resources.

2.2.1 User Population.

The question of who uses TAFIIS is important in scoping the data processing functions to be supported. Because of the goals of continuing evolution and the need to shorten the development cycle between states of service, a class of system support users has been added to the conventional definition of the user population which includes only analysts or operators involved directly with the application. In addition to the

analyst class users such as planners, operations officers, liaison, intelligence, etc., the system support user class expands the user population to include:

[1] Data base specialists

[2] Software maintainers

[3] Performance evaluators

[4] Security officers

[5] Trainers

[6] System operators

[7] Hardware maintainers

[8] Communicators

The system support user when addressed in user interface requirements expands the command repertoire to be processed by the data processing system. The skill level profile of the system user also changes dramatically when the system support users are thrown in with the persons that staff the traditional analyst and duty positions associated with command and control elements. Even the skill levels of persons in operational duty positions can vary. Some system functions can be handled by data entry clerks with very little training. Intelligence analysts will spend considerable time on the system reviewing message traffic and updating intelligence files. Operations officers may spend most of their time examining status and situation displays and communicating with remote elements. Training may vary widely in terms of length and depth. Expertise in use of particular system functions may depend on how often that function is used. Management personnel may also be expected to use the data processing system for certain functions.

Because of this wide diversity in the user population, the operating concept proposed for TAFIIS assumes that the user interface must be tailorable for the skill level and function of the particular user. This operating concept for the

user interface has several important impacts on the data processing configuration. Because there are obvious benefits in using standard user terminals and interfaces in reducing inventory cost and in minimizing the interoperability problems, some form of user interface standard is necessary. However, if the standard is applied to all aspects of the man-machine interface, then the system has very little adaptivity for different user skill levels and context of system use. From the standpoint of terminal standardization, only the electrical and external characteristics are important and not the interpretation of commands.

Both of these benefits can be accomodated with an approach to the user interface language which involves the use of a common user language substrate. The substrate provides a singular common user language across all user types and is in a keyboardable textual format that is adaptable to most types of dumb and intelligent terminals. The basis of this approach has been described in some of OSI's recent studies for the Army Research Institute [Reference 2] and is currently being evaluated for use in several large multi-user system designs.

There is a secondary benefit from the use of a keyboardable substrate that occurs during system development. Because the software designer needs only to work with the internal substrate form of user commands, he is effectively decoupled from the day-to-day vacillations in user preferences for function keys, menu layouts, display formats, etc. Fewer changes in functional command forms will result in reduced application development time.

At the same time that the user interface looks extremely varied from the user population perspective, there are also many commonalities at the generic function level from the machine implementation perspective. Such generic

functions tend to fall into the areas of:

[1] compose/edit

[2] search

[3] maintain/define

[4] retrieve/output

[5] route

[6] compute

[7] user aids

[8] status

From the perspective of the DOS, the user interface also has common requirements across the user population. These requirements include audit trail, command/message sequencing, device transparency, directory functions, and access control.

The approach in implementing the user interface can be summarized as one of providing user specificity from the user's perspective and commonality from the machine perspective.

2.2.2 System Preparation.

Although the natural tendency is to think of only the operational period following deployment, the preparation prior to deployment is necessary to the success of the rolling force package concept. The initial data base carried into the field will be an extensive composite of contingency planning data, intelligence data (such as order of battle, preplanned target information, and ethnographic data on the country), network data on theater and national interfaces and reporting requirements, and internal working formats and procedures. If the system is expected to work efficiently in the field, users must gain familiarity with the use and capabilities of the system through command post exercises and field test operations.

Any preparation that can be done in a garrison state will reduce the time needed to mobilize. Preparation also increases the initial utility of the

system once it is in place. Therefore, there is a high value that can be attributed to an operating concept which ensures the continual readiness of the command and control data processing support. It is a conclusion of this study that the operating concept should emphasize continual use of the data processing system in garrison and training operations such that the transition from peacetime to wartime use is smooth and that users are aware of how to use the system.

Intelligence activities are never in a dormant state and are in fact more critical during peacetime. Intelligence preparation of the battlefield (IPB) is gaining more attention from commanders because of the benefits of using intelligence products for contingency planning at lower echelon levels. National intelligence collection capabilities that are normally thought of as being for strategic use exclusively have been shown to be invaluable to the IPB process and to subsequent exploitation of real-time intelligence gathered in the field. This approach to intelligence preparation requires continued access and exploitation of intelligence data by analysts who will support the field commanders.

Data processing system preparatio should parallel that of aircraft systems that are maintained and supported in a continual state of readiness.

2.2.3 Message Handling.

Communications are not used in the same manner during peacetime as they would be in the field. Field communications require use of unreliable radio nets and emphasize use of communication security (COMSEC) procedures to prevent the opposing force (OPFOR) from exploiting transmissions for targeting or intelligence purposes. Unfortunately, COMSEC procedures are rigidly adhered to only for strategic operations. The tactical user has

difficulty adhering to COMSEC discipline because of inadequate communications equipment and unfamiliarity with COMSEC requirements. Because users are dependent on communications for their authority and ability to conduct operations, they will place a high premium on any communication mode, even if it is not secure. The TAFIIS Master Plan identifies joint service programs such as the Joint Tactical Information Distribution System (JTIDS) and Tri-service Tactical Communications (TRITAC) which are aimed at improving secure communications services.

The data processing configuration plays a role in communication security from several points:

[1] digital communications can be encrypted more easily and rapidly than voice

[2] digital messages can be transmitted in less time and thereby reduce the probability of intercept

[3] digital messages can utilize store and forward services and alternate routing to compensate for intermittent or unreliable communication networks; voice relays are time-consuming and error prone

[4] digital message communications can be controlled by consistent CEOI (Communications Electronics Operating Instructions) to avoid COMSEC violations

[5] digital messages can be logged and stored by the data processing system

[6] digital messages can be automatically disseminated to multiple destinations with no additional effort on the part of the sender.

There are several pitfalls in the use of digital message traffic to replace voice communications. The first is that there are both formal and informal communications between C^3 elements. The informal communications are vital to achieving the feedback which allows the system to adapt. Informal communications, however, are less adaptable to digital message format although successes have been achieved to a great extent in the intelligence community with INDICOM and OPSCOMM networks and in ARPANET with the electronic mail system. Users must be intimately familiar and comfortable with these systems for effective use in informal communications.

A second pitfall is that the language used for digital message transmission is not always suitable to the individual user. Message formats have data fields for use in communication control, message filing, access control, security classification and downgrade, and activity codes. The message language may be further complicated by a content structure designed for computer readability or computer generation. The vocabulary may be complicated by the use of abbreviations and acronyms peculiar to an activity, the service unit creating the message, or simply the preference of the sender. On top of format and vocabulary peculiarities, typing and transposition errors will occur along with possible transmission garbles. These language problems translate into interoperability problems which become extremely pronounced when digital message traffic forms the basis of communication between different service elements or allied elements.

Reliability factors are also detrimental to user acceptability of digital message handling. If the computer is storing and controlling access to message data and then the computer fails, the user may be without backup information. Even worse, the computer may not be able to restore the message data when the system has been repaired and

restarted. Although reliability problems have caused users to view digital message handling systems with skepticism, good design approaches and diversity in data protection media can overcome these problems.

These problems can be largely overcome in order to derive the operational benefits of digital communication. The approach to digital message communication must be based on linguistic principles rather than optimization of computer readability and data transmission. Most importantly, the users must be familiar with the use of digital communications and utilize digital communications as part of the normal operations in peacetime. If this approach is taken as a generalized concept, users will know how to use TAFIIS data processing facilities when they are deployed and will have adapted the user interface to their specific requirements for communication and data base usage.

2.2.4 Adaptivity Mechanisms.

Preparation of the data processing system prior to deployment ensures that the system will have immediate utility. Preparation of the system will also include the provision for contingency operation and adaptation. Once the system is in place it must adapt to the environment, the missions in process, and survive the opposing force threat. Over the long term, the information system must adapt to the addition of new elements, missions, and functional organization. The adaptation mechanisms which are proposed in the concept of operation include:

[1] Communication network adaptivity

[2] Distributed Data base management

[3] Robust information structures

[4] Decentralized resource management

2.2.4.1 Communication Adaptivity.

Communication adaptivity involves adjustment of information flow to the available physical communication network structure and adjustment to the changing information needs of individual users.

The field environment is highly dynamic from the perspective of the communications services available to the data processing system. Communication bandwidth between centers will vary as links are established and as outages occur. Loading will vary with the 24-hour operational cycle and over the long term as the force level changes. The mix of missions and support activities will vary with the phase of battle.

Two types of adaptivity mechanisms are proposed in the operating concept for dealing with the tactical constraints placed on information flow. The first concept is to use monitoring mechanisms in the communication network to detect when the load is too high for the available communication bandwidth. This information would be fed back to the data processing system so that an adjustment could be made to mode of network operation to reduce the demand for network communication services.

The second concept is an approach to adjusting the logical information flow to the needs of users as their requirements change due to:

[1] change in geographic area of interest

[2] change in functional responsibility (mission, duty position, etc.)

[3] change in the tactical situation (type of engagement, threat)

[4] requirements to assume the area of responsibility for elements or individual users that are inserted or removed from the system.

The concept for adjusting the

Information flow to these types of changes cannot be reasonably charged to the information sender because hundreds of recipient users could be involved in the distribution of a given item of information. Instead, the information is made available to every major communication node in the network and individual users may gain access to this data on the basis of interest and access authority.

The problem of how to disseminate information efficiently to multiple nodes (that may number in the tens) is discussed as part of the data base distribution and data criticality definition issues.

2.2.4.2 Software Adaptivity.

During deployment, software adaptivity will include directory updates, addition of capacity for additional users, installation of applications modules to support new missions, and reconfiguration to continue operation in degraded modes.

The hardware configuration is not expected to remain stable in a deployed configuration because of the high attrition of computer equipment in the hostile field environment. New interfaces to communication devices are expected to be commonplace as the communication networks are expanded to include new elements or are reconfigured for new dispositions of elements.

Information structures will have to adapt to the specifics of the missions and information requirements of particular users. The user population will not be stable due to shift changes, personnel changes, and reorganization. Interoperability requirements with other service or allied elements may require compromise or modification of data formats used in communication.

In the proposed concept of operation, all software upgrades made in the field will be treated the same as data base

management problems and information or data files transferred to field elements to accomplish the upgrade will be treated as tactical information flow. The obvious impact of this concept of operation is that software must be treated as critical data and must be in a form which is compatible with field communications services.

2.2.4.3 Long Term Adaptivity.

Adaptivity in the garrison environment is also a critical system requirement. As new weapon systems, communications capabilities, and intelligence collection capabilities are introduced into the TAF inventory or as obsolete systems are phased out, there will be significant changes to the data processing system. Although the network configuration will change on a much slower basis, the garrison configuration will undergo changes in directories, software modules, and hardware.

Each time a change is made which affects information structures, operating procedures, or reliability the system will have to be reverified for operational use. Many large systems have been plagued by architecture problems which make each upgrade a painful and time-consuming process of retesting and removal of bugs introduced by the upgrade. This retest period generally results in a significant period of system downtime and poor reliability.

As a general concept of operational upgrading, it will be assumed that both old and new functional capabilities can be maintained in the data processing configuration at the same time. This means that readiness would not be severely impacted by the introduction of new capabilities and that the old system would be available as a fallback. This concept places a burden on the communications system to handle duplicate message traffic. Capacity of the data processing system must be sufficient to support two concurrently

active configurations. It also places a burden on the system configuration control mechanisms to handle multiple software, hardware, and data base versions concurrently. From the user's perspective it means duplication of staffing to run both systems.

None of these implications are inconsistent with the goals of TAFIIS or the general operating concept being presented for the following reasons:

- TAFIIS will have extra capacity to meet expansion requirements and backup contingencies.

- Communications networks will be constantly handling duplicates and retransmissions because of reliability problems.

- Operating groups must be capable of providing multiple staffing for every duty position to meet 24-hour day tactical manning requirements. Parallel operation of two systems could be used as a training vehicle for cross-training of operators for multiple duty positions and for preparation for degraded mode operation.

2.2.4.4 Evolution Mechanisms.

An important question is how to produce the higher skilled staff that will be required to perform the complex activities of software and hardware integration and testing required to achieve continual system upgrades. The Air Force, like other services, is faced with personnel shortages in high skill areas such as data processing and hardware maintenance. Software development and testing is too complex to be undertaken by inexperienced personnel. For these reasons, an operating concept has been assumed that is based on the initial phase of system upgrade being performed by a centralized development and staging facility. Centralized software maintenance and hardware testing would concentrate the available

high skill personnel.

Some additional benefits to be derived from the centralized maintenance concept might be in promoting standardization of software modules and configuration control in areas requiring interoperability with other systems.

Since some amount of software maintenance must still be done in the field, provisions should be made for simplifying these procedures so they can be done by lower skilled personnel. This area is of particular concern to the DOS study because the operating system is the key factor in determining the complexity of software maintenance. A corresponding question is how to make the skills of the central facility available to the field and how to make field problems known to the central facility.

2.3 Desirable System Functions

The functions of the TAFIIS data processing architecture can be defined in terms of generic processing functions, and in terms of the specific roles of the operating system, applications software, system software, hardware, and communication. The focus of this study is specifically on the role of the operating system although boundaries must be defined between the different components.

2.3.1 What Is the TAFIIS Data Processing Mission?

Other ways of addressing this question are "Why is data processing required in the tactical environment?" or "What can automated data processing do that the user can't do equally well for himself?". The most basic operations that the computer can do for the user include remembering, computing, and electrical interfacing. The several hundred machine instructions which manipulate these basic operations can be organized into complex functions through several levels of intermediate language abstraction. From the

perspective of the data processing system (not the user's) these complex functions generally fall into the areas of:

[1] composition and editing of data

[2] search for data records

[3] define/maintain data base

[4] retrieve and output data

[5] routing of data between users

[6] user aids/computation

These general categories may translate into hundreds of machine-recognizable commands. In the context of a specific user and command argument combinations, the command repertoire is virtually unlimited in scope.

Unlike the robust communication between humans, communication with a machine must be explicit and without ambiguity. Although the initial cost of automation is high, the advantage of the computer is that once a computational sequence has been defined explicitly, it can be remembered and repeated with a much higher speed. The goal of the system designer is to arrive at some midpoint between the lowest level machine instructions and the user's notion of what he wants the machine to do.

In this respect, the state of the art in man machine interaction has advanced significantly due to research in natural language information processing techniques and the quantum reductions in computer hardware costs. Early computer systems did not consider natural language interfaces with computers because of the high costs of maintaining large dictionary structures and computation of lengthy parsing algorithms. Even storage space restrictions have had a detrimental effect on the user interface. In the attempt to save storage space in the computer, the system designer forced the user to remember arbitrary coding schemes and

remember the context of data values with no field names or descriptions in the data records.

Systems which were developed in this manner many years ago have left a legacy of non-human oriented message and data record forms which persist in many data processing systems yet today. This is not to say that brevity in data records cannot be useful (and it can be in tactical communications). However, these forms can and should be developed along human-oriented linguistic principles and not in accordance with outdated computer cost constraints.

Because of the importance of information structure utilized in digital message communications and in the man-machine interface, special attention has been given to this subject during the course of this study. The results of this study are presented in Appendix A, "Implications of Advanced Data Structures for Tactical Communications".

2.3.2 The Role of the DOS In the Data Processing Architecture.

The components of the TAFIIS data processing system architecture include:

[1] Processor and peripherals

[2] Communications devices

[3] User terminals

[4] Operating system software

[5] System software

[6] Applications software

The general nature of each of these components are easily understood in the context of a single processor system. When distributed processing is involved with multiple processors, the context of these definitions are ambiguous depending on the perspective from which the system is viewed. As a case in point, a remote computer system may access the data base of a second host computer by emulating the

Interface of a local terminal. The remote system can be viewed as a set of components or as a terminal. Similarly, a programmable communication device may be viewed as a piece of hardware during use but should be viewed as a cooperating system with its own operating system and software modules from the perspective of protocol development and maintenance.

The only point at which there is an awareness of the real features of component characteristics is from the perspective of the operating system which ties the components into a processing sequence. The operating system is of little importance in a singular process sequence such as process control where the algorithm is static. The operating system gains immense importance in a multi-programming environment where the software structure is not static. Distributed processing introduces a new dimension where neither the software nor the hardware configuration is static.

In the concept of operation assumed for the TAFIIS data processing system, the operating system is assumed to perform the types of functions listed in Table 1.

The distributed operating system is assumed to encompass the functions of the operating system in each processor plus any extensions required to operate the data processing resources in a distributed configuration. The extensions required to support a distributed processing configuration generally fall into the areas of:

[1] Directory services to keep track of users, software modules, and data.

[2] Allocation of resources shared by multiple nodes such as communication services, global data bases, excess capacity (terminals, backup processors, input/output devices)

[3] Scheduling of tasks involving interprocessor interaction

[4] Access to global system software

[5] Performance monitoring (processor status, communication service status, interprocessor interaction status, user status, device status, resource utilization)

[6] Degradation handling and system recovery (error detection, error reporting, fault isolation, restart, and reintroduction of failed unit after repair)

[7] Interprocessor communication (event occurrence, data mapping, data transmit and acknowledgement, remote terminal access)

[8] Multi-level data security, access control, release control and audit trail.

Table 1. Hierarchical Classification of OS Services

1.0 SERVICES TO PROCESSES

    1.1 Process management

        1.1.1 Process Initiation

        1.1.2 Process termination

        1.1.3 Error Recovery

        1.1.4 Interprocess mediation

            1.1.4.1 Priority Assignment and Management

            1.1.4.2 Coordination Primitives

            1.1.4.3 Communication

        1.1.5 Environment Management

            1.1.5.1 Storage Management (for specific processes)

            1.1.5.2 Exceptional Condition Management

            1.1.5.3 "Privileged" Process Services

            1.1.5.4 Process Limit Monitoring

    1.2 Resource Management

        1.2.1 Processor

            1.2.1.1 Scheduling, Conflict Resolution, Deadlock Prevention

            1.2.1.2 Allocation

            1.2.1.3 Protection

            1.2.1.4 Error Detection and Recovery

        1.2.2 Timing Services

            1.2.2.1 Scheduling, Conflict Resolution, Deadlock Prevention

            1.2.2.2 Allocation

            1.2.2.3 Protection

            1.2.2.4 Error Detection and Recovery

        1.2.3 Main Storage Management Global Resource Management

            1.2.3.1 Partitioning

            1.2.3.2 Segment Control

        1.2.4 Secondary Storage Management (Global Resource Management)

            1.2.4.1 Scheduling, Conflict Resolution, Deadlock Prevention

In actuality the distributed operating system must be developed within the capabilities and constraints of the hardware and software of the constituent processors in the distributed processing network. The directives which fall into the domain of the DOS rather than the constituent operating system (COS) in fact must be handled through the COS and executed by tasks which are scheduled by the COS. For convenience, the directives and software used to deal with functions which fall into the category of distributed processing can be defined as being in the DOS rather than the COS domain. The convenience of this definition lies in defining differences in performance requirements between DOS and COS functions and clearly identifying the overhead of distributed processing operation.

2.3.3 An Approach to Data Processing Requirements Definition.

To this point, only the general function and operation of the TAFIIS data processing system has been defined. Performance requirements which can be used to determine size and cost of the data processing architecture include capacity, speed, and reliability. Determining performance requirements for these three factors is a non-trivial task because of the many operating environments and loading levels that the system must operate under. Time is also a consideration from the perspective that components can be replaced and anticipated operational contingencies may never materialize.

A full requirements definition either from the perspective of functionality or performance is neither possible nor desirable. One of the pitfalls of prior large system developments has been the assumption that requirements are stable and a system can be developed to satisfy those requirements. As a result, at the end of a ten-year development cycle, the system will not be responsive to requirements which have changed during the development period. If the developer attempts to follow the ups and downs of system requirements on a day-to-day basis for a very large system, schedule delays ensue because of the complex interfaces which tie the components of the system together.

A final factor in the development process is that the system affects the environment in which it operates. Electronic warfare tactics stimulate the development of countermeasures and those stimulate the development of counter-countermeasures. The existence of an effective data processing system will cause that system to be a high value target for the enemy which will further increase the threat to the information system survivability. Requirements will continually change because the threat will continually change.

2.3.4 How Well Should the TAFIIS Perform Its Functions?

The more qualitative performance measures of flexibility, adaptivity, predictability, and operability may hold out a better guideline for performance requirement definition than capacity, speed, and reliability. The performance criteria of survivability is more of a human element process than one of measuring redundancy, data protection, and alternate data routing. Mobility, degraded mode operation, and self-sufficiency which are major factors in C^3 element survivability are highly dependent on the human element in the operating concept. Backup operating modes will work only if personnel are trained to use them. Elements can be mobilized only if the entire element and its communication services and other support can be relocated also. Greater mobility implies smaller self-sufficient elements. Frequent mobilization means that operations may be intermittent unless backup elements can

immediately assume the responsibility of the moving unit. In all of these concepts there is an implicit emphasis on highly-skilled and cross-trained operators who can not only operate independently but can also assume the functions of other elements. There does not appear to be this trend in the current TAF doctrine or training. The configuration of the current Tactical Air Control Center (TACC) is large, highly centralized, and relatively immobile because of complex communications and colocation with other elements. Because of the vast numbers of equipments and personnel, the TACC is readily identifiable by the OPFOR and is a high value target.

Placement of the TACC in a rear area will improve its survivability and at the same time reduce its effectiveness because of reduced direct communications to tactical units and forward C^3 elements.

Communications availability may be in many cases the overriding factor in determining the organization and disposition of major command and control elements. Communications capability and the survivability of those communications will be a continuing design issue in the data processing architecture analysis. Communications technology for computer networking is expected to change significantly during the life cycle of TAFIIS. Likewise, the threat to those communications services will also change.

2.3.5 Goals for the TAFIIS Data Processing.

The general goals for the data processing system of TAFIIS as stated from the user's viewpoint would include the following:

[1] Make information accessible to users who need it

[2] Improve the throughput of time-sensitive information

[3] Support the local data base needs of users

[4] Make available global data bases which are needed for planning, coordination, threat assessment, targeting, intelligence production, and friendly force status monitoring

[5] Provide reliable dissemination of messages carrying requirements, tasking, warnings, and status information

[6] Provide extensive degraded mode operating capability and rapid recovery of system functions after failure.

Almost all of the above items reflect deficiencies in the current operating capability.

3. STRAWMAN ARCHITECTURE

This section is devoted to the description of a strawman data processing architecture which addresses the operating concept and goals of the TAFIIS Master Plan. This architecture is presented as a means of identifying and discussing issues involved in the use of distributed data processing in the tactical environment and the approach to the design of a distributed operating system (DOS). The issues and rationale for choosing the strawman architecture are presented in Section 4.0.

3.1 Top Level Architecture

A physical view of the strawman data processing architecture is shown in Figure 1. The components of the distributed processing architecture are:

- The users -- both analyst and system support classes

- The processing "cell" composed of one or more colocated processors and supporting one or more users.

- Network communications termed the "mininet" for interconnecting processors within a single cell.

- Network communications termed the "maxinet" for interconnecting cells and external elements.

- The operating system -- hierarchically organized by constituent processor functions (COS), mininet functions (MINIDOS), and maxinet functions (MAXIDOS).

- The functional software components -- functional threads composed of task groups in specific cells and tasks in specific processors.

- Data -- structured according to its context and handling constraints within the maxinet, mininet, and constituent processors.

The strawman architecture uses distributed processing to achieve the following operational objectives:

- Geographic dispersal

    - for field of view (Control and Reporting Center, CRC; Control and Reporting Post, CRP; Air Surveillance Radar Team, ASRT)

    - for colocation with tactical units (Tactical Unit Operations Center, TUOC; Airlift Control Element, ALCE; Direct Air Support Center, DASC)

    - for protection in rear areas (Air Force Component Headquarters, AFCH; Tactical Air Control Center, TACC; Airlift Control Center, ALCC; Tactical Air Base, TAB)

- Capacity flexibility -- multiple processors to extend the capacity of individual TAFIIS elements to handle a wide range of mission support activities, number of users, communications, and data base requirements for varying TAF deployments or changes occurring after deployment

- Survivability -- physical dispersal is necessary for protection against single strike loss.

- Activity and data security -- activities and intelligence data involved in sensitive collection operations must be protected by physical isolation.

- Modular implementation -- system components should retain a level of standalone operability that facilitates independent development and evolution without creating interoperability problems.

3-1

COMPONENTS	QTY
USERS	X00
CELLS	X0
REMOTE CELLS	X0
TASK GROUPS	X000
TASKS	X0000
MAXINET	1
(MODES)	X

CELL 2

CELL 1   CPU 1

COMM.

MAXINET

LONG-HAUL
COMM.

CPU 2

REMOTE
CELL

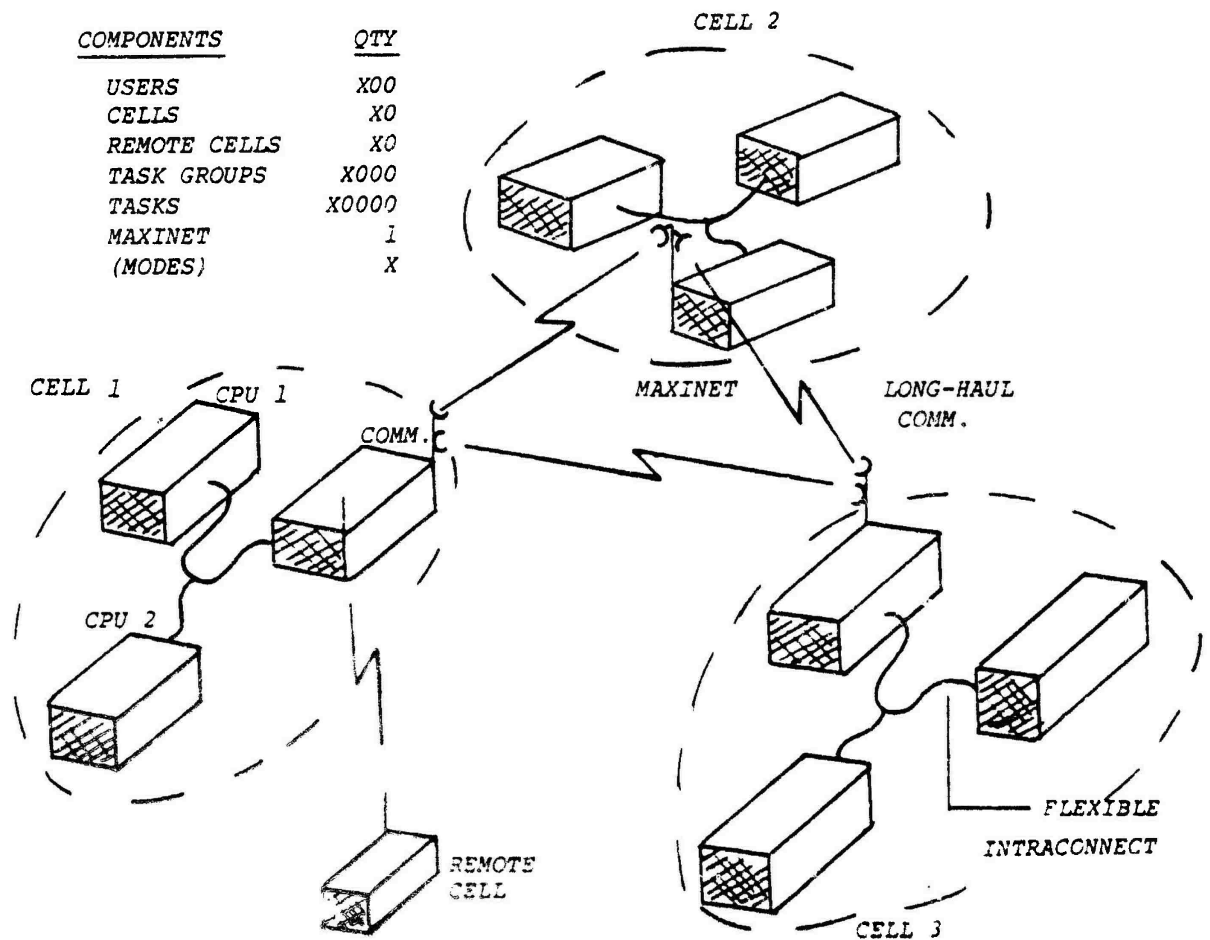FLEXIBLE
INTRACONNECT

CELL 3

Figure 1. Physical View of the TAFIIS Data Processing Architecture

3.1.1 Physical Dimensions.

The number of cells in the distributed processing network could reach into the tens in a large scale deployment with multiple Army corps elements and multiple AF wings. Elements such as the TUOCs, DASCs, and CRPs would be duplicated to cover the larger geographical area and additional tactical unit interfaces. Smaller contingencies might involve a single command and control element and one aircraft detachment. Sizing within elements must be based on load factors such as geographic area of interest, number of active missions, and level of threat from opposing forces. Neither function nor loading is uniform between any two TAF elements.

3.1.1.1 Number of Users.

The total user population of TAFIIS could number in the hundreds. Not all users would require continuous access to terminals and users in a tactical environment would operate in shifts. User terminal loading projections must take into account peaks in activity cycles for planning, operations, and reporting. At a major element like the TACC there could be as many as a hundred users with many unique functional and terminal access requirements.

3.1.1.2 Communications.

Two types of communication networking are required in TAFIIS: long-haul communications with links up to hundreds of kilometers long, and local links in the 100 meter to 1 kilometer range. The interconnection of elements through long-haul links will be referred to in the strawman architecture description as the "maxinet". The organic communications of TAFIIS can provide multichannel voice and digital networks which span an area in the order of 10^6 square kilometers in size.

External elements that would be tied into TAFIIS by the maxinet include:

- Theater elements
- World Wide Military Command and Control System (WWMCCS)
- National support elements.

These elements would tie into the maxinet via Defense Communication System (DCS), AUTODIN, or Intelligence Data Handling System Communications (IDHSC) links. These links provide worldwide communication services in addition to the organic communications of TAFIIS.

Networking of data processing facilities that are colocated at a TAB or at the AFCH can be achieved with high-speed bus technology. Networking of processors at a local level will be referred to as the "mininet". Special high-speed links would also be available to tie in forward radar surveillance teams and airborne platforms. These links may be considered as either a part of the mininet or the maxinet depending on whether the element is a subordinate unit or a command element, respectively.

The number of nodes or mininets that would occur in a TAF deployment would depend on the geographic constraints, force level, mission, and threat level. A maximum expected number of nodes would be in the tens to account for multiple DASC, CRP, and TAB elements.

It is assumed that a moderate degree of success will be achieved with the Adaptive Communication Control System (ACCS). The ACCS will provide a high degree of adaptivity and robust interconnectivity to the maxinet.

Even with the ACCS the maxinet will be susceptible to intermittent availability and saturation problems. Electronic warfare (EW), sabotage, anti-radiation missiles and mobilization will all contribute to maxinet performance problems. Node outage from mobilization, attrition or link outage will require frequent network reconfiguration. Reconfiguration

time after mobilization is more likely to be a function of repair time for transport damage and skills of communicators rather than time to recover routing tables and directories.

Long-haul communication links are likely to be sized more on the basis of equipment availability and past experience rather than on a dynamic allocation basis. Link capacity will be gradually adapted to needs by addition of channel capacity or new links. Maxinet links may be implemented with microwave, troposcatter, satellite, airborne relay, or line-of-sight radio nets. Access to voice radio nets is a matter of radio availability, electromagnetic propagation path, and interference from terrain or other radiation. However, voice nets are easily overloaded and are susceptible to jamming. Long haul communication links implemented with microwave or troposcatter systems will have multi-channel and multi-mode capability but will have longer setup times when redeployed.

3.1.2 Software Structure of Strawman Architecture.

The software structure of the strawman has been adapted for the maxinet and mininet physical structure. The software structure was selected in recognition of the constraints of the maxinet and mininet communications services and an attempt has been made to reduce the sensitivity of system operability to maxinet communication availability. As a result, the operating system is structured into three logical levels with different performance characteristics. As shown in Figure 2, these three levels are:

[1] constituent operating system (COS) of individual processors

[2] mininet distributed operating system (MINIDOS)

[3] maxinet distributed operating system (MAXIDOS)

Applications software is structured from a top-down perspective in terms of directory and resource allocation functions performed by the operating system. At the lowest level, the COS recognizes software at the task level and allocates execution time, program residency, and input/output bandwidth from the resources of a single processor.

At the MINIDOS level, task groups which perform specific functions must be recognizable because of relocation requirements for backup modes. Resources are allocated only for the purpose of longer term load balancing and prevention of saturation of common local resources such as the mininet bus and the pipeline into the maxinet. The MINIDOS must maintain directories of primary and backup copies of all critical data including: message traffic, data bases, and software. The directory services do not have to be at the same level of detail as at the COS. For instance, the COS will recognize software at the task level while the MINIDOS need only maintain directories of complete task group locations. The protocols used for storing and retrieving data recognized by the MINIDOS must encase the COS protocols in order to guarantee data protection and recovery from errors at the single processor level. The MINIDOS would bear the burden of recovering a functional process which was aborted due to loss of a processor, user terminal, or input/output device.

The MINIDOS is also the main factor in controlling the allocation of system resources. By controlling the directory of task groups which execute applications plus controlling the bandwidth of communications and I/O channels, the MINIDOS has effective control over all local processing resources and one node of the maxinet communications network.

LOGICAL NETWORKS
OVERLAYED ON PHYSICAL
COMMUNICATION
NETWORK

SURV.

WX

TACTICAL AIR
SUPPORT NET

NATIONAL AND

THEATER ELEMENTS

MAXIDOS

I/O BW
CONTROL

CELL 3

CELL 1

CELL 2

MAXIDOS

MAXIDOS

MINIDOS

COS

COS

TASK
GROUP
1

TASK
GROUP
2

TASK
GROUP
3

USERS

REMOTE ELEMENTS
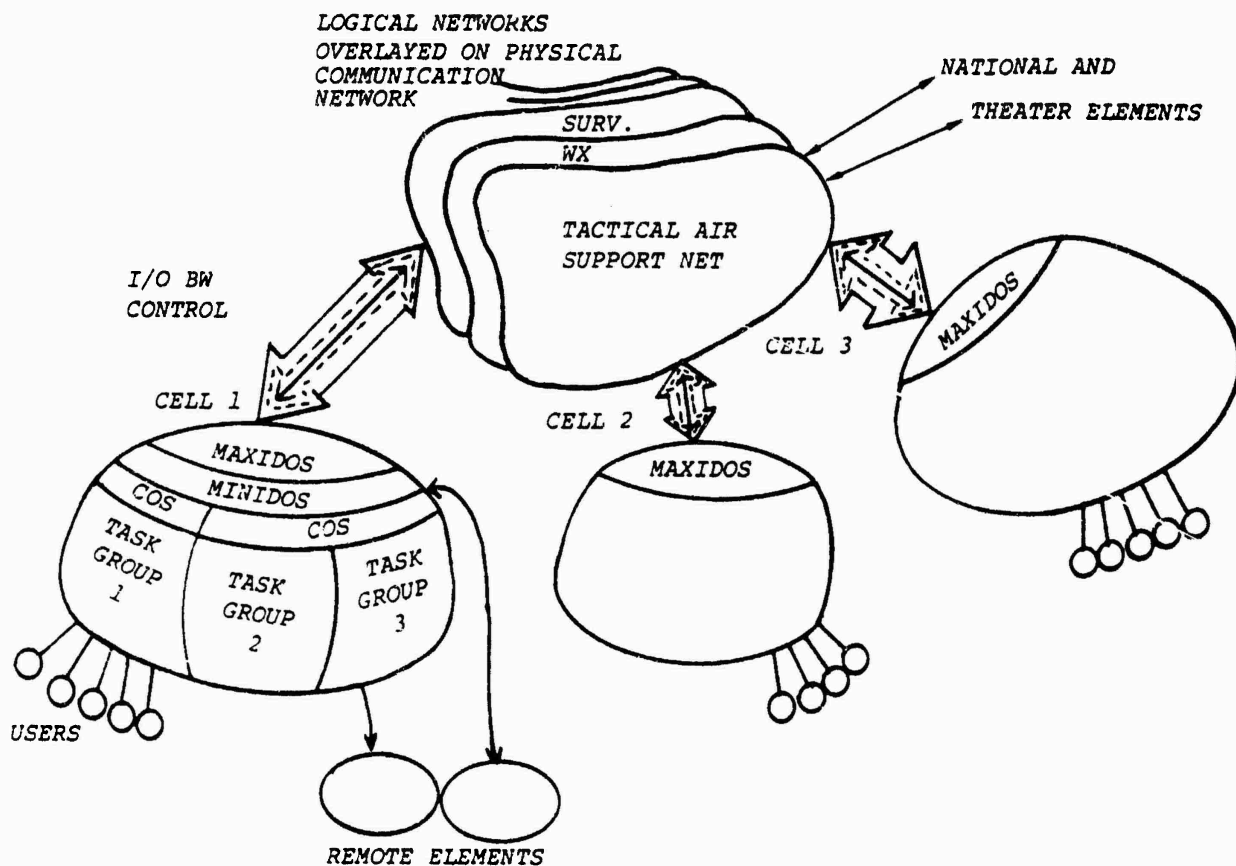
Figure 2. Logical View of the TAFIIS Architecture

3-5

At the MAXIDOS level directories of users or functional task groups need only be maintained at a logical level since physical mapping can be performed by the MINIDOS. For this reason, the entirety of users, hardware and software associated with the mininet is designated as a single logical "cell" to the MAXIDOS. Common resources controlled by the MAXIDOS are maxinet communication bandwidth and global data.

The interrelationships of the three levels of operating system and functional software is illustrated in Figure 3. The most significant feature of the interactions between the different levels of operating system is the communication bandwidth available to perform inter-task communication. The overhead of the operating system increases from the COS to the MINIDOS and from the MINIDOS to the MAXIDOS level because of the increased complexity of protocols, directory services, and contention.

In order to minimize the overhead costs of the operating system, the strawman operating concept is based on keeping as much processing as possible at the lower levels of the operating system hierarchy.

3.1.3 *Data Structure for the Strawman.*

Global data are defined to be any types of information that are of interest to more than one cell and are not specifically directed from one physical task group to another.

This broad definition includes message traffic, data bases, directories, files, status information, performance data, and system configuration data as candidates for inclusion in global data.

Since this broad definition of global data leads to a very large and complex directory of global data and access controls, the strawman software architecture attempts to simplify the

management of global data by defining arbitrary logical subnetworks of users associated with specific global data types. The maxinet becomes a composition of many types of logical subnetworks overlaid on a single physical structure.

The control of data access within the MAXINET has the added dimension of multi-level security since some cells may operate at higher security classification levels.

External elements and intelligence elements within deployed TAF will have the capability of handling SI/SAO compartmented information through the maxinet. Access to this data must be rigidly controlled and all dissemination must have accountability.

Specific types of logical subnetworks which may occur within the maxinet structure are:

[1] Weather

[2] Personnel

[3] Logistics

[4] Operations Plan/Frag order

[5] Enemy situation

[6] Friendly situation

[7] Surveillance and Identification

[8] Warnings and alerts

[9] Radiation management

[10] Scramble orders

[11] Order of battle

[12] Mission status

[13] Tactical Air Support Requests

[14] SI/SAO messages

[15] In-flight report net

[16] Intelligence Collection Plan

[17] Fire Support Coordination

[18] Software trouble reports

[19] Management information
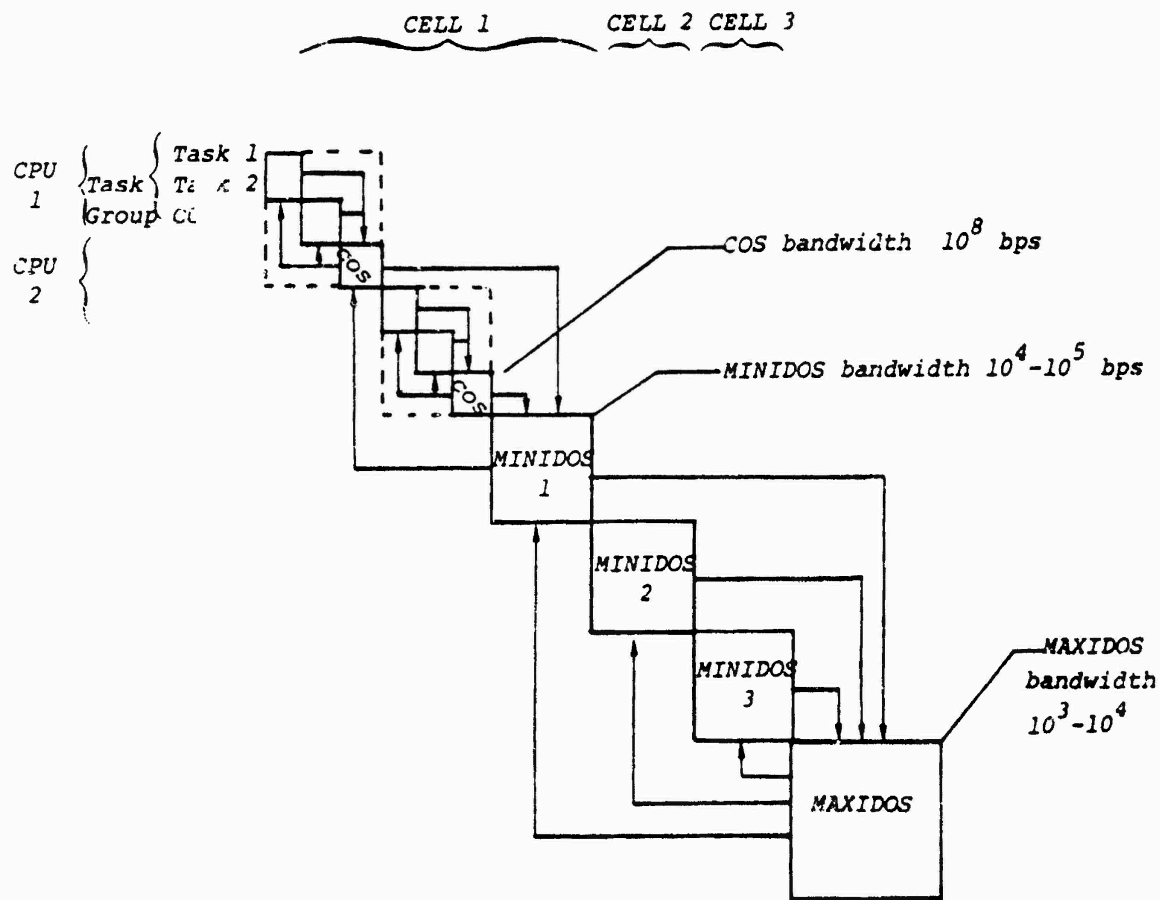
Figure 3. Interrelationships in the Three-level Operating System

Each subnetwork has its characteristics in terms of participants and how those participants interact in terms of data updates, data value addition, data usage, data exchange, and operational backup. The use of global data within these subnetworks by participants has wide variation in terms of timeliness, area of interest, and level of detail. For this reason, it is assumed that the control over use of global data is done by each individual user. A predetermined policy must exist between users in the subnetwork as to what data can be updated by which users and which users can access which data.

3.1.4 Security Impact on Architecture.

From the perspective of the data processing architecture, a cell will either have SI/SAO access or it will not. This decision was made because of the encryption problems for handling of SI/SAO data in the same physical environment which supports users without access to this classification level. The maxinet will therefore restrict the dissemination of SI/SAO data to only those cells which exclusively have user populations with appropriate clearances.

Within both types of cells there is still a need to distinguish classification levels of data for handling and retransmission.

In those cases where an SI/SAO cell were colocated with a collateral level cell, the two cells could share the same communication node of the maxinet.

3.2 The Mininet

The mininet logical functions are aimed at satisfying the following objectives in the strawman architecture:

[1] Providing easily extensible system capacity by making it possible to cluster multiple processors into a cooperative operating configuration. This capacity flexibility must make it possible to extend the number of user

terminals and addition of functional capability.

[2] Providing device transparency to common resources such as storage devices, communications, and input/output devices.

[3] Providing reallocation of resources for degraded mode operation. Available resources reallocated may be in standby configuration or in use for lower priority functions.

[4] Providing a multi-tasking operation to support a broad set of user support processing.

[5] Providing a software structure which facilitates incorporation of new processing capabilities, some of which may involve the use of microprocessors or other hardware elements to perform front end, back end, or special computation.

[6] Providing uniformity in intertask communications to facilitate evolution of modules in processing strings and with the ability to support old and new versions of tasks during cutover periods.

The mininet concept emphasizes the ability to provide flexible capacity for many levels of employment and numbers of users. This concept is dependent on wideband bus technology such as addressed in the Flexible Intraconnect program. If the mininet incurs a failure at the pipeline into the maxinet it will be isolated electrically but can continue to operate on the basis of voice communications and locally available data. If the mininet incurs a failure in the bus connecting the various processors, the entire data processing facilities of the mininet may become inoperable. It is assumed that each element within the TAFIIS operating domain will have some form of totally manual backup.

3.3 The Maxinet

The maxinet concept attempts to address the issue of critical information flow in the tactical environment given the problems of communication availability, limited capacity, and potentially long throughput times. The system goals of mobility and survivability can be addressed through analysis of the operability features of the maxinet.

3.4 Cell Architecture

In the strawman architecture the unit recognized by the MAXIDOS is the "cell". The cell physically corresponds to all processing components which are interconnected by a single local bus system. In general the cell would also be associated with a major communication node of the maxinet. Multiple cells can exist at a single communication node if the cells use physically separate bus systems but share a common communication center for maxinet communications.

Were it not for the problem of physical isolation between SI/SAO and collateral activities, the distinction between maxinet node and cell would not be necessary. This type of configuration is likely to occur at the AFCH where all-source fusion centers will be colocated with planning and operations elements and also at the TAB where IMINT or SIGINT production facilities are likely to be based.

Individual cells will have particular problems in interfaces, interoperability, user interface, and capacity. Factors which will cause uniqueness in the configuration of cells include:

[1] DASC operability with Army elements

[2] CRC/CRP operability with the Airborne Warning and Control System (AWACS), Naval Tactical Data System (NTDS), and other surveillance and identification elements (including allied elements)

[3] TACC capacity for a large and diverse user population

[4] Intelligence element secure interfaces and rapid communication

[5] Logistic element integrated data base requirements.

Common functional characteristics of cells include:

[1] Access to intelligence products

[2] Access to operations plan

[3] Access to friendly situation data

[4] Message handling

[5] Local data base support

Cell architecture must deal with the problems of hardware component multiplicity and diversity, loading variation, special external interfaces, interelement operability, user interface specificity, and reliability. Across specific cell configurations must be a measure of control on software modules governed by configuration management systems and user interfaces governed by top-level interoperability requirements.

From the perspective of the data processing system, cells are heterogeneous processing elements of TAFIIS because the configurations will be dissimilar in hardware components, function, and data. A major dilemma in formulating the cell architecture is to maintain the responsiveness of data processing to the individual TAF element while achieving overall goals of interoperability and maintainability. Interoperability is closely related to the the data structures and protocols defined for the maxinet.

The architecture of the maxinet and the mininet presume that programs and data will have a generic form for transmission through or mobility within the maxinet. Within the mininet, it is assumed that

the programs or data structures can be adapted for local performance needs.

Maintainability is assumed to be a combination of:

- familiarity of users with the system
- availability of expertise to solve operating problems
- availability of backup operating modes (also familiar to the users).

3.5 Design Issues

The strawman architecture does not try to present a complete solution to every issue. In fact this is not possible considering that the requirements are not firm nor have the cost-performance tradeoffs been established. The following design issues have been addressed in this study and are briefly explored in the following paragraphs:

[1] How to design a user interface which meets the specific needs of a duty position and also meets interoperability criteria

[2] How to determine the criticality of data in order to provide sufficient data protection and responsive data access

[3] How to provide for the gradual evolution of TAFIIS requirements and capabilities

[4] How to operate the data processing system such that it can be done reliably with field personnel

[5] How to provide the control of information flow such that it is adaptive to the changing tactical network

[6] How to build an information structure for TAFIIS information that can be used from the many different perspectives of TAFIIS users and persevere through an

extended system life cycle

[7] How to configure a survivable system that depends heavily on an integrated data base and is subject to intermittent communications between elements and element outage

[8] How to provide reliable data processing operation in an environment with frequent hardware failures, hostile enemy action, unreliable communication, and limited numbers of highly skilled data processing personnel.

Interoperability is not stated as a separate design issue because it cannot be independently addressed.

The following paragraphs address the open issues and possible tradeoffs that must be considered.

3.5.1 How to design a user interface.

The user interface is an issue in both the data processing architecture formulation and in the design of the operating system. The design of the user interface will impact resource allocation, survivability, operating costs, and system performance. It has been frequently recognized that the human factors guidelines for developing the man-machine interface have been outpaced by user terminal technology, processing capacity, and computational linguistics technology.

Too frequently, the performance of an otherwise excellent data processing system is degraded by flaws in the user interface. Typical problems include:

[1] The system does not recognize different user types.

[2] The system does not recognize different user skill levels.

[3] Multiple concurrent user activities cannot be handled by the system even though the user multiplexes his time.

[4] The system does not handle multiple vocabularies for commands and arguments.

[5] The system does not provide adequate performance (capacity, response time. availability, data protection).

[6] Commands cannot be optimized on the basis of frequency of use.

[7] Commands cannot be tailored for context in which they are used.

A concept which OSI has been developing over the last two years in connection with complex user interface designs for intelligence applications is based on a structure illustrated in Figure 4. This concept is based on a dual form for user commands to the system and messages from the system to the user. The two forms are the internal substrate against which standard syntax rules are applied and the external form which may have different manifestations determined by the user.

As shown in Figure 4, the external version of a user command may be a function key, menu selection, form fill-in, light-pen or cursor selection, or keyboarded alphanumeric string. All of these forms would map to a singular and explicit internal command substrate. The substrate command form would carry all necessary information for function invocation, audit trail, access authorization, and user identification.

Similarly, the system responses to the user would have a singular internal form and a variable external form. The variable external form could allow adjustment for variations in highlighting or alert mechanisms between different devices, preferences of individual users in display format, and ranking of output by significance to the user.

This concept has additional development cost in terms of the translator and tailoring mechanisms. Additional processing capacity is required to support the intermediate conversion from internal to external form. Storage capacity and file services are required to maintain the conversion tables and display formats used in the mapping process.

The advantages of this concept are numerous:

[1] The functional software component designer is isolated from the day to day variance and preferences of individual users during the development stages.

[2] Software maintenance is easier on functional modules because dependency on the external form is removed from the functional working of the software.

[3] It is easier to achieve terminal device transparency because applications modules are not directly interfaced to the terminal device.

[4] Functional testing can be performed with simple TTY type terminal devices using the keyboardable form of the command substrate.

[5] The command repertoire is open-ended because the internal command form is totally explicit and unambiguous.

The operating system is intimately involved in this user interface architecture because it is responsible for the internal routing of commands and responses to commands. This is of particular importance if remote users are involved or if access to a remote data base is required. Within this architecture there are questions regarding the scheduler and whether this function should be an operating system, system software, or application function.
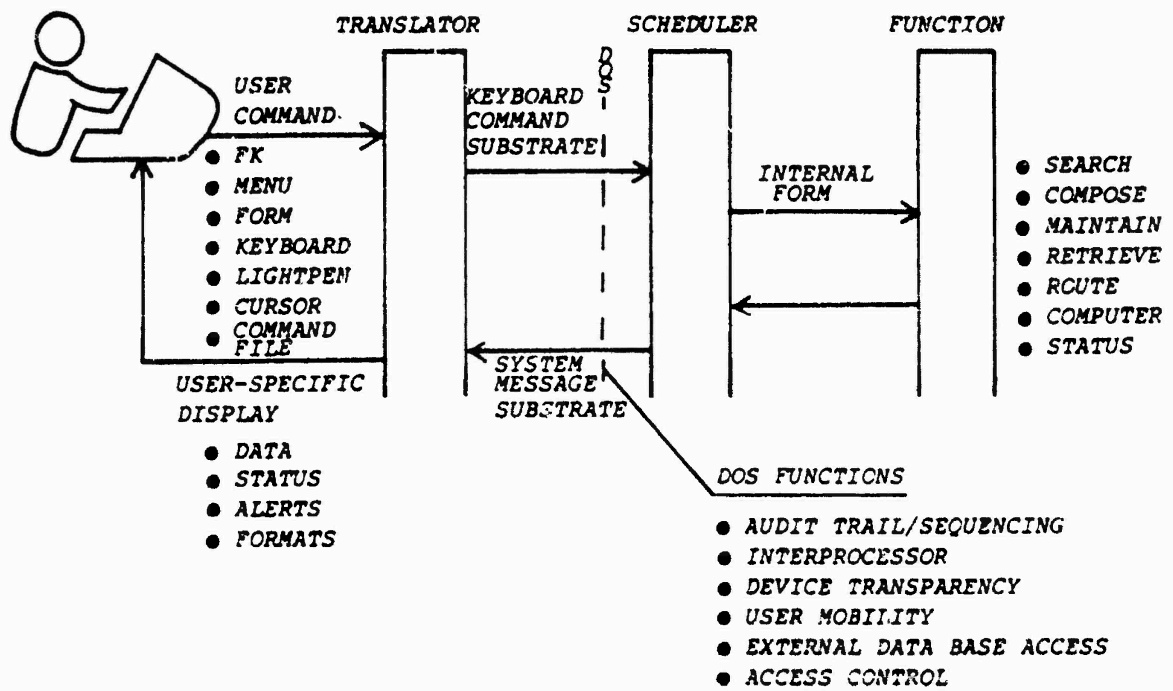
Figure 4. User Interface Architecture

3.5.2 How to determine the criticality of data.

The requirements for data protection and transmission speed are related to the criticality of a data item. In order for data criticality to be useful in allocation of resources or in prioritization of data transmissions there must be a consistent, system-wide definition of data criticality.

In a distributed data base environment with many users, the definition of data criticality is more difficult and must be interpreted in terms of each specific user. The general factors affecting data criticality definition from the perspective of one specific user are:

[1] ability to operate without the data

[2] time to recover the data from a secondary source

[3] reliability of the secondary data source

Operability without a new data item is complicated if the user has a previous copy of the data base to which the data item refers. Criticality will vary depending on how different the new data is relevant to what the user already has in the existing local data base. For instance, an intelligence report which updates the position of an enemy unit is not as critical as the first report which identified its existence. Also, this report is more significant to the user who is responsible for covering the area in which the unit is reported than to another user on the net who does not have responsibility.

Another complication is that not all users of information have the same response time requirements in the use of data. For example, an operations planner has much more flexibility in his use of situation data than does the targeting coordinator.

In using data criticality to allocate data protection resources and communication bandwidth, it is apparent that the definition must be from the receiving end. The sending element may be able to assign an intrinsic data criticality measure to the information content for its initial transmission but not for its total distribution. For this reason, data criticality would not be particularly useful in control of information dissemination through the maxinet. Each individual user might have to see the information before criticality can be determined. Full dissemination of every piece of data to every element would be intuitively inefficient considering the limited long-haul communication resources of TAFIIS. A pragmatic approach has been selected which is more along the lines of human information handling rather than classical data base management. A singular solution to data base queries does not always exist as is presumed to be the case in most data base management schemes. In the stark reality of the tactical environment, the user must deal with incomplete information, loss of communications, and multiple data values for the answer he is seeking. The natural human instinct is to supply missing information with hypothesis, evaluate data values on the basis of past performance of the source, and periodically try to reestablish communications with reliable information sources.

The important question being addressed by this approach is whether protocols and information dissemination algorithms can be implemented which can adapt the transfer of information between data processing elements in much the same manner as a human would adapt. This question is discussed in more detail along with the general discussion of decentralized control techniques in Section 4.3.

3-13

3.5.3 How to provide for the evolution of TAFIIS.

The principal questions that arise in meeting the TAFIIS Master Plan goal of an evolving modular system configuration include:

[1] How to use the current inventory of equipment

[2] Evolution of functional modules on different schedules

[3] Communication service evolution

[4] Operating system service evolution

[5] User population changes (function, skill, number)

[6] Hardware evolution

[7] Requirements change

The TAFIIS data processing concept presented thus far places a heavy emphasis on the operating system and information structures as the common denominators to maintain the system integrity through the evolutionary process. The data processing inventory will have two major and counteracting forces which will determine its composition. The force for change is the constant improvement in hardware and software technology which if introduced into the tactical environment could lead to greater operational capability. New technology must be introduced to counter an ever changing enemy threat. Counteracting these forces for change is the inertia created by investment in existing equipment inventories and training of personnel to operate those equipments.

Inter-service standardization of equipment would add to inventory stability except that it has never been achieved to any great degree except in voice communications modes. Attempts to standardize information structure has evidenced the following:

[1] Communications protocols and message formats have been standardized for common user communications services such as AUTODIN, INDICOM, OPSCOMM, NWMCCS, and IDHSC networks.

[2] Only external formats have been standardized (neither data base internal formats nor message content formats have been successfully standardized).

[3] Content of data bases such as order of battle, targeting, or installation data bases has not been standardized.

[4] Rigid formatting rules such as seen in the JINTACCS message formatting standard seem doomed to failure because of the difficulty in use.

Various information forms are cross-interpretable between users and computer systems given that the context is explicit and linguistic rules can be applied -- most often with the aid of a human. The primary issue appears to be one whether to standardize format and context of information or utilize a self-defining information structure. This question is addressed in detail in Appendix A. The primary tradeoffs are the efficiency of processing gained with standardized data structures versus the operability benefits of catering to the skill levels, language, and requirements of specific users. The referenced study concludes that self-defining data structures would have longer term benefits in system operability and evolution but at the cost of larger data structures and additional processing steps. With the advent of microprocessor technology which can supply the needed computation power for more complex processing, the technology trend would seem to favor the self-defining data structure. A related cost factor is communication costs for transmitting the larger data structures.

3.5.4 How to make the system work in a field environment.

The ease of operation which is crucial to the usability of a data processing system in a field environment must be evaluated from the perspective of the system user. Like all services, the Air Force is faced with limited numbers of skilled personnel and restricted training resources for upgrading the skill levels of users. Data processing systems are historically complex to operate and require a cadre of highly skilled personnel to maintain the operation. By including those individuals who are responsible for configuring and maintaining deployable data processing systems in the user population the problem of complete system operability cannot be overlooked. A simplified operating concept could be interpreted to mean that the system used in the field will be operated the same as the system that is used for training, exercises and for contingency plan preparation. Operability impacts the reliability of the system and the adaptivity of the system to new situations.

The principle question in operability and in the goal of simplifying the operating procedures is to what extent the system can be integrated into predeployment usage such that operability is not a problem and performance of the combined data processing and human elements of the complete information system is adequate.

3.5.5 How to provide adaptive information flow control.

One of the operating constraints of the tactical environment that is the most difficult to predict or simulate is that of communication bandwidth. Bandwidth describes the maximum data rate that can be achieved in a single channel, the total transmission capacity as a function of time. Bandwidth descriptors can be used in predicting minimum response times for interactions which involve interprocessor communications and in determining total data transmission capacity for a multi-user communication service. Predictability of communications service in the tactical environment is extremely complex because of factors which affect bandwidth availability and queuing delays due to contention for shared resources. These factors include:

[1] loading cycles from planning and operations schedules

[2] equipment failure rates

[3] error rates

[4] media instability

[5] electronic warfare disruptions

[6] radiation control

[7] node attrition

[8] node outage during mobilization

If loading is not adjusted to the available bandwidth, then response time degradation will ensue, even for high precedence traffic. The adaptivity features necessary for the data processing configuration to survive in this type of communication service environment include:

[1] robust routing of data between elements to bypass individual link failures

[2] adaptive information flow rates to avoid saturation of degraded communication facilities

[3] bulk data transfer to recover long-term network outages

[4] device interface transparency for robust information generation and reception (including voice and physical media)

3.5.6 How to structure information for TAFIIS.

Many issues surround the selection of an information structuring approach for the messages and data bases that are

to be handled within TAFIIS. All of the following types of information structures will be handled by the operating system.

[1] record messages

[2] data bases

[3] user-to-user messages

[4] user interface language

[5] software

[6] system configuration tables

[7] performance data

[8] system messages

In the context of data handled by the distributed operating system, all types of data can be regarded as a message, i.e. a standalone unit of information.

3.5.7 How to configure a survivable information system.

A survivable information system concept must consider all sources of information including hardcopy, human memory, and computer readable data bases. The threats to the survivability of the information system include:

[1] loss of local hardcopy

[2] loss of communications with individuals who have information in their head

[3] failure of data processing which supports data base maintenance and usage.

These types of losses may be due to enemy action, electronic warfare against communications, equipment failure, or suspension of operation during mobilization. Data processing capability has an important impact on the survivability of all three types of information. Data processing can support the creation and maintenance of hardcopy information sources that can be used in the field without any communication or computer resource support. Order of Battle (OB) data, preplanned target lists, maps, weaponeering data, country ethnographic data, terrain analysis, operating doctrine, and other types of technical data are frequently used in hardcopy form because they are fairly static over short time periods. Intelligence analysts who have access to computer system support almost always have a hardcopy backup file of key reference material.

The human memory is depended on for all short term operational needs and decision making. Organization of operations or intelligence staff involved in rapid decision making roles (such as targeting) is normally done to take advantage of the memory resources of individuals in the group. Computers support the use of group information resources by providing electronic mail and digital conferencing capability, and by providing group displays.

Automated data bases have become the backbone information resource for strategic command and control and intelligence operations by providing means for handling enormous volumes of detailed data. The strategic level data processing services aid in the integration of requirements and production of specific data for the area of interest of specific users and indexes or summaries which simplify the use of the voluminous data base. These systems have availability problems even in the rather benign environment of strategic computer facilities.

In the tactical environment the use of computer resources to support automated data base functions is hampered by the reliability problems of communications, storage, and output devices required to support the functioning of the system. Very limited success has been achieved to date because of high cost, limited performance, long set up times and limited mobility. Miniaturization of computer components and high density storage devices are gradually reducing these

limitations.

Cost is still an extremely important factor because of the complexity of militarizing data processing equipment and integrating it into the operational environment.

Issues which must be addressed in achieving a survivable information system supported by data processing include:

[1] How to apportion the data processing budget between support to the human element in information handling, the preparation of backup hardcopy information, and providing redundancy in automated data base capability.

[2] How to balance diversity in operation against complexity in training for degraded mode operation.

[3] How to achieve self-sufficient operation of elements when isolated by communication failures without giving up the advantages of centralized planning and resource allocation.

These issues are complex because they are all long range issues and involve future cost tradeoffs and threat analysis. The TAFIIS system development approach must remain extremely flexible in the respect to achieving survivability and no options should be excluded. Emphasis must be placed on the human element in terms of being able to use backup information resources to achieve the desired survivability level. This issue is directly related to the overall issue of adaptivity to the specific tactical situation.

3.5.8 *How to provide reliable data processing operation.*

Data processing reliability has been established as an issue separate from that of providing information system survivability. Reliability encompasses a

variety of specific issues in regard to the data processing architecture:

[1] data protection from hardware failure

[2] data protection from contamination

[3] data protection from unauthorized access

[4] availability of functional capability

[5] availability of external interfaces

[6] availability of critical data

[7] operability in degraded mode

[8] recovery time from failures

User distrust of data processing in the military environment is most frequently traced to reliability problems. Simply because a computer system has failure logs which show a .99 availability figure does not mean that the system has adequate availability from the perspective of the user. From the users perspective, the system could be inadequate because of poor response time under peak loading, contention for shared devices, software errors which necessitate processing repetition, lack of data base currency, difficulty in executing functions, or other problems affecting operability.

Many of these problems are never adequately corrected because they are not regarded as system failures or there may be no way of altering the system configuration to alleviate the problem. In most cases, the problem lies in the system development process and the disregard for the use of performance monitoring and feedback in the operating system control of resources.

A second issue in system reliability is the basic reliability of the components. In hardware design, the reliability of components can be made only so good before the quality control process becomes prohibitively expensive. This

is even moreso true in software components where testing and debugging time is the most dominant factor in software reliability. However, time is the most expensive resource in the system development process and the TAFIIS system implementation plan calls for reduction in the time to deploy a new system configuration.

In hardware, the solution has been the development of fault tolerant systems which use various mechanisms to maintain operability under single or multiple device failures. At the system level, where processes involve hardware, software, and human elements the design of fault tolerant processes can be much more difficult.

Evidences of fault tolerant software devices include:

[1] redundant data storage

[2] checkpointing

[3] reasonableness tests

[4] exception case handling

One of the most critical areas of fault tolerant processing is in the cutover from an old configuration to a new configuration which will be a frequent occurrence in the modular evolution of the system or in system extension in field conditions. This will impact heavily on operating system requirements.

3.6 Toward an Open-ended System Configuration

Because of these open issues in the design of a suitable data processing system for TAFIIS there should be a much stronger emphasis on the development approach. The classical approach in which requirements are defined at the beginning of the project followed by a lengthy implementation period is not responsive to the changing world situation nor is exploitive of hardware and software state of the art.

An alternative is to view the development process as an on-going production process which constantly provides deployable data processing configurations for using elements. This perspective is more consistent with the open issues mentioned above.

The concept of continual requirements definition and system development is viable only if there is continuity in the view of the user and consistency in the physical and logical interconnect of components. In terms of the design approach proposed in this document the emphasis falls in the area of software development and system integration and validation. The issues presented above also point out the need for new types of technology in the system development process which are discussed in Section 5.0.

4. RATIONALE FOR THE STRAWMAN ARCHITECTURE

This section presents arguments supporting the design choices made in the definition of the strawman architecture for the tactical C^3 distributed operating system.

The diverse arguments and considerations are presented in separate subsections corresponding to different specific aspects of the scheduling, control and protection functions performed by an operating system. Subsection 4.1, however, deals with the appropriateness of the overall architectural concept as a solution to the tactical command and control problem.

The common format followed in each one of the subsections has been chosen to emphasize the feasibility and adequacy of the proposed choices and their preferability over alternative solutions. Open architectural issues are discussed and the nature of the information required for their resolution is identified, whenever applicable. The final part of each subsection briefly reviews the technological methods and tools required to fully develop and implement the Strawman architecture, indicating requirements for technological advance and its associated risks.

The concluding section addresses the problem of the development approach for the strawman architecture and its associated technology requirements and risks. The combination of the architecture risks plus the development approach risks are the basis for the technological risk evaluation presented in Section 5.

4.1 Overall Architecture

4.1.1 Rationale for the Selected Configuration.

The primary arguments supporting the choice of an architecture are based on:

- a number of "cells" or "clusters" of relatively tightly coupled processors interconnected

- by a low bandwidth communications network.

Geographical proximity and communications availability are the primary considerations for this clustered design approach. The nature of tactical information handling problems indicates that availability of a geographically distributed data processing and exchange capability could substantially enhance the ability to properly manage the resources of the tactical air force.

The geographical distribution of any automated system implemented to support tactical C^3 functions is necessary as the nature of the missions and tasks to be performed preclude the colocation of all information users and producers in the vicinity of a centralized data processing and storage capability.

The interconnection of the geographically dispersed computing sites is required in order to provide access to data, that due to the particular characteristics of the information collection and flow processes in a tactical environment, is not available at all locations. Further, the need to assure survivability of critical information elements poses additional requirements for the interconnection of processing sites into a distributed computing facility.

The above arguments are very general in nature and, while clearly supporting the need for an interconnected distributed capability, do not specifically provide a rational foundation for the chosen architecture. In order to provide that rationale, specifically calling for tight coupling at the local level and loose cooperation at the global level, it is necessary to examine the nature of tactical information processing in further detail.

Tactical information can be generally characterized as being composed of of two primary elements:

- An "air situation model" describing the state of affairs in the real world by means of data values and relations of interest to the tactical commander

- Messages, conveying information, that require processing (both by humans and computing equipment) in order to determine possible modifications to the air situation model.

Users of this information, however, exhibit wide differences as to the general scope and level of detail of the specific data elements to which they require access. Analysis of the TAFIIS Master Plan documentation [Reference 1] reveals, however, that certain groups of users can be identified as sharing a number of common characteristics such as:

- their common interest in information relevant to the performance of specific tasks in a certain geographical area (e.g. close air support)

- their physical geographical proximity

- their common data processing needs

- their common need to access or produce information at a uniform level of detail (e.g.: aggregated information pertaining to a more or less broad area of tactical resource managements versus specific control of individual assets)

- their need to be aware about the existence and capabilities of other producers and consumers in their group

- their common perception of different groups of users (i.e., different tactical organizations) as a common functional unit concerned with the management of related resources and thus required to maintain adequate information on their status.

The choice for a tightly coupled architecture in support of each of these user groups is therefore a logical consequence of their common needs and requirements for the sharing of processing and information resources. Their physical proximity allows the use of existing or proposed technology (i.e., bus technology, the flexible intraconnect) in order to provide the large bandwidth required by the efficient centralized allocation of computational resources in a multiprocessor tightly coupled environment. Stringent control, characterized by availability of up-to-date, consistent (across cell processing elements) system tables is required to allow graceful resource reconfiguration in response to changing environmental conditions (particularly to variations in mission scope resulting in changes to workload or functional priorities).

On the other hand, information processing needs between groups of users are limited essentially to exchange of data elements as required by existing policy and changing mission needs. Therefore, at the global level primary needs are in the area of global data base management characterized by:

- data retrieval of information elements between cells

- update of other cell elements, particularly for backup purposes

- data base segment recovery.

In order to perform these global data management functions, the system must rely on the usage of electromagnetic channels which must be shared with other tactical functions and which are susceptible to interference or failure. Under this environmental circumstances, continuous availability of adequate communications can not be assured and the resulting uncertainty, at each cell, about the true status of the overall computing resources necessarily indicates a design where each cell has a large degree of functional autonomy. In

4-2

other words, intercell functional dependence will require assumptions about their ability to carry effective communication which, on the basis of a realistic analysis of the state of the art and expected advances in digital network technology, cannot be guaranteed. Therefore, the global set of cells forms a "cooperative federation" of processing elements [Reference 3].

It is important to remark however, that the amount of collaborative behavior and dependency between cells is larger than that usually associated with experimental networks such as the ARPANET. For example, cells are required (as further detailed below) to provide services to other cells at all levels of priority rather than being primary processors of local information with secondary requirements for global request support. The requirement for autonomy stems from the need to assure their viability and usefulness at a level commensurate with availability of processing resources, rather than from the desire to resolve global conflicts by assignment of functional priority to local needs.

4.1.2 Evaluation of Alternative Solutions.

As stated in the previous discussion, global solutions that emphasize continued communication availability and extensive intercell coordination cannot be guaranteed to operate under the environmental constraints of the tactical environment.

In addition to these basic communication oriented considerations, other arguments also indicate that at the global level a loose federation is desirable over more centralized forms of control.

First, there are basic differences between cells both in terms of functional processing needs as well as scope that indicate that tight interconnections of dissimilar cells will not lead in an straightforward manner to

advantages such as operational performance gains due to processing task distribution and sharing over the larger processing system. Any such advantages will be offset by disadvantages related to the management and control of a more complex resource and performance losses related to inability to exploit functional specificity. Second, any system which implements an information flow/processing structure which is substantially different from that of the basic command and control structure which it supports is bound to induce user reluctance to depart from the use of proven procedures to go into methods that disagree with his long established perception of organizational relations.

At the local level, alternative solutions that tend to decentralize resource management by increasing the autonomy of each processing element seem, at this stage, to be less desirable due to increased difficulties in the coordination of processes such as system reconfiguration and recovery. It is conceivable, however, that a further level of control hierarchy (at the local level) may be required to exploit lower level functional commonalities, either at the tactical function level (e.g., a processing element devoted to operations planning) or at the ADP functional level (e.g., a data base computer). Security considerations discussed below seem to support this requirement. At this stage, further discussion of local (mininet) design must be postponed until specific processing requirements for each cell are known.

4.1.3 Open Architecture Issues.

By far the most extensive need for additional detail lies on the identification of specific characteristics for each cell architecture. At this stage limitations on the scope of this work as well as lack of detailed requirement availability preclude further development.

At a global level, the ability of a processing/storage entity to become a "cell" and be recognized by other cells must be further defined. The related problem of cell coalition (i.e., merging of several cells) deserves similar attention. Since at the global level, the extent of services that each cell receives from others is dependent not only on resource availability but on global policy defining missions, privileges and priorities, it is clear that any specific development addressing the issue of cell identity and behavior must be firmly rooted in broad organizational principles. Further studies on information flow dynamics in a changing tactical environment, stressing in particular the extent of the information required, (i.e., not only its type but also the instances required) are needed in order to provide specific answers to these questions.

4.1.4 Technological Needs.

Specific developments in operating system technology required to assure successful deployment of a tactical C^3 system are addressed in detail below in the context of specific issue problem/discussion, at the overall and general architectural level this subsection's discussion comments will be confined to the identification of technological needs to produce the information required to specify a candidate architecture to higher levels of detail.

Three major technological areas require further advances in order to enhance the understanding of the complex requirements of a tactical system and to assist in the design of architectures to meet those requirements:

- *Automated data base requirement analysis and design tools:* Particularly those that emphasize both types and extent of information as perceived by the user. In the terminology of the ANSI/SPARC [Reference 4] model.

Emphasis is required in the development of languages for the specification of data semantics emphasizing relations between production, process and use of information at the "external" level of specification. Tools are required also to analyze those requirements so as to assist in the determination of desirable architectural choices [References 5,6].

- *Automated system analysis and design tools:* Enhanced tools are required to specify and analyze the performance of concurrent and simultaneous processes and the effect of their concurrent actions on information [References 7,8].

- *Distributed system simulation/emulation tools:* Enhancements are required for tools allowing fast and flexible specification and simulation/emulation of distributed architectures. Standard scenarios must be developed to be used in conjunction with specific architectural evaluations.

Due to the extent and complexity of tactical data handling and the need to understand the details of such transactions in order to further specify architectural characteristics, development of automated advanced data base and information processing requirement specification tools must be considered to have higher priority over the development of architecture evaluation methodology.

4.2 Interprocessor Communication

In the Strawman architecture presented in this report, the design has been performed primarily at a logical level of abstraction dealing with each cell as a single entity capable of communicating with other cells via a communications network (maxinet). It has been assumed that the underlying physical network will rely primarily on electromagnetic links, which must be shared with other tactical systems and which

is subject to failure, damage, interference or destruction and thus cannot be assured to be continuously available.

Average bandwidth availability of the maxinet has been assumed to be in order of 10-50 Kb/sec. No specific assumptions have been made or design choices performed related to the nature of the interconnection pattern. Any specification of actual communication system characteristics at the physical level must necessarily follow studies to precise the detailed nature of each cell and its information requirements.

At the logical level, however, the Strawman architecture specifies that each cell will contain interfaces (either concentrated in a communications front end processor, or distributed throughout the cell processing elements which present an homogeneous interface to the maxinet. As further precised below in the discussions of resource sharing and data base management, each cell interface keeps a map of the status of other network resources as measured by its ability to communicate with other cells.

The choice for a common interface (i.e., a common data/message exchange language) between cells is based on the expected complexity of the actual time-varying configuration of each cell. At any moment in time, the actual configuration of each cell reflects its functional requirements as well as the need to process dynamically changing amounts of information. In order for other cells to be free from the resource consuming task (related primarily to the need to exchange large amounts of cell status information), it is desirable that communication between cells be performed using a common language specifically designed towards support of distributed data management functions such as:

- data segment retrievals

- data base element creation/modification/update

- data base element availability verification

In order to further free cells from the need to ascertain the nature of languages supported by processing elements in other cells, their physical data representations and other low-level configurational details, the intercell communication language is based on high-level, "external" (in the ANSI/SPARC terminology) views of tactical information using a nomenclature previously defined and which is common across all cells. In this way, the actual conversion between cell languages and data representations as well as the actual determination of the most appropriate intercell procedure to be used in response to an intercell information exchange is determined by each cell. Therefore, cells do not need to consider the internal operation of other cells (beyond their perception of data availability expressed at a high level of data description) in order to process intercell requests.

It is important to rework that information detailing a cell's ability to support data exchange requests (e.g. availability of certain data base segments) is considered here to constitute part of the data that interprets the data base. Therefore such information can be requested and transmitted using the postulated data exchange language and protocols without the need to further complicate the intercell communication.

The concept of interfacing cells through an "intelligent gateway" (i.e., the MAXI-DOS functions residing at each cell) is also supported by the need to provide interoperability between the TAC C[3] system and other related information systems (e.g., WWMCCS). By adapting the MAXIDOS interface philosophy and by implementation of is common tactical-data oriented language, each

of these systems becomes in effect another cell capable of joining the maxinet in order to provide/receive services from other tactical processing cells.

The nature of interprocessor communication between the processing elements that compose a cell are still undefined at this stage, as the actual details are highly dependent on the internal physical and functional organization of each cell. It seems clear, however, that a major portion of intercell DOS (the MINIDOS) resources will have to be directed towards reformatting and translation tasks required by the need to interface heterogeneous processing, man/machine interface and storage elements. Again, the actual approach to be followed is highly dependent on the eventual structure of each cell and our organizational requirements to provide implementations that incorporate existing computational resources.

4.2.1 Evaluation of Alternative Solutions.

Approaches to the interface of multiple processors which rely on the use of specialized communication/translation network node(s) must contend with the requirement to assure continuous high bandwidth communication to these network interface computers. The nature of the tactical environment clearly indicates that that requirement will not be generally met. Further, any such approach will increase intercell dependence (in this case between cells and the network interface computers) while placing high level requirements on the network interface computer survivability.

Any scheme which relaxes the requirement for a data oriented, high-level homogeneous interface between cells increases intercell dependence and unnecessarily complicates the solution of technological problems associated

with interoperability, being therefore less desirable than the proposed architectural approach.

Conversely, interprocessor communication languages that rely on data descriptions at lower levels of abstraction (e.g., logical or physical levels) require additional cell knowledge about the internal organization of other cells being less desirable also from the point of view of human understandability of intercell messages.

An additional advantage of the use of high level, user-oriented interfaces between "cells" is the capability of direct interface of humans into the maxinet (through simple non-processing communication terminals possibly having encode/decode capabilities) to whom intercell messages may be selectively disseminated.

At the MINIDOS level, approaches that emphasize language and data description commonality seem less desirable, however, as imposition of homogeneous protocols reduces the ability to exploit functional specificity on cell design. Further, the practical requirement to utilize available computational resources may be incompatible (due to the wide variety of approaches being utilized at present) with any design based on standardized, TAC-wide, intercommunication protocols.

4.2.2 Open architectural issues.

The nature of intercell communication, at the MAXIDOS level, has been specified in the design at a high level of abstraction. While specific priorities have not been proposed, the nature and scope of the exchanges between MAXIDOS at different cells has been specified.

In order to perform the physical actions resulting in the transmission of requests and data between cells, the cells must rely on the capabilities of a physical communication network that implements

the logical intercell network (i.e. the maxinet). An analysis of available and evolving technology for the physical interface of geographically dispersed computing sites indicates that actual interchange will be likely established by means of circuits established using digital electromagnetic links and automated switching nodes. The interface between cells (actually between their MAXIDOS functions) and the underlying physical network must be specified when more detailed designs are evolved.

Due to the nature of the decentralized approach proposed, which stresses dynamic resource allocation at the global level on the basis of observed environmental characteristics, it is important that future developments explore the details of the information exchange between cells and the processors (front ends, switching units, etc.) constituting the physical network. For example, the ability of the physical network to inform cells about dynamic changes in communication capacity will greatly increase the effectiveness of the resource allocation decisions performed by each cell, which, otherwise, will be hampered by lack of required information to ascertain the probability of success of allocation decisions (e.g., such as the probability of success of efficient file recovery using a remote processor).

Relationships between resource allocation at the cell level (i.e., the decision to request actions from another cell) and at the physical network level (i.e., the decision to transmit a certain priority message segment or packet) must be studied. Physical network implementations must be able to apportion resources so as to reflect both the priority needs established by cells (on their basis of their perception of environmental realities, function criticality and known policy) and their own perception of resource availability

(which may differ from those perceived by individual units). The exact nature of mininet interprocessor protocols and implementations has not been explored as the extent of required capabilities is clearly dependent on the nature and scope of the supported functions within each cell. Further studies that attempt to precise this exchange must deal separately with each particular possible type of cell.

4.2.3 Technological needs.

The overall design philosophy at the global level has been based on the idea of *implementing* a processing environment where full availability of computational and communication resources is not required to assure production of useful results. Unlike classical processing environments where successful completion of a computing function is dependent both on availability of valid inputs as well as that of a function-dependent, a priori defined amount of resources, the maxinet/ MAXIDOS design stresses use of algorithms that adapt their behavior in response to perceived (or estimated) changes in input validity and resource availability (i.e., adaptive resilient algorithms).

The approach on which the MAXIDOS design is based therefore stresses continuous observation and estimation of environmental processing parameters (typically translated into measures useful to aid MAXIDOS resource allocation process, such as "coherence measures" complemented by use of algorithms that are able to produce results of increasing usefulness in direct relation to availability of resources.

Although the proposed nature of the collaboration and decision making processes to be performed by cells closely resembles that used by human decision makers in various fields of endeavor (among which command and control is a particular relevant example), the approach has not been yet

4-7

automated in the form of a distributed operating system as the concept radically departs from established exact computation, inflexible behavior (i.e. successful process completion versus error), more adaptive processing that characterizes present day operating system functions. For this reason, a number of technological developments are required in order to produce the necessary language and algorithms that constitute the MAXIDOS.

Further, as the MAXIDOS is essentially a high level system function interfacing with lower level systems (such as cell processing elements or maxinet resources) it is necessary to define the nature of the interaction between the MAXIDOS and other systems in order to assure that such subsystems provide adequate assistance to MAXIDOS decisions-making functions (i.e. resource allocation) in the proper estimation of processing status and resource availability, as well as in the performance of processing activities in accordance with MAXIDOS decisions. In the specific area of interprocessor communication the following developments are required:

● *Development of data-based user languages for information retrieval/update on the basis of information value (criticality):*

Present data base access languages do not place adequate emphasis on the differentiation of data and information. This state of affairs results in systems that are not responsive to user information needs [References 9,10]. In order to fully provide MAXIDOS decision functions with the informational elements required to perform efficient allocation of resources it is necessary that users be provided with languages that clearly define the nature of the needed *information* (rather than data elements) and its value. The question of the amount of flexibility to be provided in the user

specification of other processing alternatives (e.g., it is very important that I know ......but if not possible I will settle for ...., etc.) versus automatic determination and evaluation of such alternatives by the system must also be studied.

● *Critical data handling protocols/algorithms:*

One of the main concepts in the proposed maxinet design is that of the use of algorithms having the property of adapting its behavior to resource availability producing, for example, partial results which are still valuable to the computational purposes sought by algorithm execution, even where all conditions for algorithm completion cannot be met. At the intercell communication level, algorithms that perform intercell communication using resource availability as a dynamic variable must be developed. These algorithms must behave so as to assure that both: (1) critical transfer functions are performed first or are allocated essential resources, and (2) intercell transactions result in transfer of valuable information between cells even if only partially completed.

For example, a file transfer procedure designed with these principles in mind should stress stepwise transfer of data elements and relations in the file in a sequence that reflects the relative importance of data elements to the purposes sought by the transmission.

● *MAXIDOS/maxinet interfaces:*

Clearly the efficiency of MAXIDOS functions as decision makers is highly dependent on the quality of the information about resource availability available to them. Equally important in assuring that performance of MAXIDOS decision making function leads to adequate resource utilization is the development of processing

4-8

techniques which perform their functions in accordance with assumptions about their behavior made by those allocating procedures. The relation between MAXIDOS functions and those maxinet processes directly interfacing with them must be further defined and studied.

In particular, emphasis should be placed in the definition of:

[1] *Dynamic bandwidth allocation algorithms:* Providing both information to the MAXIDOS about available communication capacity as well as allowing MAXIDOS request of additional communication resources.

[2] *Saturation prevention algorithms:* Detecting potential exhaustion of resources and providing information to MAXIDOS functions about the likelihood of such an event.

[3] *Maxinet Physical Transfer Protocols:* Studying the nature of the physical transfer procedures required to implement resource allocation decisions of each cell's MAXIDOS while resolving intercell conflicts on the basis of an "a priori" defined global policy.

[4] *Network error handling procedures:* Defining the relative responsibilities of maxinet and MAXIDOS functions in the detection and recovery of transmission errors. The relative importance of utilizing automated maxinet retries versus invocation of further MAXIDOS functions must be evaluated for a variety of cases.

Technological requirements for interprocessor communication for the MINIDOS/mininet are not so clearly defined as those for the MAXIDOS. The

cause for this can be traced, as before, to the multiplicity of possible intracell configurations supporting diverse tactical C^3 functions. Two areas, however, can be explored at this point in order to further specify and determine the total approach feasibility:

● *Heterogeneous processor interfacing:* In Section 4.1 the rationale for the assumption of nonhomogeneous processing element usage in a typical cell was presented. Interface procedures and communication protocols facilitating the flexible interconnection of heterogeneous devices must be developed. Those procedures must provide for:

[1] reformatting/translation of bit strings elements between heterogeneous processors

[2] reformatting/translation of OS level requests between heterogeneous processors, or, alternatively, translations of such requests to a common interprocessor language

● *MAXIDOS/MINIDOS Interface:* Closely related to the problem of interface of heterogeneous processors is that of defining the nature of interfaces between MAXIDOS AND MINIDOS functions. For a typical cell, relevant questions to be determined include:

[1] the resident locus (or loci) of MAXIDOS functions within the cell

[2] required reformatting/translation operations between languages or representations used in cell components and MAXIDOS requests/data.

4.3 Resource Management

4.3.1 Rationale for the Selected Architecture.

The proposed concept for resource management at the global (i.e. maxinet)

4-9

level can be briefly characterized as relying on probabilistic resource management under conditions of imperfect, incomplete and inconsistent state observations by a number of loosely coupled decision elements (i.e. the MAXIDOS functions located in each cell).

The choice of probabilistic resource management as the theoretical basis for the development of resource allocation algorithms at the global level has its basis in considerations related both to the state of the art as well as to operating environment conditions.

The nature of the environment where the tactical C^3 system is to operate precludes the development of control algorithms relying in the development of consensus among two or more cells before an actual resource allocation is performed. Since availability of and communication with other cells cannot be assured due to the unreliable characteristics of the communication and processing resources, the allocation procedures contained within each cell must reach an independent assessment of the state of the system in order to develop a rational basis to arrive at control decisions. It is important to remark that, even in the unreliable environment under consideration, cells are not precluded from requesting information from other cells in the maxinet thus obtaining a, hopefully, more accurate estimate of the system state. However, participation of other cells as correspondents in any such exchange agreement, is not postulated here as being necessary to assure continuous operation of the system.

Indeed, in the proposed approach, each cell has a number of choices regarding possible allocation decisions to be taken at any instant in time including that of obtaining additional information (either about system state or data base elements) from other cells. Such decision is taken, along the lines of

classical decision analysis, after consideration of the probability of success, expected value and risk associated with each available choice (including that of obtaining additional information or signaling possible intentions to other cells). Even when a cell's action was to result in an unfavorable outcome (e.g., requested information fails to be received), the cell is still able to continue operation after revising its previous estimates of the probability of success for that action.

While the resulting control scheme is highly decentralized, as there are multiple independent decision makers, it is also important to mention that there exists only one set of goals for system operation commonly observed by all cells, (i.e. the global policy defining a "common good"). Further, each cell controls only a partial set of the total amount of resources available to the global system. The *essential* difference with many centralized and decentralized control schemes, however, is the fact that the nature of the environment imposes a dynamical *nonclassical* information scheme to the system, in the sense that different decision makers have different information about the state of the system.

An evaluation of existing knowledge shows that development of the required algorithms can be reasonably expected to evolve from foundations provided by decentralized control theory [References 11,12], decision theory [References 13,14,15] and market signaling theory [References 16,17]

The same technological evaluation shows that in order to control algorithms, the notion of efficient control will have to be necessarily revised due to the basic analytical difficulties found even for example decentralized control systems [References 17,18] when efficiency is expressed by means of optimality measures as commonly done with classical control theory.

4-10

Subsection 4.3.4. below will further expand this point.

At the mininet or intracell level, the resource management approach is highly dependent on the nature of the mission performed by the cell and has been left, therefore, unspecified. On the basis of potential deployment of suitable high bandwidth communication resources (i.e. the flexible intraconnect), it can be hypothesized, however, that intracell resource management algorithms will be based on control principles requiring high coordination between the allocation procedures resident at multiple processing elements.

Therefore, it is expected that intracell processing will feature centralized control procedures with a resultant ability to assure data consistency and high level adaptive performance of computational tasks. Singularities and exceptions to this centralized control rule can also be expected in certain cells due to the need to physically segregate a component of the cell resources without the services of a high bandwidth intracell channel (e.g. remote radars in an integrated surveillance network controlled by a CRC/CRP). In these cases, the segregated portion should not be considered to constitute a separate cell as its interface requirements are confined to exchanges with the original cell without need for full exchange capabilities with all other cells federated into the maxinet.

4.3.2 Evaluation of alternative solutions.

Solutions that require assumptions of availability of reliable, fast communication resources between cells as well as cell survivability are not consistent with the nature of the environment in which the tactical C^3 system is expected to operate and therefore have not been considered here.

It is of interest, however, to consider the eventual possibility of instances where cells become colocated thus allowing fast, reliable information exchange between them. In this case, a local centralized controlling scheme encompassing both cells can likely be postulated as a desirable alternative to enhance, at least to a limited extent, the computational properties of the system. Several reasons, however, indicate that this is not a desirable capability:

- The nature of cell colocation will most likely be temporarily limited. Therefore, implementation of centralized algorithms wherever colocation is possible, will entail frequent content switches between radically different resource management philosophies.

- Even when colocated, the requirements for all exchange will still be limited to data base exchanges that will likely be reasonably controlled by the proposed decentralized probabilistic management algorithms.

- When colocated, the higher reliability between cell exchange will increase the probabilities of success of information exchange will reduce the associated costs, thus favoring the choice of allocation alternatives (by all colocated cells) that promote cell survivability and data quality.

Resource management techniques proposed for the mininet have only been tentatively identified to a very high description level. Not many alternatives can therefore be precised to the extent of allowing rational comparisons between architectural characteristics.

The question of possible cell partition into portions interconnected through low bandwidth lines without new cell formation was discussed at the end of the previous subsection.

The basis for the rejection of the concept of interfacing the partitioned

segments (as new cells) through the maxinet was the lack of specific requirements for the segregated portion (e.g. new CRP or FACP) to communicate with cells other than that from which the new segment was separated.

4.3.3 Open architectural issues.

The proposed probabilistic resource management scheme (global or maxinet level) relies on use of independent decision makers that produce their resource allocation choices on the basis of information that is not consistent across decision making elements.

The principles of development of these procedures, their mode of operation and their general characteristics and properties remain matters requiring further study. This point is further discussed in the next subsection when precising relevant technological requirements.

Clearly the proposed management schemes have been precised only at the conceptional level dealing basically with the overall management philosophy to be used by the system. As further levels of description are considered, the nature of the supporting functions required to implement both centralized and decentralized control procedures must be specified.

4.3.4 Technological requirements.

Major technological development needs exist in the area of decentralized control of distributed systems under environmental conditions such as those present in tactical environments. Of particular importance are those developments that either provide the foundations for the development or pursue the development of algorithms that allow collaborative resource management using inconsistent or partially incorrect information. Specifically, technological developments are required in two broad areas:

- *Probabilistic management of processing resources.*

On the basis of existing theoretical and applied developments in the areas of team management, decision analysis, game theory and associated fields, the basis for development of probabilistic computational resource management algorithms must be developed. In particular, the following specific issues must be addressed:

[1] Which state variables must be estimated or measured in order to allow efficient decision making on the basis of those estimated values?

[2] How can the "coherence" (in the sense of relevant information consistency between cells) be estimated by each cell? What are the properties of alternative coherence measures that allow their comparative evaluation as to their usefulness in making efficient decisions?

[3] Can the adequacy and/or efficiency of decisions and that of the underlying strategies be characterized in terms of paradigms that are both realistic as well as more amenable to useful analysis than those provided by optimal control and/or optimality considerations? This is a particularly important issue whenever system adequacy is measured by the degree to which a given strategy compares to that which is optimal under a given optimality definition. [Reference 18] The complexity of the optimality computation for large systems with multiple operating environments appears to preclude classical approach.

[4] What are the criteria to be used by decision algorithms to prefer certain choices over

4-12

others? In particular, when it is more convenient to expend further resources about system state or to signal potential courses of action to other decision makers rather than to proceed to allocate resources on the basis of estimated states?

[5] Since clearly the scale of the proposed system will likely forbid analysis of extensive properties of strategies (e.g. optimality, etc.), it can be expected that studies will necessarily be confined to families of allocation procedures, that on the basis of heuristic considerations, can be expected to provide adequate resource control. As part of the required technological developments it is necessary to develop rational basis for the comparison and evaluation of as well as experimental basis for the comparison and evaluation of alternative heuristic approaches.

- *Development of resilient algorithms.* Existing operating systems rely on procedures which can be successfully activated and completed only when certain prior assumptions on input data validity and resource availability are satisfied. Whenever the processing required by these procedures has been completed, only two possible types of result can be obtained:

[1] the execution has terminated successfully and appropriate correct outputs have been obtained.

[2] an error condition has been detected and the quality of the output (as well as that of other system variables) cannot be guaranteed.

Use of this type of procedure at the global level in the tactical environment is most likely to result in a high number of completion with erroneous results as availability of the required resources cannot, in general, be guaranteed. Further, in the majority of cases, the use of resources until execution termination does not result in the attainment of partial goals which, while falling short of the intended goal at execution inception, are still useful to users.

Operating systems such as the MAXI-DOS, operating under conditions of low resource availability and reliability, must rely on the use of a new type of procedures designed to operate in a stepwise fashion using resources in a carefully graded manner and achieving a set of recognizable partially successful goals at different stages of their execution. Partial file transfers provide an example of useful partial goals of data handling procedures at the MAXIDOS level. If the stepwise design of any such partial transfer sequence included considerations based on the relative value of each segment transfer to the relevant maxinet parties, then the sequential information exchange can be designed so as to assure that any instantiation of the procedure will lead to the best possible results commensurate with the resources available during that particular activation.

Technological requirements are needed to develop algorithms having the property of resilient execution and adaptation to dynamic changes in resource availability. Functional areas of interest include all activities performed by the MAXIDOS including retrieval, update (with particular emphasis on algorithms that assure "partial" consistency of replicated copies), backup transfer and recovery of data base portions.

4-13

4.4 Security

4.4.1 Rationale for the selected architecture.

Analysis of existing information exchange practices and patterns in the tactical environment shows that there exists a substantial requirement for generation, storage, processing and transfer of information elements of diverse levels of security classification. Further, tactical information must be assessed by users at multiple levels of clearance in compliance with a well defined security policy.

At the global level, the MAXIDOS is essentially concerned with direct transfer of information between cells. The proposed design approach to secure transfer of classified information relies on several basic concepts:

a. the use of a well defined, dynamically reactivated authentication procedure performed by each cell to identify other cells as such, and to eliminate the possibility of cell impersonation by an unfriendly agent.

b. the transfer of information to other cells using a cryptographic approach dependent solely on the nature of the security classification of information and on the potential existence of a relevant need-to-know by a user of the information-receiving cell.

c. use of cryptographic schemes that vary dynamically with time so as to assure classified information protection even on the event of enemy capture of a complete cell resource.

d. the assignment of policy-compliant information dissemination responsibilities within each cell to intracell control procedures (i.e., the MINIDOS)

The rationale for (a) is clear and intended to prevent both cell impersonation by an unfriendly agent as well as to prevent disclosure of information to passive elements able to listen to the network transactions between cells. Dynamical reactivation authentication procedures between cells indicated in (a) above, is an unfortunate complexity required to minimize secure information loss in the event of enemy takeover.

It must be remarked, however, that the dynamic configuration management algorithms postulated for the maxinet rely on continuous measurements and rechecking by each cell of the global system variables that define system status in order to have a solid informational basis to perform resource allocation decisions. This continuous skepticism is required by the volatile nature of the parameters that characterize availability of system resources at the global level. The detailed discussion of the rational basis for this approach constitutes the major portion of Section 4.5 below and therefore no additional arguments will be provided here.

Design choices (b) and (d) above are based on the fact that, irrespectively of the amount of data protection to be required of the MAXIDOS, still the MINI-DOS in each cell must perform all basic access control functions at the local level (i.e. rights of local users versus classification of information). Once the MINIDOS has determined the legality of a transaction, it would be redundant to require rechecking by the MAXIDOS (either that resident at that cell or that of a remote cell required, for example, to provide classified information). Further, global access control at the MAXIDOS level requires each cell to keep an accurate map of the identity of all relevant global users, while performing individual authentication on each one of them. Both performance as well as feasibility considerations show this to be an undesirable approach.

Finally, the use of dynamically changing cryptographic schemes is intended to decrease both the value of a potential enemy cryptographical solution as well as the amount of information disclosed through the maxinet in the event of an enemy cell takeover.

At the intracell level, security requirements are dependent on the specific cell being considered and have not been specified in detail. At a general level, however, it can be expected that the intracell security problem will be much more difficult to treat than the global security problem due to the need to consider further resource sharing modes (such as the use of system programs by untrusted processes) and to the potential lack of physical isolation capabilities (such as those existing between cells) to effect logical isolation between users and protected resources.

4.4.2 Evaluation of Alternative Solutions.

Solutions that require performance of authentication and validation of user requests at the MAXIDOS level require MAXIDOS resource management both for local as well as remote cells. Besides adding complexity to the MAXIDOS design and operation, it is doubtful that the communication resources of the .naxinet could adequately provide sufficient support to assure efficient performance of the required access control functions. MINIDOS functions, on the other hand, are already concerned with a wide variety of resource protection requirements at the local level and are, therefore, a more logical choice for the authentication and verification of remote information requests, which usually can be tested for legality on the basis of the nature of the request.[1]

_____

1. Notwithstanding, the present approach does not forbid requests for information where the MINIDOS may require examination of the

The proposed approach information protection is therefore, more desirable and consistent with the basic hierarchical management philosophy of the overall design, that frees global functions, to the largest possible extent, from having to consider the internal processing characteristics of individual cells.

4.4.3 Open architectural issues.

The actual nature of the authentication and encoding functions to be used at the maxinet level must be precised in detail. In particular, the nature of the schemes used to vary coding schemes and keys with time should be established.

Authentication mechanisms must be designed so as to minimize the possibility of cell impersonation, particularly those that employ an enemy overtaken cell.

All aspects of intracell security architecture have not been specified in this initial study, due to their basic dependence on the eventual nature and characteristic of each cell.

4.4.4 Technological requirements.

Technological requirements in the area of security are closely related to issues found when attempting to precise the proposed architecture to a more detailed level. The specific technological needs detailed below, have therefore, a close correspondence with the open architectural issues detailed in the previous subsection:

● _Dynamic data encryption/secure communication protocols._ Algorithms for the dynamic production and modification of cryptographic keys must be developed. These algorithms must be responsive to scheduled as well as

_____

actual data before allowing local user access to it. This eventuality should be avoided, however, as it promotes maxinet transfer of unnecessary information.

nonscheduled system events. The capabilities provided by public key systems [Reference 19] to alleviate the problems associated with global key distribution (usually requiring a substantial deal of intercell coordination) must be examined.

- *Dynamic authentication algorithms.* Network cells must assure the authenticity of other interfacing parties in the maxinet. Developed algorithms must assure not only the identity of each cell, but also likely in combination with physical/human security procedures) they must help to detect the possibility of enemy cell takeover and operation of a federated cell. The possibility of performing authentication in combination with other security related activities, such as key dissemination should also be studied.

- *Multilevel multiuser distributed security.* Methods, techniques and procedures to develop and certify secure operating systems must be developed. Since most of the cell architectures have not been defined, the studies should concentrate on defining design principles and architectural characteristics that promote secure operation and which could be used systematically as guidelines to develop specific cell configurations.

4.5 Configuration Management

4.5.1 Rationale for the Selected Architecture.

The need to manage a time-varying amount of tactical processing resources has already received Air Force recognition in the form of the development of the concept of a "rolling force package". This concept is specifically aimed at assuring dynamic modification of the tactical system configuration in response both to changes of the specific missions performed by the tactical force and to changes in the tactical environment.

At the information processing level, the changing nature of processing requirements and environment is translated into the need to deploy a system capable of continuous and efficient reconfiguration of its processing, storage and communication resources.

In the proposed architecture, reconfiguration needs exist both at the cell and global levels. At the cell levels, local configurations may require modification in order to be responsive to dynamic variations in the processing and storage workload. At the global level, cells may join or leave the maxinet as the tactical organizations they support become operational, inoperative or perform physical movements in the tactical field. Further, cells may become connected or disconnected from other cells during processing and therefore the maxinet configuration perceived by one cell may be radically different from that perceived by others.

The hierarchical management breakdown between MAXIDOS control of global tactical data resources and MINIDOS allocation of processing and data resources at the local level recognizes the essential differences in reconfiguration needs and resource capabilities at the intercell versus intracell environments.

The MAXIDOS must be concerned with global management of unreliable resources through unreliable, noisy, low capacity channels. The MAXIDOS must choose from an extensive set of potential decisions which must be evaluated using incomplete and/or ambiguous data. Decision algorithms should provide a rational basis to perform independent resource allocation decisions, possibly opting to signal intentions or demand additional information from other cells. In the global context, reconfiguration is very much a cell dependent relative concept which expresses the variations in global resource availability perceived by each

cell. At this level, reconfiguration issues are related to the correctness of each cell's view of the world and its consistency with other cell's views (coherence). The nature of the tactical environment makes development of a correct consensus among cells an ideal goal achievable only when perfect environmental conditions occur.

On the other hand, at the local level, it can be assumed that all the resources required to assure correct and consistent perception of cell status (i.e. consensus) by all its processing elements are available (together with all the potential advantages to be gained by their efficient usage). At this level, issues such as the development of design concepts that promote internal cell reconfiguration with minimal disruption to ongoing processing deserve to be considered.

Focusing on the determination of the amount of resources required to achieve those reconfiguration goals are likely to be of high practical value to assist in the detailed design of internal cell architecture, while the somewhat dual question (at least from a classical optimization theory viewpoint) of efficient usage of a *fixed* amount of resources could be expected to provide answers of value in the development of global control algorithms.

From a processing point of view, reconfiguration needs at the maxinet level are also very different from those present when managing the internal configuration of each cell. At the maxinet level, the primary requirements for cells to monitor system configuration is related to the need to assess remote cell capabilities either as a primary or backup source of tactical data base elements. At the cell level, at least in principle, processing elements may be assigned and reassigned to support diverse computing functions, task processing may be shared by several processing elements and the assignment of

local data base elements to storage devices can possibly be changed due to external (e.g., higher information volume from cell inputs) or internal (e.g., decision to provide better performance) reasons.

The design presented here properly recognizes the fact that the nature of the reconfiguration problems and needs is very different at the global and local levels and correspondingly emphasizes approaches likely to produce the best possible results in each processing domain.

4.5.2 Evaluation of alternative solutions.

Solutions requiring a correct and consistent assessment of system status by all cells in the maxinet must necessarily rely on the use of extensive exchange of management information (e.g., coordination messages) between units in order to assure those integrity/coherence goals. Further, processes such as cell union or secession or global data reorganization must be both preceded and followed by intercell exchanges intended to assure a commonality of processing intentions and a consistency of individual views about the starting, intermediate, and final system states.

Algorithms based on any such approach must necessarily experience long delays before being successfully completed. Further, disruption to communication and processing capabilities, such as those expected in the tactical environment, are likely to prevent completion of the majority of the attempted exchanges. In all the cases included in that majority, failure to complete is tantamount to complete failure of the reconfiguration effort with a corresponding waste of the resources utilized until the moment of failure.

The proposed approach is based on the use of algorithms that attempt to produce the most efficient results

commensurable with the resources available to the procedure during its execution. The actual results obtained by activation of the procedure, whether totally unsuccessful, partially or completely successful, could be used by each cell to revise its view about the future desirability of reusing the procedure under similar circumstances.

On the other hand, at the local level, resources required to gain concensus and coordination between processing elements will be available and therefore, the performance and correctness advantages derived from achievement of a correct consensus can be realized.

4.5.3 Open architectural issues.

Most of the open related architectural issues in the area of configuration management involve the definition of additional detail about the nature of the algorithms and procedures to be used for each of the possible reconfiguration activities to be performed.

At the global level, those activities include:

● Cell union to maxinet

● Cell secession from maxinet

● Primary reallocation of data elements to cells

● Reallocation of backup responsibilities to cells.

At the local level, the list is considerably more extended possibly having to address such issues as task partition and allocation, file migration and computing function reallocation.

4.5.4 Technological requirements.

In order to initiate the examination of the detailed design and feasibility issues briefly described in the previous subsection, it is necessary to perform a series of related research and development activities devoted either to uncover new reconfiguration/status monitoring procedures or to establish

the usefulness of existing algorithms in the tactical C^3 system.

Considerations based on the nature of the tactical environment and discussed extensively above indicate that new developments will be required to allow eventual implementation of useful reconfiguration procedures at the MAXI-DOS level while research directed towards implementation of MINIDOS capabilities should focus on applicability of existing concepts and relevant modifications.

Consistent with the above arguments, two major technological areas can now be precised as requiring additional near-future development:

● *Global configuration management procedures.* The purpose of these activities should be the development of new reconfiguration algorithms that, while promoting the development of correct, coherent views of system configurations, are able to function at the best level of performance commensurate with existing resource availability. The studies should focus on the practical value of achieving correctness and coherence in order to effect efficient reconfiguration versus the possible risks incurred by implementing reconfiguration approaches on the basis of incomplete, imperfect information.

● *Cell reconfiguration techniques/algorithms.* The usefulness of existing techniques and their impact in resource requirements at the local level must be examined to determine their suitability for implementation as part of specific cell's MINIDOS. The studies should include the development of distributed design principles (particularly in the area of physical data base distribution) which promote efficient, low-interference reconfiguration.

4-18

4.6 Data Base Management

4.6.1 Rationale for the Selected Architecture.

The basis for hierarchical breakdown of data base management responsibilities at the global and local levels is the same used to allocate control responsibilities between MAXIDOS and MINIDOS functions. As these arguments have already been provided in detailed form elsewhere, they will not be repeated here.

For similar reasons, arguments supporting the need for development of novel probabilistic resource management algorithms at the global level, also support the need for probabilistic data base management algorithms.

It is important to remark, however, that the need to develop operating system algorithms having adaptation and resiliency properties is the basis for the notion of data criticality. One of the fundamental comparisons between alternative distributed data base management approaches is how the algorithms employ data criticality in in decision making.

The desirability of a high-level information (rather than data) oriented language at the global level (i.e. MAXIDOS/maxinet) follows from the homogeneous nature of the information handled by every federated cell (i.e. tactical air messages and air situation models) and is also justified by the need to provide limited human backup of cell information processing activities in the event of ADP equipment failure in some cell. The use of a common information handling message frees MAXIDOS functions resident at individual cells from consideration of details about the individual characteristics of the diverse DBMS that may likely be found across the cells federated into the maxinet.

At the local level, however, the need to provide extensible, modifiable cells capable of being initially deployed utilizing existing resources imposes consideration of translation/conversion schemes [Reference 20] providing for use of a common data base distributed over multiple heterogeneous processors.

4.6.2 Evaluation of alternative solutions.

The low capacity, unreliable characteristics of communication through the maxinet precludes the use of approaches requiring the definition of a "control consensus" between the different cells before an actual allocation decision is made.

In these approaches, extensive coordination messages must be exchanged between the cells, thus requiring existence of fast, high capacity, high availability communication channels.

The same arguments indicate that approaches assuring consistency of multiple distributed copies of data base portions require existence of information exchange capabilities which will not be available at the global level. Consistency, as well as some of the other desirable system properties already discussed, must be considered to be an ideal goal generally attainable only to a limited extent, which is related to the availability of computational resources. Thus, a whole spectrum of possible degrees of property attainment must be considered whenever user perceived characteristics of the system are discussed. This quality assessment approach is markedly different from that used for existing data base management systems which, to a large extent, are characterized in terms of having valued variables (i.e. consistency, integrity).

In the proposed design, duplicate copy consistency (for specific data base segments) must be defined or

4-19

determined to be valuable to the system objectives and its attainment will be conditioned by existence of required resources, estimates of probability of success and relative values of other processing goals.

The use of intercell communication language defined at lower levels of data abstraction (e.g., logical level in the ANSI/SPARC model terminology [Reference 4]) require further MAXIDOS involvement into the nature of local data base implementation details while reducing human understandability of the contents of intercell exchanges. The approach utilized in the design, emphasizing usage of conceptual schemes, user-oriented primitives and self- defining data elements, is also a better choice as a (generic type of) target language to which data representations and requests originating at the local level must be translated in order to be exchanged over the maxinet.

4.6.3 Open architectural issues.

At the global level, a number of open questions must be answered in order to further precise the nature of the data definition/manipulation capabilities of the MAXIDOS. The following paragraphs briefly examine those open issues.

The definition of the flexibility of data retrieval requests with respect to the location of the data required to resolve any given query has not been attempted. Specifically, it must be determined whether queries requiring retrieval from more than one cell (as well as all performance related query decomposition processes) will be required/allowed. The advantages of providing more flexible and encompassing retrieval schemes are obvious but must be measured against the need to limit performance of resource consuming transactions due to the unreliable behavior of communication and processing elements at the global level.

Basis for the dynamic distribution and reorganization of the global data base in response to workload or mission changes have not been precised. In particular, the rationale for local storage versus remote access of specific elements has not been specifically studied for typical tactical data base elements. Resolution of these matters must await both the results of precise data usage studies as well as the results of detailed tactical information/data definition and evaluation studies such as those mentioned before in Subsection 4.1.4.

The nature of mechanisms to be used to measure workload or to register mission changes that may affect system resource allocation must also be defined at a higher level of detail. Availability of adequate measurement/feedback mechanisms has been assumed as are the basic tools to derive data criticality assessments. These assessments are required by the MAXIDOS decision functions to determine the relative value of resource allocation choices.

In summary, open data management issues primarily involve the definition of additional design detail which further specifies the nature of the required data management functions both at the global and local levels. Due to the relatively low level of development of distributed data management concepts, resolution of most of these issues must await technological developments, particularly in the area of feasibility evaluation and demonstration.

4.6.4 Technological Requirements.

The previous subsection identified a number of open architectural issues requiring resolution in order to provide a more detailed specification of the tactical information system characteristics. To resolve most of these issues, it is necessary to further advance the understanding of distributed data base

management, thus providing a rational foundation for the performance of relevant design choices. Specifically, these types of major general technical developments are required:

- *Extensible distributed data management systems.* These studies will mostly benefit the design at the local (MINIDOS) level by providing principles and methods for the development of efficient data base management systems over heterogeneous processing elements integrated into a system capable of flexible modification by variations to its hardware or data base structures. Studies should pay particular attention to the following problems:

  - Data base elements conversion/translation

  - Data base conversion/translation

  - Dynamic data element distribution/redistribution

  - Integrity/Consistency Assurance

  - Query decomposition

  - Efficient data base partition

- *Probabilistic data management algorithms.* These studies should focus on the determination of feasibility of probabilistic management of data resources and can be expected to have major applicability at the global (MAXIDOS) level. Relevant questions to be answered include:

  - What are the rational or experimental basis for the derivation and measurement of data criticality? Can linguistic techniques for the analysis of data bases be effectively used as an aid in criticality determination?

  - What are the relevant variables (i.e. beyond data criticality) required to perform effective management of resources using probabilistic schemes?

  - What is the value of obtaining some form of consensus or coherence between decision elements before resource allocation?

  - What are relevant performance parameters to be measured in order to determine the extent to which data quality properties of systems (i.e. integrity, consistency, availability) have been achieved at a given time during system operation?

4.7 An Approach to Evolutionary System Development

One of the themes presented in this document is that of placing more emphasis on the development process. In the same manner that one would place more emphasis on an automobile factory than on the first automobile produced, data processing system development must be considered more in terms of an on-going production cycle. Several concepts must be accepted if this is the case. First is that development resources should be more concentrated and more emphasis should be placed on specialized technology for software production. The nature of system procurement followed by the government tends to cause fragmentation of development resources. Each individual project must support its own development resources and those resources remain only for the duration of development period. Price competition between contractors tends to prohibit maintenance of expensive support facilities.

A second premise that must be accepted in this production concept is that there are cyclical production schedules with firm requirements for working models. The TAFIIS data processing configuration should be treated as a product which is deliverable on future schedule determined by contingency plans and alert status in response to world crises. A data

4-21

processing system to support a Tactical Air Force contingency plan might be deliverable within a week, augmentation within a month, and a replacement system within a year. In the following paragraphs CSI will describe a concept for a Centralized Development Facility which addresses many of the development problems identified in the discussion of the data processing architecture and technology requirements.

4.7.1 The Centralized Development and Staging Facility Concept (CDSF).

The CDSF attempts to bring together in one development facility the technology resources, real data, users, and operating environment constraints. This concept is not the same as that of the SDC Software Factory where emphasis was placed on uniform requirements definition and software production. The software factory, like other types of system development concepts, does not take into account the significance of requirements volatility or the value of intermediate working versions. The CDSF concept tends to reduce the distinction between the development, training, and operational processes and likewise the staff assigned to those positions. The creed of this facility might be stated as "make it work". The centralization is aimed at improving the transfer of technology from the research stage to the operational stage and improving the feedback of operational needs to the system developers. It is OSI's premise that real data and personal contact between users and system developers is needed to achieve this flow of information and the maintenance of a continuous production cycle.

4.7.2 The Location.

The CDSF would ideally be located at an air base with the following attributes:

[1] Large computer room with following facilities:

- Maxinet communications (ARPANET, AUTOD N, tactical radio equipment sets)

- Computers (emulations of current systems, breadboards of upcoming systems)

- User terminals (programmer, trainer, student, maintenance, etc.)

- Classrooms

- Vendor facilities (software and hardware)

- Data storage (unclassified, collateral, SI/SAO)

- Mininet communications (wideband bus, high-speed links between development computers

[2] Tactical configuration staging compound

[3] Storage area for contingency data processing components

[4] COMSEC monitoring unit

Colocation of the CDSF at an active air base would allow operational testing of ground-air communications as well as ground based communications. The CDSF could function in various modes:

[1] system operator training (i.e. making the equipment work)

[2] data base maintenance (preparation of real data for contingency plans)

[3] software maintenance (repair or upgrade of task groups)

[4] hardware maintenance (repair or inventory upgrade)

[5] functional test (thread tests with software, data, and user components)

[6] multi-thread tests (interference, contention, and loading evaluation)

4-22

[7] user training (initial introduction or updates on data processing usage)

[8] integration of deployment configurations (hardware, communications, data, software, users)

As pictured in Figure 5, the tactical configuration staging compound would be physically segregated but in close proximity to the main computer facilities. The close proximity allows immediate access of users to the staff and support services of the CDSF whenever a problem is encountered. Users can split their time between operation in a controlled and friendly environment for training and data base preparation and testing and evaluation of the field configuration.

Physical segregation of the tactical equipment allows the enforcement of field procedures for operation and maintenance of deployable equipment, data, and support facilities. Both users and system developers will be faced with the realities of performing upgrades to data and software components through field communications. This concept of transferring the development system to the field system is intended to evolve procedures which can be continued once the system has been deployed remotely from the CDSF.

4.7.3 Concept of System Evolution.

If the CDSF always has the mission of being able to stage a working data processing configuration within a specified and limited time window, then the concept of system evolution has a more specific meaning. From a cold start, if the Air Force was required to deploy a working information system in 30 days there would be an obvious emphasis on communications equipment, intelligence data, and contingency planning. Very little attention would be given to untried or unreliable data processing equipment. For contingencies a year away, there would be time to introduce new

hardware which could be operated on a standalone basis and provide local support to specific elements or users. Extensive inventory changes requiring joint force cooperation and lengthy integration testing or unit training would not be included because of the high risk.

Once local data base support has been established for individual elements then digital message processing and electronic mail services can be effectively utilized in the field environment. Exercises in Germany have demonstrated that high speed digital communications with a computer on one end and manual message handling on the other end is not compatible. The computer is not able to effectively filter, buffer, or interpret the needs of the user sufficiently to prevent saturation of the manual message handling process. On the other hand, an automated message terminal can buffer all incoming messages at maximum line rates and the user can select messages from the buffer at his own rate.

Once effective digital communications have been established between elements of TAFIIS and those elements have local data base capabilities, then users can make effective usage of external data bases.

4.7.4 The Human Element.

An important aspect of colocating designers, data base specialists, communicators, intelligence analysts, and operations planners with field users is that informal communications networks can be established through personal association. These informal communication channels are vital in the continued maintenance of the hardware, software, and data base components of the information system. Much of the technological expertise in these areas is oral and would not be committed to paper even in the best of conditions.

Figure 5. Centralized Development and Staging Facility (CDSF) Concept

Knowing who might have the answer to an operational problem, whether it is in intelligence, operations, or maintenance, is a major factor in the adaptivity of the system.

Informal technical communications should be supported after personnel and equipment have left the staging area through electronic mail, newsletters, and telephone conversations.

Another aspect in the CDSF concept is that the facility operation itself is a training and personnel development device. In the way that strategic intelligence facilities provide a means for advancing the education and career of intelligence analysts, the CDSF can provide an assignment location to advance the education and career opportunities of the blue-suit data processing specialist. Rotation of data processing personnel through the CDSF and field assignments will benefit both the field operation and development process by building a cadre of highly-skilled professionals to support the system. The key to building this type of staff appears to be a combination of giving individuals assignments with an active mission and in a field which continually advances their professional skills. The CDSF concept is an alternative to assignments with tactical units where technical support is very limited and personnel are isolated from new developments and on-going activities in the C^3I community.

4.7.5 Technological Support.

A prerequisite for the CDSF location should be that it be near one of the centers of industrial data processing technology. Continuous contractor support will be required to introduce new software and hardware technology. Contractor support can be much more effective if immediate access to real data and interfacing hardware components. The CDSF provides a way in which developers can be brought face-to-face with users, data, operating constraints so that latent design flaws are not perpetuated through the system as a result of misunderstandings or incomplete data. Consolidation of development facilities can allow for better utilization of high cost technology support such as:

[1] operating system

[2] system software

[3] hardware maintenance

[4] performance monitors

[5] hardware integration test equipment

[6] configuration management

[7] documentation production

[8] programmer training

[8] software quality control

These technological support items if fragmented over many sites would be much more costly and not nearly as effective in producing a working and deployable system.

4.7.6 The OPSEC Problem.

Operations security is a major problem for tactical systems which will have access to sensitive intelligence data or that will operate in conjunction with new weapon systems or intelligence collection platforms. Secure data storage is not an insignificant problem and is a problem for contractors involved in development and for tactical systems in a garrison mode. For systems in garrison, limited secure storage may mean that the users and data bases cannot be kept current. This factor impacts directly on readiness of units. For contractors in a development mode, it may mean that data base structures or user interface designs are inadequate because they are unaware of how the user will handle real data.

Another aspect of operations security is communications security (COMSEC). The benign environment of the unclassified development and training facility does not prepare either users or operations support staff with the knowledge of how to deal with a high threat electronic warfare environment. Handling of real data in the CDSF during system staging can provide an opportunity for COMSEC training and COMSEC procedure monitoring and evaluation under controlled conditions. The main computer facility would have multi-level operation to support different stages of software and data base development as well as provide training to users with different security clearances.

4.7.7 Standardization/Interoperability.

The communications and operating system concepts used in the CDSF will have the greatest effects on standardization and interoperability. The CDSF policy regarding protocols, information structures, and dissemination control will establish the adaptivity of the system for:

- application module evolution

- software maintenance

- hardware upgrade

- inter-service operability

The development of the user interface within the control of the CDSF has several interesting aspects. First is that the CDSF can control the end-user command language and display formats. Second is that the CDSF will be facing users of many different skill levels and duty positions in its requirement to stage all aspects of system operability. With this centralization of user interface development, the CDSF configuration control mechanisms can be used to enforce commonality in the user language and at the same time evaluate the benefits of tailoring display and command formats to individual users. Again, the immediacy of the operational

evaluation to the system development process will allow experimentation and evolution of an effective man-machine interface.

4.7.8 Inventory.

Evolving the massive inventory of communications, data processing, and other support equipment is probably the biggest impediment to the continued evolution and upgrade of TAFIIS. The CDSF concept can help to alleviate this problem by minimizing the inventory of data processing equipment in a garrison state. If most of the equipment needed for contingency plans is centralized at the CDSF storage facilities, then only that amount needed for deployment and augmentation need be maintained. However, if complete configurations are assigned to garrison units then a significant amount of duplication will occur. In an actual deployment, the equipment will tend to be consolidated according to situational needs and control of equipment does not necessarily remain with the original tactical unit. The data processing equipment inventory needs extremely tight control to ensure interoperability and to achieve a higher turnover rate to utilize current hardware technology. The centralized inventory management system controlled through the CDSF seems more appropriate to achieve this goal than distribution of inventory.

4.7.9 Training.

Training provided to users through the CDSF would be oriented toward communications, data base development and maintenance, and data processing system operation. This training would not replace the specialized training in military operations, intelligence, or equipment maintenance that would be prerequisite for filling duty positions.

Centralization of the data processing inventory would have an impact on tactical unit training because fewer facilities would be available in garrison

4-26

status for that purpose. This would tend to imply that more unit training would be achieved through the CDSF or that users should have remote access to the CDSF facilities and support staff while in garrison roles.

Assignment to the staff of the CDSF would be in itself be a training opportunity for Air Force personnel with data processing MOSs.

4.7.10 Open issues.

The CDSF concept has many open issues regarding the cost and organizational viability of the concept. Some of the pertinent questions that must be asked include:

[1] How long would training classes last?

[2] Would personnel be assigned to the CDSF or be TDY during training or staging operations?

[3] How much of the data processing inventory should be controlled by the CDSF versus allocated to tactical units?

[4] How do you prevent various stages of system developments from interfering with each other?

[5] How critical is centralization of all support activities and what are the effects of providing some types of technical support and software development through remote facilities?

The last two questions are particularly pertinent to RADC experiments in the use of networking to integrate the results of software developments from multiple contractors. OSI's own experience in system developments involving multiple computer networks and multiple contractors indicates that the operating system and information structures are the key factors in integration. This is consistent with the operating concept and design issues described in this document.

4.7.11 Technological Requirements.

The viability of the CDSF concept depends on the ready availability of the following kinds of technology:

- Technologists (intelligence, operations, tactical warfare, data processing, data base management, networking, operating systems, software development, linguistics, etc.)

- Automated development and staging tools (program development, testing, integration, data analyzers, performance monitors, interface emulators, etc.)

- Access to real data needed in requirements definition (equipment specifications, data dictionaries, intelligence, contingency plans, protocols, etc.)

- Simulation/emulation of network communication services

- Configuration management tools

- Functional thread design language

- Dynamic and static operating system monitors

- Multi-thread performance enhancement

- C^3 community digital mail

One of the unique aspects of the CDSF approach is that there is much more emphasis placed on moving incrementally through the development process and maintaining continuity from version to version. There are several benefits which can be derived from the continuity of system configuration such as building on proven software and building a system technology base which will have long term benefits in the development process. It is expected that the CDSF will promote a professional atmosphere among technologists, trainers, and users. It would also be expected that Air Force data processing personnel would rotate through project technical, training, and user assignments over

a several year period.

The exploiting of experience in detailed aspects of the information system components and feeding that information back into the development process might be termed a heuristic procedure for system upgrade. This would apply to software, data base, hardware, and procedural aspects of the system.

5. TECHNOLOGICAL RISK ASSESS-MENT

In this section, the technological requirements presented in Section 4 evaluated in terms of their risk, payoff and expectation time for usable results. The evaluation performed uses a format closely related to the one used for technological risk assessment in the area of strategic C^3 system research and development needs [Reference 21]. The results of the evaluation performed are given in Table 2. The rows of this table correspond to the technological requirements identified in Section 4. The first column of the table identifies the technological need. The second and th' l columns are intended primarily f :der convenience and furnish a sequential identification number for each technological need (used for cross reference in the comments column) and a cross reference to the specific subsection in Section 4 where the need is described in detail, respectively.

The following three columns describe the risk, payoff and expectation time for usable results associated with each technological needs. The meaning of these terms and their values (high, moderate, low for risk or payoff, or number of years for expectation time) is explained below. The last column of the table ("comments") is devoted to present a brief description of the technological need and whenever appropriate, of its relation to other identified technological requirements.

Risk is a measure of the probability of failure by a competent research group in achieving results usable in the design, development or operation of a tactical C^3 distributed operating system.

- *High* risk means that the technological need requires use of difficult research, often along paths

that are not clear at this time.

- *Moderate* risk means that the technological need requires difficult research in a partially understood area along somewhat apparent research paths.

- *Low* risk means development of nonexistent but otherwise straight forward results.

Payoff is a measure of the value of producing technological research and development results to further the implementation of an efficient tactical information system.

- *High* payoff means that development of appropriate results is essential in order to implement an efficient tactical system.

- *Moderate* Payoff means that development of results is important to implement an efficient tactical information system, although incomplete solutions will allow development at acceptable levels of performance.

- *Low* payoff means that the technology is useful but not essential to the goals of tactical C^3 system development.[1]

Expectation time for usable results is the number of years (assuming finding of R & D studies at appropriate manning levels) required to produce results transferable to the design, development and implementation of a Tactical C^3 Distributed Operating System.

_____

1. No "low payoff" technological needs have been identified or evaluated in this study. The informal explanation of low payoff is included both for the sake of completeness and an aid to further precise the other evaluation values (e.g. high, moderate).

TABLE 2. TECHNOLOGICAL REQUIREMENTS EVALUATION

TECHNOLOGICAL NEEDS	TAFIIS DISTRIBUTED DATA PROCESSING TECHNOLOGY REQUIREMENTS					
	SEQUENCE NO.	REPORT SECTION	RISK	PAYOFF	EXPECTATION TIME	COMMENTS
Tactical Air Situation model development	1	4.1.4	Moderate	High	2	Important to overall understanding of information structure
Automated tactical dp requirement collection and analysis	2	4.1.4	Low/Moderate	High	1.5	Currently supported by RADC; intended to emphasize user-oriented data relations while developing further automated analysis tools
Distributed system simulation/emulation	3	4.1.4	Moderate	Moderate	2	Oriented toward development of evaluation tools and testbeds for the testing of global management concepts
MAXIDOS/MAXINET interface study	4	4.2.4	Moderate	High	2	Intended to uncover useful information exchanges that enhance decision making efficiency of MAXIDOS and performance of MAXINET
Criticality-oriented data handling languages	6	4.2.4	High	High	3	Related to No. 1 but focussing on actual user/cell languages to specify and handle data by criticality
Critical data transmission algorithms/protocols	8	4.2.4	High	High	3-5	Focussing on physical global net procedures to handle data on the basis of criticality related to No.'s 4,9,10
Heterogeneous processor interfacing	7	4.2.4	Moderate	High	2	Focussing primarily on low level protocol problems; No. 16 deals with similar questions at the logical db level
MAXIDOS/MINIDOS interfacing	8	4.2.4	Low	High	1	Primarily intended to uncover constraints on internal cell design derived from the requirement to interface cell with other cells
Probabilistic resource management	9	4.3.4	Moderate/High	High	3-5	Concerned with general concept development; see No.'s 14 and 17 for more specific studies

TABLE 2 (CONTINUED). TECHNOLOGICAL REQUIREMENTS EVALUATION

TECHNOLOGICAL NEEDS	TAFIIS DISTRIBUTED DATA PROCESSING TECHNOLOGY REQUIREMENTS					
	SEQUENCE NO.	REPORT SECTION	RISK	PAYOFF	EXPECTATION TIME	COMMENTS
Adaptive resilient algorithms	10	4.3.4	High	High	2-3	Primarily oriented toward data handling at the global level; related to to general aims of No. 9
MAXINET encryption techniques	11	4.4.4	Moderate	Moderate	2	Primarily intended to simplify key distribution problems as well as to minimize compromise in the event of enemy cell takeover
MAXIDOS-level authentication algorithms	12	4.4.4	High	High	3-5	Primarily intended to provide reliable identification/ authentication/ reauthentication procedures at the global level
Multilevel/Multiuser distributed security	13	4.4.4	High	Moderate	4-6	Focussing on use of distribution (at the local level) as aid to resource protection rather than as an additional complication
Global configuration management	14	4.5.4	High	High	3-5	Related to the general aims of No. 9 but focussing on reconfiguration issues at the MAXINET level; emphasis on determination of advantages gained by promoting inter-cell coherence
Cell reconfiguration procedures	15	4.5.4	High	High	3-5	Focussing on local reconfiguration in response to workload changes, failures, or performance improvement needs
Extensible distributed data base management	16	4.6.4	High	High	3-5	Focussing on interfacing of a dynamically varying number of heterogeneous dbms (at the local level)
Probabilistic data management algorithms	17	4.6.4	High	High	3-5	Related to the general aims of No. 9 but focussing on data management issues at the MAXINET level
Heuristic procedure development	18	4.7	Low	High	1-2	Applications are in tailoring of user interface and in extending lifetime of applications software

6. REFERENCES

[1]   Tactical Air Forces Integrated Information System (TAFIIS) Master Plan. Volumes I-VIII, X, Tactical Air Forces Interoperability Group, TAFIG-77-1, September 1977. (ADB023853L)

[2]   Experimental Evaluation of MIQSTURE; An Online Interactive Language for Tactical Intelligence Processing. Operating Systems, Inc., ARI Technical Report, September 10, 1979.

[3]   Saltzer, Jerome H.,Research Problems of Decentralized Systems with Largely Autonomous Modes. ACM Operating Systems Review, Vol. 12, No. 1, January 1978, pp. 43-52.

[4]   ANSI/X3/SPARC Study Group on Data Base Management Systems: Interim Report. FDT (Bulletin of ACM SIGMOD) 7, No. 2 (1975).

[5]   Fry, J.P., and Teorey, T.J., Design and Performance Tools for Improving Database Usability and Responsiveness, in Databases: Improving Usability and Responsiveness. B. Shneidermann (ed.), Academic Press, New York, 1978, pp. 151-189.

[6]   Bachman, C.W., Implementation Techniques for Data Structure Sets. Data Base Management Systems, ed. by D.A. Jardine, North-Holland, Amsterdam, 1974.

[7]   Teichrow, D.,and Hershey E.A., PSL/PSA: A Computer Aided Technique for Structured Documentation and Analysis of Information Processing Systems. IEEE Transactions on Software Engineering, SE-3, 1, 1977, pp. 41-48.

[8]   Nutt, G.J., Evaluation Nets for Computer Performance Analysis. Proceedings of the 1972 Fall Joint Computer Conference, AFIPS Press, Montvale, N.J., 1972.

[9]   Langefor, B., A Theoretical Analysis of Information Systems. Auerbach, Philadelphia, 1973.

[10]   Sundgren, B., Theory of Data Bases, Mason-Charter Pub., New York, 1975.

[11]   Von Stackelberg, H., The Theory of Market Economy. Oxford University Press, Oxford, England, 1952.

[12]   Cruz, J.B. Jr., Leader-Follower Strategies for Multilevel Systems. IEEE Transactions on Automatic Control, Vol. AC-23, No. 2, April 1978, pp. 244-245.

[13]   Luce, R.D. and Raiffa, H., Games and Decisions. Wiley, New York, 1957.

[14]   Raiffa, H.. Decision Analysis. Addison-Wesley Publishing Co., Reading, Mass., 1970.

[15]   Keeney, Ralph L. and Raiffa, H., Decision with Multiple Objectives: Preferences and Value Tradeoffs. John Wiley and Sons, New York, 1976.

[16]   Ho, Y.C., Kastner, M.P. and Wong, Z., Teams, Signaling and Information Theory, IEEE Transactions on Automatic Control, Vol. AC-23, No. 2, April 1978, pp.305-312.

[17]   Ho, Y.C., Kastner, M.P., Market Signaling: An Example of a Two-Person Decision Problem with Dynamic Information Structure. IEEE Transactions on Automatic Control, Vol. AC-23, No. 2, April 1978, pp. 350_361.

[18] Witsenhausen, H.S., A Counterexample in Stochastic Optimal Control. SIAMJ Control, Vol. 6, No. 1, 1968, pp. 131-147.

[19] Diffie, W. and Hellman, M., Privacy and Authentication: An Introduction to Cryptography. Proceedings of the IEEE, Vol. 67, No. 3, March 1979, pp. 397-427.

[20] Kimbleton, S.R., Wood, H.M. and Fitzgerald, M.L., Network Operating Systems - An Implementation Approach. Proceedings of 1978 National Computer Conference, AFIPS Press, Montvale, N.J., 1978, pp.773-782.

[21] Research in Network Data Management and Resource Sharing - Research Plan. CAC Document No. 164, JTSA Document No. 5510, Center for Advanced Computation, University of Illinois at Urbana - Champaign, Urbana, Illinois, June 1975.

IMPLICATIONS OF ADVANCED DATA
STRUCTURES FOR TACTICAL
COMMUNICATIONS

APPENDIX A

TABLE OF CONTENTS

1. INTRODUCTION

As the communication of information in the tactical environment becomes increasingly digitized, the problem of textual information saturation looms large. Most of the information which is currently transmitted by telephone and extracted for tactical op/intel use will become digitized in the near future. While this provides an important automated capability for transmitting and recording information, it offers a concomitant challenge in terms of what to transmit and what to record, how to summarize or aggregate recorded information, how to utilize such information to alert or predict, how to evaluate the credibility of information, and how to determine when recorded information is no longer useful in a volatile tactical environment.

Although some thought has been given to the distillation of verbose textual information, very little has been devoted to the related problems mentioned above. Moreover, the solution currently proposed for distilling the content of tactical messages (i.e., the JINTACCS message format, illustrated in Figure 1) is one that puts the burden of mechanical information distillation functions on the user, who is reduced to counting characters and specifying microfields to make things easier for the computer; thus, this particular approach to information distillation may create a variety of new problems in attempting to solve one that has existed for a long time.

It is clear that dealing with data elements rather than the typical natural language form of tactical messages is more efficient in terms of communications costs, as well as data base generation and maintenance costs, and associated information exploitation costs. The difficulty is to develop an approach to data element specification which does not require mechanistic behavior of the human information generator.

Such an approach has been developed by OSI for RADC under Contracts F30602-76-C-0194, 77-C-0060, 77-C-0144, and 78-C-0274. This approach reflects state-of-the-art achievements in artificial intelligence and language processing technology, which have considerable potential for tactical information handling systems of the future. In the course of these contracts, a methodology has been defined for creating and maintaining data structures called templates, which are compact structures for representing information on entities and events described in the text of intelligence messages. A testbed system which allows both interactive and automated generation of template data structures for air activities and satellite/missile events has been constructed. Although the technology was originally developed for Indications & Warning (I&W) applications, it is also relevant to tactical information processing problems. The discussion of the following section describes this technology and treats its implications for tactical C^3 communications.

```
UNCLAS
RELEASE/NOTAL
EXER/EXERCISE:BRAVE SHIELD 95
MSGID/MISREP/366TFW/0622014
SEGMENT/MISSION ID
MSNID/MSNNO:AR125/REQNO:27/FRAG:68/OPORD:524/RINGO 11
SEGMENT/TARGET STRUCK-SIGHTED
1EA/DE/TGT-ID       /TARGET-LOCATION /TGTTYP/SUBTYP/SPD    /DIR
    01 TAIGE17567   -L482025N0142450E BRIDGE BRGFTS
    02 ACS3579       U33UVP440440     VARMOR VTANK   40KPH NNE
1EB/DE/ON-TIME/OFFTIME/PCTDAM/PCTDES/PCTCOV/ORD    /QTY /CAL
    01 221530Z 221545Z   25            MK-83    3
    02 221715Z 221725Z   20     40     ROCKEYE  10
NARR/01 ONE SPAN OF 4 SPAN BRIDGE DROPPED
    02 TEN TANKS SIGHTED 2 DAMAGED 4 DESTROYED FFFF
SEGMENT/AIR INTERCEPT
1EC/DE/MILDTM /INTCP-LOC       /EN-ACFT-TYP        /ENG /DES /DAM
    01 221740Z L481200N0135000E MIG21J              4   1   1
1ED/DE/ALT    /ELEV
    01 T15KFT
NARR/01 MIG21J FLT EMPLOYED LOW LEVEL POP UP TACTICS INDICATING THE
PRESENCE OF GCI CONTROL. WHEN MIG LEADER WAS SHOT DOWN THE
REMAINING MIGS BROKE CONTACT. FFFF
SEGMENT/SURFACE-TO-AIR FIRE SAMS
1EE/TYP-SA-FIRE/SA-FIRE-LOC      /ALT     /INT/NO-SAM/MIS-DIS/EA
    SA8        L480100N0142050E T15KFT   MED    3 02504H  Y
    SA2        U33UVP350125     T1KFT    LT     2 05009L  N
AMPN/SPLIT S EVASIVE TACTIC LOW LEVEL EGRESS USED AGAINST SA8 FFFF
SEGMENT/AIRCRAFT LOST
1EF/DE/ACFT-NAME            /QTY /CAUSE       /CREW-DISPO
    01 F4E RINGO13       -  1 MSLS-AG    MIA
1EH/DE/TMLST/LOCATION        /LQFR
    01 1548Z L480102N0142054E EST
SEGMENT/TARGET WEATHER
1EG/TGT-ID       /TARGET-LOCATION /WEATHER
    TAIGE17567   L482025N0142450E BROKEN
    ACS3579      U33UVP440440     BLUE SKY
RMKS/RINGO 11 A FLIGHT OF 4 F4E FFFF
```

Figure 1. Example of Mission Report in

JINTACCS Format

2. A BASIC INFORMATION STRUCTURE FOR C³ REPORTING

In a tactical environment, an intelligence analyst asks "What's happening?" "Where are the OPFOR units?" "What is the level of threat?" "What does this mean in terms of what has happened before?" Most important, "What will happen next?" His counterpart in ops is also concerned with these questions, as well as the key question "What do we do about it?" "What are our courses of action?"

An information structure for the tactical C³ environment should thus accommodate this set of information parameters. The information structures called templates in the discussion of the previous section are n-ary relational structures which are based on the information parameters implied in the set of questions mentioned above.

A template is essentially composed of a set of information parameters or descriptors which represent the type of information that answers the set of questions shown in Table 1, which also illustrates the corresponding descriptors of a prototype template.

This is somewhat of an oversimplification of the template concept for convenience of presentation, since complex descriptors within the template are actually represented by pointers to other types of templates: e.g., object templates. Thus for the message given in example a), an object template reflecting an aircraft description is essentially embedded in the event description by a pointer reference, as shown in Table 2:

TWO SAAF CAPETOWN-BASED SAG22 ACFT ARE OPERATING OVER THE INDIAN OCEAN.

It is interesting to note that message (a) is incomplete in terms of the prototype template specification outlined in Table 1. Conspicuously absent is a time descriptor.[1]

For the instantiated or filled-in template to be complete, a time descriptor element must be satisfied. The absence of this element can be signalled to the analyst to indicate that this element must be supplied for the information representation to be complete. Thus, *intra*-template relations -- the set of relations connecting descriptors *within* a template -- provide an important means of alerting analysts to the missing information elements in data structures constituting subparts of the network of templates which represents an analyst's information model of a particular tactical situation.

Event summaries, which are subjects or titles and thus key sentences of a message, are often characterized by missing elements, as in example (b):

REMOVAL OF GENERAL WALUSIMBI, COMMANDER OF THE 2ND UGANDAN FIELD ARMY

The most striking missing element is the relation of Agent -- who removed General Walusimbi from the position of field Army Commander? In attempting to fill out the template, this is an important element which the analyst should search for in the text of the message. If the agent of the removal is not given in the text of the message, related

_____

1. Also a separate template because of the complexity of most time descriptors, which are derived only partially from explicitly stated times as in the Table 1 example; but must generally be reconstructed from the tense of the internal verbs, time operators such as "currently", which point to information in the message header or the textual context of the time referent, and the internal structure of a time descriptor, which may read "at 10 minute intervals for a period of six hours".

Table 1. Information Parameters of a Prototype Template

what	event type	airspace violation
who	agent (or object plus owner)	a Ugandan MIG-21
when	time of the event	at 0235Z 25 April 1978
where	location at which event took place	6 miles from the Kenya border near Suam
to whom	patient, or entity affected by the event	Kenya
why	interpretation of the event	probable reconnaissance mission

Table 2   Content Representation for Sentence (a).

```
+---------------------------------------------------------------+
|                                                               |
|                    EVENT DESCRIPTION                          |
|                                                               |
|   UNIT REPRESENTED:   EVENT                                   |
|   EVENT TYPE:         BE ACTIVE                               |
|                                                               |
|   OBJECT: ─────────────────────────────────────────┐         |
|                                                     |         |
|   REGION:             THE INDIAN OCEAN              |         |
|                                                     ▼         |
|      +---------------------------------------------------+    |
|      |                                                   |    |
|      |             AIRCRAFT DESCRIPTION                  |    |
|      |                                                   |    |
|      |    UNIT REPRESENTED:   OBJECT                     |    |
|      |    OBJECT TYPE:        AIRCRAFT                   |    |
|      |                                                   |    |
|      |    TYPE:               SA622                      |    |
|      |    SUBORDINATION:      SOUTH AFRICAN FLEET AIR FORCE|  |
|      |    BASE:               CAPETOWN-BASED             |    |
|      |    SET SPECIFICATION:  TWO                        |    |
|      |                                                   |    |
|      +---------------------------------------------------+    |
|                                                               |
+---------------------------------------------------------------+
```
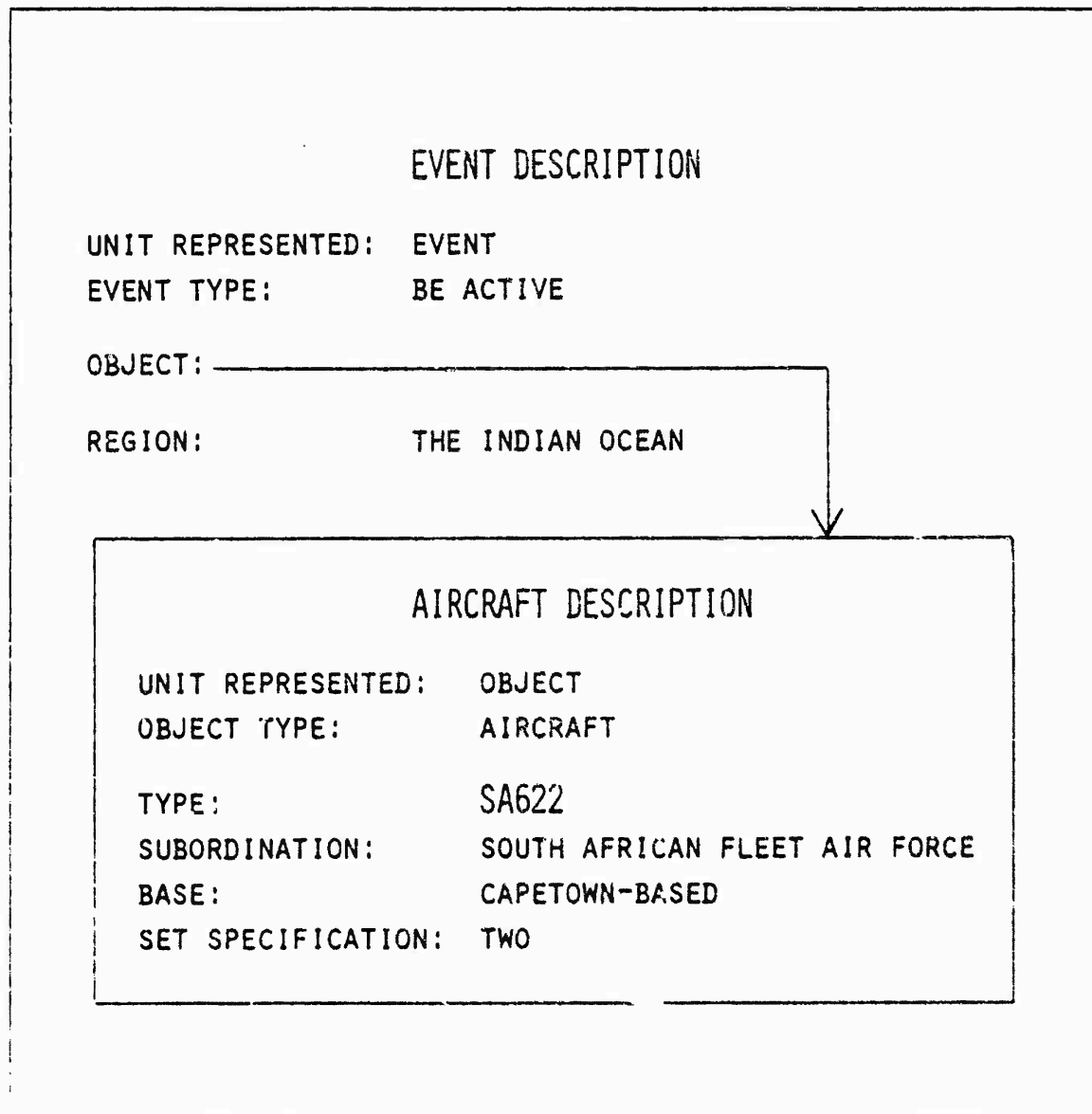
templates can furnish a set of possibilities for the analyst to consider. For example, who APPOINTED, (nominated, named) Walusimbi to the position of field Army Commander? APPOINT is a template which represents a necessary predecessor event to a REMOVE event. On the other hand, a REMOVE event is a possible, but not obligatory successor of an APPOINT event, for the PATIENT, or EXPERIENCER of the APPOINT action -- in this case, General Walusimbi -- may resign, die in office, or leave the country. The explicated network of *Inter* -template relations, both obligatory and optional, thus provides an additional means of alerting the analyst to the implications of an event, as well as to related events which may furnish data elements missing in the template currently being filled out (e.g. the probable person or organization who removed General Walusimbi from the command).

Moreover, because of the inter-template relation between REMOVE and REPLACE, a sequential event which has a high probability of occurrence, analysts can be alerted to look at potential replacements and predict the impact on the given tactical situation, which will result in the formation of alternative strategies by operations and planning personnel.

In summary, the application of template technology to information processing in a tactical C^3 environment can provide an important analytical aid from the following five points of view:

1. Templates constitute a powerful means for distilling the content of verbose textual messages into a compact format, which is not so compact as to be difficult for tactical C^3 personnel to generate and understand.

2. Templates are logical information structures which can be used to represent the analyst's and

commander's information models of the tactical situation.

3. *Inter-* and *intra* -template relations can assist the analyst in recognizing missing elements of information and predicting future events, allowing operations/planning personnel to define alternative courses of action in advance.

4. Templates provide a discrete representation of information which lends itself readily to statistical analysis, as in indications monitoring applications.

5. Event data available from the TIME and LOCATION descriptors of templates can be exploited to drive automatic plotting of ground movements and aircraft tracks.

Quite apart from the analytical, planning, and information handling benefits for which the technology was designed, the use of such structures for information representation can be expected to have a profound effect on requirements governing the design of distributed operating systems, databases, and data communications in the following areas:

1. More compact on line databases (verbose message texts from which templates are developed can be maintained as offline, backup, information);

2. Reduced intersystem data communications volume with higher information quality (capability for disseminating information in packages smaller than messages)

3. Definition of a basis for detecting and eliminating redundant information.

4. Definition of a basis for summarizing information and generating reports.

3. THE TEMPLATE AS A LOGICAL DATA STRUCTURE

In the air activities domain, there are templates for describing various classes of events (e.g., the class of "flights", "arrivals", "departures", "deployments", etc.); templates for representing the attributes of classes of persons (e.g., the class of aircraft pilots, the class of controllers); templates for other complex entities, such as dates, with their attributes of day, month, and year. In the missile and satellite domain there are templates for such event classes as "launch", "impact", and "re-enter", and various classes of space objects.

In general, an indefinite number of properties may be specified for any given class of entities. However, not all properties have the same degree of usefulness in a given context. The properties selected for inclusion in a template, must, therefore, be sensitive to the task the template is designed to support. Accordingly, templates include only that information which is particularly relevant and useful to the task at hand, and not the full range of facts one might find in the encyclopedia.

The template then, is the fundamental organizing knowledge structure for the representation of events and other entities. It is an abstract description of a class of individuals and as such it constitutes what we will refer to as an INTENSIONAL description, made up of a distinctive set of descriptors, or arguments.

"Individuals" are different unique entities in the world being modeled. As a simple example from the air activities domain, a particular flight is an individual event with the participating aircraft and its specific direction, time, and date as descriptors or arguments, uniquely defining the given event Thus an "individual" is a particular token or instance of a type, which is characterized by a set of descriptors.

The class of individuals to which an intensional description applies is called the EXTENSION of the general concept described by the template. Descriptions of individual events are called Event Records. Thus, the set of event records describing events of the same class, i.e., event records having all obligatory descriptors characterizing the event template representing that event class, constitute the EXTENSION of the concept described by that template.

Templates are an important component of the Event Representation Language (ERL) - currently under development. They reflect the priorities of the intelligence reporting system and support the analyst's view of the world. They are part of the domain-specific knowledge components of the system, and are stored in the permanent data base. They are the permanent knowledge structures from which representations of individual events and objects are derived for the purpose of storage and retrieval in an event data base. As described in the preceding section, they can also serve as the basis for identifying missing or contradictory information, for predicting future events, for summarizing data base contents, and for making various plots, including aircraft, ship, and ground force tracking, event/time/space displays, etc.

3.1 The Template Descriptor System for Air Activities.

As noted in the preceding section, a template is composed of a set of *descriptors*, or properties used for the description of events. Also, the set of descriptors characterizing an event template is essentially the *what, who, where, when,* and *why* information parameters of any event. The definitions of the descriptors and their organization reflect the representation of events and activities of various kinds.

Thus, flight reports include a description of the object(s) which is (are) doing the flying and frequently mention other relations such as the source of the flight, its direction, the area overflown, the destination, and the mission.

Human agents, such as pilots and navigators, are very seldom mentioned in flight reports. For this reason, they are regarded as presuppositions of the flight event: that is, the notion of a flight event necessarily involves the human agents or flight crew of the aircraft.

Table 3 shows the air activities descriptor system.

Table 3. Air Activities Descriptor System

A. **Motion related descriptors**

Agent(A)	animate instigator of the action
Object(O)	the entity that moves or changes or whose position or existence is being described
Source(S)	the location of the object at the beginning of a motion
Goal(G)	projected or actual destination of the object at the end of the motion
Direction(D)	direction of motion of object at time of observation
Path(P)	path or area traversed during motion
Extent(E)	extent of motion
Limit(L)	limit of motion
Altitude(Alt)	altitude of object at time of observation
Region(R)	general location of the action
Status(Sta)	begin, continue, end
Time specification(T)	time of observation or duration of the event

B. **Event related descriptors**

Mission(M)	purpose of flight

C. **Aircraft related descriptors**

Type
Class
NATO designation
Country of origin
Subordination
Homebase
Staging base
Set specification
Configuration

4. IMPLEMENTATION OF TEMPLATE-ORIENTED INFORMATION REPRESENTATION

4.1 Format and Content Structure of Tactical Messages.

Although a large volume of traffic covering a variety of subjects is generated each day, electrical messages are characterized by certain uniformities in format and content. These uniformities assist the analyst in interactively creating and updating templates (Section 4.2), and can be exploited for automated extraction of event/time/location data and generation of data base elements. (Section 4.3)

4.1.1 Format Characteristics of Typical Messages:

The most obvious uniformities involve format. There is a distinct set of formats associated with all military message traffic. The specific formats are mainly distinguished by different types of header and trailer lines and different sequences of header information. As opposed to these formats, which are specific to particular sources, the following types of format are general descriptions which cut across all sources, and all of which occur in tactical traffic.

4.1.1.1 Formatted Summary Traffic:

The formatted summary traffic consists of periodic summaries of events involving aircraft, ground vehicles, sensors, etc. The formats are arrays, where the rows represent individual events involving objects -- e.g., aircraft, weapons systems, military organizational elements -- and the columns contain properties of the event, including identification symbols, geographical coordinates, data/time group of the particular event, and other information.

Further format and content characteristics of this traffic are that the type of object and event described are directly derivable from the report title, there are no initial text lines, and the order of the columns of information is invariable. An example is given in Figure 2.

4.1.1.2 Semi-Formatted Summary Traffic:

on the other hand, *semi-formatted summary traffic* characteristically consists of a few lines of text which identify the nature of the events summarized, followed by a formatted segment containing the summary. Some of the text portions may be quite standardized, as in the summaries of submarine contacts presented in example (a) and (b) of Table 4[2]

others may vary widely.

4.1.1.3 Unformatted Traffic:

Unformatted traffic consists of messages which contain no formatted information in the body of the document (although there is a formatted header of some kind). Table 5 presents some examples of such messages. In general, these reports contain information as to the time and location of a given event, and may contain additional data giving the context of the event sequence or chain of related events, properties of the objects or events, the sources of the information, and some interpretation of the event.

The majority of these messages have a characteristic content structure which may be represented by the following formula, where the parentheses enclose optional elements and curly brackets enclose alternatives:

S represents references to the source of the information, which is an optional element: the initial sentence of example (b) contains a source reference.

_____

2. Headers and trailers are not shown in these examples.

```
ZCZCJKL082
PP XFNALA
DE XANADU #2247 1911845
ZNY MMNSH
KXKX PP DEF
P 1019052 JUL  74 ZED
FM ALCAN
TO ROMEO FOUR
ZEM
UNCLASS? NO WRONG DISSEM
ZZYYKALQ62
2/CG/0341-74
DAILY CANADIAN ACTIVITIES SUMMARY REPORT (DCASR)
YYGG
```

1. AIR ACTIVITIES

TYPE

T10	42180	SASKATOON	RGINA
BX5	32465	VANCOUVER	VICTORIA
BX5	36702	EDMONTON	CALGARY
H42	18543	VICTORIA	CAMPBELL RIVER
T10	40367	CALGARY	REGINA

2. SUBMARINE ACTIVITIES

ID	LAT/LONG	CLASS
W1	1620N13450W	E
W2	4045N15950W	H
W3	3030N16000W	Y
W4	4045N13170W	E
W5	1100N13930W	G

Figure 2. Example of a Message Containing Two Formatted Summary Sections.

$$(S) \ E \ L \ (T) \ \left(\left\{ \begin{array}{c} E' \ (I) \\ I \end{array} \right\} \right)$$

E symbolizes the major event being reported in the message, while *L* represents the location and *T* the time of the given event. In example (a) of Table 4 all this information is contained in the single sentence constituting the body of the message. In the majority of messages of this type, a description of the event and the location of the event occurs in the first sentence of the text. Information as to the time of event is often omitted, and must be derived from the header and from verb tense.

E' symbolizes further information on the event, e.g., properties of the objects involved, as shown by second sentence of example (b) and other occurrences in the event chain, as shown in example (c), Table 5.

I represents some interpretation or evaluation of the event, as in the third sentence of example (c), Table 5. The interpretation is often omitted, but when an interpretation is not made, more detailed information on the event is usually given.

messages containing more than one paragraph are generally structured along the same lines, where each paragraph reports a separate event, with associated location and time data, details, and interpretive comment.

4.1.2 Content Characteristics of Typical Messages:

Messages dealing with the same kinds of events are usually very similar in content.

In addition to the similarity of information relating to events of the same type, reporting procedures contribute to the uniformity of special content subsets -- e.g., air activities -- of the message traffic. These procedures

specify typical frequencies of reports for certain types of events, and describe guidelines for the production of such reports, including the kinds of information which should be presented: especially the properties -- type, class, call sign, home base -- of the object -- aircraft, submarine, ship -- as well as the location and/or movement reported on, and the time of the event. All of these standards contribute to a uniformity in event reporting. In fact, reports issued periodically in response to a collection requirement for a particular type of event tend to be very similar. The person generating such a report is not seeking to be original or imaginative in his reporting, but to communicate factual information as concisely and precisely as possible; it is therefore not surprising that the use of words and sentence structure is to a considerable extent standardized.

This standardized use of vocabulary and syntax in effect constitutes a special subset of English developed for the reporting of events of intelligence interest.

4.1.3 Implementation Considerations Based on Complexity of Content and Format:

The standardized features of this special language for event reporting allow automated as well as interactive exploitation of event information contained in the message text.

The table shown below presents an approximate measure of level of difficulty of implementation, corresponding to the level of complexity of the data in terms of the format ad content characteristics described above, and in terms of the contemplated degree of automation.

TABLE 4

(a) Summary of not SA or known friendly submarine contacts off the
    west coast of the Union of South Africa as of 02/0330Z:

        W1   3230S 0850E W

        W2   3310S 0945E Nuclear

        W3   Contact lost Y

(b) Summary of not SA or known friendly submarine contacts of a 02/0240Z:

        E4   3130S 0650E   Y

        E7   Contact destroyed   Y

        E10 3275S 0850E   W

(c) While none are being listed as positive contacts, not SA or known
    friendly submarine activity is particularly active in the
    Mozambique Channel and west of the Cape of Good Hope.
    Summary of not SA or known friendly submarine contacts as of
    02/23100Z:

        W5   2025S 4210E   Y

        W6   2250S 4130E   Y

        W7   2330S 4025E   Y

        W8   Contact Lost   Nuclear

        W9   3125S 0750E   W

        W10 3250S 0830E   Y

TABLE 5

a. At 301400Z the Angolan amphibious force from the Nova Lisboa area was located in the vicinity of 2125S 4010E.

b. Malagasy fishermen have reported that several Angolan naval vessels were in the vicinity of 2075S 4040E. Identification was not possible due to low visibility.

c. A single high performance aircraft entered Biafran airspace just north of Enugu 0500N 0490E. The aircraft flew to the vicinity of Makurdi 0600N 0645E then exited generally along the same route. Flight profile flown indicated a recon mission.

TABLE 6

	Interactive	Automated
Formatted	1	2
Unformatted	2	3
Semiformatted	2	3

Clearly -- because the knowledgeable human operator is driving the operation -- interactive processing of formatted material represents the simplest implementation case. Of course, this is still not a trivial undertaking because the human analyst must be able to treat the formatted data in terms of some information model which is relevant to his analytical task.

For example, assuming that flights of aircraft from a given point to a given point are of interest to him, an analyst can easily transform the data shown in tabular form in Figure 2 to a "fly" template of the form given in Table 3. Using the "fly" template and its associated object template describing the aircraft, it is immediately obvious that these formatted summary messages omit a substantial amount of information concerning both the flight event, and the aircraft, which, if significant may trigger the analyst to seek the additional information. If the missing information is not significant, a tabular summary such as the one presented in Figure 2 could be scanned automatically, and the structures generated for a TAC C^3 data base.

In the case of unformatted or semiformatted messages, the analyst scans the material and parses it into a cognitive structure as shown in Figure 3. The parsed information is accommodated by the data structure shown at the left edge, which is essentially a prototype template for representing air penetration activities.

In the following discussion, 4.2 describes an interactive approach to creating and filling out templates, while 4.3 treats an existing experimental system for automatically filling out templates.

4.2 An Interactive Approach to Template Creation and Maintenance

In order for template-oriented data-structuring techniques to succeed, it is necessary that the analysts themselves must be capable of generating new templates to extend the information model as required, and that they will maintain the information model dynamically (by updating the template and associated templates) in terms of which the model (or some component of the model) is realized. Thus the processes involved in generating, maintaining, and manipulating templates must not be excessively complex or time consuming. Template definition and processing procedures must not appear unwieldy or counter-intuitive.

In most cases for tactical analysts, components of their information model of a tactical situation which are not contained in a situation display or an order of battle file are inexplicit. It will therefore be necessary to elicit such concepts by an inductive process, beginning with the intelligence messages of interest, building on the entities and events of interest reflected in the contents of these messages to explicate the analyst's intuitive information model of a tactical situation. Part of the explication process involves the definition of inter and intratemplate relations.

4.2.1 Specification of Intra- and Inter-Template Relations:

The specification process involves both explication of relations between elements within a template (intratemplate relations) and those holding between templates (intertemplate relations). These relations are explicit

A-15

A SOMALI AIRCRAFT PENETRATED ETHIOPIAN AIRSPACE NEAR ASMARA (1310N3940E).

LOGICAL SEMANTIC STRUCTURE

PENETRATE (X1, X2, X3, X4, X5)
- X1   AGENT = AIRCRAFT
- X2   OBJECT = GENERAL AREA PENETRATED
- (X3)   LOCATION = SPECIFIC AREA OF PENETRATION
- (X4)   TIME OF PENETRATION
- (X5)   EXTENT OF PENETRATION
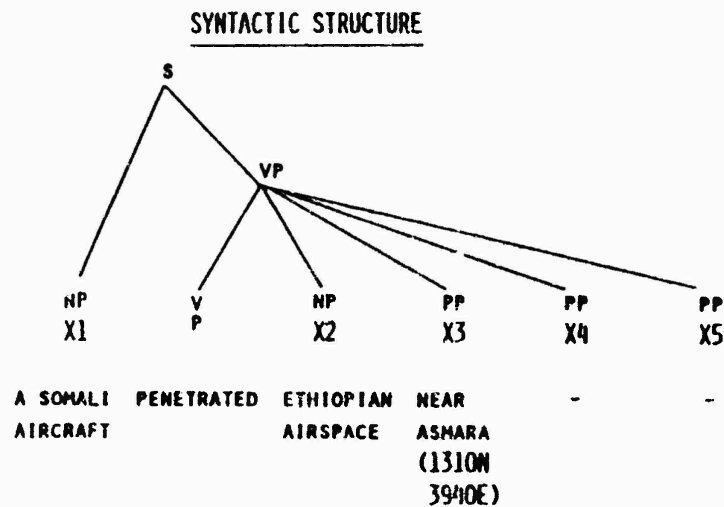
SYNTACTIC STRUCTURE

Figure 3. Structural Analysis of an Input Message Segment

representations of cognitive associations of various types which are utilized in inferential processing. It is important that these be explicated in order to assist the analyst in formulating inferences of varying degrees of subtlety. The following descriptive catalog of inference types appears in Rieger:[3]

[1] Specification inferences: A representation for an action has certain slots for information which must be filled in (specified). So, for example, it requires an inference to infer that if John hit Mary a hand was used to do the hitting.

[2] Causative inferences: What were the likely causes of an action or state?

[3] Resultative inferences: What are the likely results (effects on the world) of an action or state?

[4] Motivational inferences: Why did (or would) an actor want to perform an action? What were his intentions?

[5] Enablement inferences: What states of the world must be (must have been) true in order for some action to occur?

[6] Function inferences: Why do people desire to possess objects?

[7] Enablement-prediction inferences: If a person wants a particular state of the world to exist, is it because of some predictable action that state would enable?

---

3. Rieger, C., "The Commonplace Algorithm as a Basis for Computer Models of Human Memory, Inference, Belief and Contextual Language Understanding", in Schank and Nash - Webber (eds.), *Theoretical Issues In Natural Language Processing*, Cambridge, Mass. 1975.

[8] Missing enablement inferences: If a person cannot perform some action he desires, can it be explained by some missing prerequisite state of the world?

[9] Intervention inferences: If an action in the world is causing (or will cause) undesired results, what might an actor do to prevent or curtail the action?

[10] Action-prediction inferences: Knowing a person's needs and desires, what actions is he likely to perform to attain those desires?

[11] Knowledge-propagation inferences: Knowing that a person knows certain things, what other things can he also be predicted to know?

[12] Normative inferences: Relative to a knowledge of what is normal in the wold, determine how strongly a piece of information should be believed in the absence of specific knowledge.

[13] State-duration inferences: Approximately how long can some state or protraction action be predicted to last?

[14] Feature inferences: Knowing some features of an entity, and the situations in which that entity occurs, what additional things can be predicted about that entity?

[15] Situation inferences: What other information surrounding some familiar situation can be imagined (inferred)?

[16] Utterance-intent inferences: What can be inferred from the way in which something was said? Why did the speaker say it?

Of the types of inference listed above, *specification inferences* are reflected

in intratemplate relations (4.2.1.1); other types of inferences in intertemplate relations (4.2.1.2).

4.2.1.1 Intertemplate Relations.

Much of the exploitation of intratemplate relations is based on implicit linguistic and real world knowledge. Thus, an aircraft must fly *somewhere* -- there must be a destination of the flight within the operating range of the aircraft type. The aircraft must have an owner (e.g., Uganda), a home base, (e.g., Gulu), and a niche in some organizational structure (e.g., subordinate to the 2nd TRW and the USAF.

Similarly, taking an example from the area of political affairs, an *assassination* template necessarily implies -- to anyone who understands (i.e., has knowledge of) the concept -- an assassin or agent; a victim or patient; and an instrument, or means of performing the action. Although all these relations may not be explicitly represented in a text -- since only mention of the patient is obligator -- the reader infers the existence of agent and instrument, and knows that he must have information on these elements, as well as others involving other types of inferences, for this information to be complete. More subtly, the victim is usually a person of political importance, the act is performed for impersonal reasons, and the agent generally has no personal relationship with the victim. Thus, an implicit attribute of any individual filling the slot of patient in this template is a political association of some kind: i.e., holder of an office in the government, the opposition, or some revolutionary group. Although the patient or victim is therefore a person who is *known*, the assassin or agent is typically *unknown*. This in turn has implications for intertemplate relations relating to motivation for the act, discussed below.

In addition to explicating the attributes of entities and individuals which can fill

the descriptor slots in a template, determining functional synonyms of the predicate is also a subtask of defining intratemplate relations. Because the predicate *assassinate* has such a specific meaning, synonymous expressions are limited to the derivatives of the verb, or the known 'assassin'.

However, in most cases, predicate slots maybe filled by a variety of expressions, which are functionally synonymous for the concept represented in a given template. For example, all the following expressions are *functional* synonyms for the predicate in a template which might be characterized as and y is the object or institution under direction, in the sense of "general Lara is Director of the War College":

> x holds the directorship of y

> x holds the office of director (of y)

> x is director of y

> x heads y

> x is the head of y

> x manages y

> x is manager of y

> x administrates y

The concept of "functional synonymy" is a pragmatic designation, intended to approximate matching processes in human cognition, whereby items functioning as synonyms in a particular context are perceived as equivalent, although in a different context this may not be the case. Thus, the equivalence of the first three expressions in the preceding list to the second three clearly does not apply in cases where the context of holding the office of director is that of a board of directors of a corporation, since only the director who is also chairman of the board is equivalent to a single head or chief of an organization. Similarly, "manage"

and "administrate" may function as synonyms in a broad context, but not is a specific one.

4.2.1.2 Intertemplate Relations.

Relations obtaining between templates may be obligatory or optional. Obligatory relationships are in general derivable from linguistic and real world knowledge (e.g., a plane cannot land unless it is flying, cannot fly unless it has taken off, etc.). On the other hand, the elaboration of optional relations between templates is based on the analyst's specific knowledge of alternatives in a given situation, ideally with some probabilistic weighting. For example, the crash of an aircraft during training maneuvers in a border area may result in a search and rescue mission, and this may in turn result in a diplomatic incident, and so on.

Similarly, an assassination event may be preceded by a *hiring* event, and if successful, by a *payoff*. Thus, the ultimate agent is not the paid assassin, but the agent involved in the hiring event, or the person or organization that employs him. Although the definition of an exhaustive network of such intertemplate relations is neither feasible nor necessary, the specification of intertemplate relations should be sufficiently detailed to permit determination of the value of such relations as a means of augmenting the analyst's inferential and predictive abilities.

Another type of obligator intertemplate relation involves the class inclusion relations represented by hypernyms (the class terms) and hyponyms (the member terms) of city and event terms. For example, the notion of *assassinate* is included in the notion of (is a hyponym of) *murder*, which is included in the notion of *kill*, which is a hypernym of both. The definition of such relations is necessary to insure adequate recall against a data base of instantiated (filled-in) templates. Thus,

if an analyst wants to know whether any killings occurred in Kinshasa during a specified period, the retrieval should include all the hyponyms of the given term.

4.2.2 Design of Interactive Tutorials for Template Generation

4.2.2.1 Design of Tutorials for Filling in Existing Templates

One set of interactive tutorials and associated support functions is specifically aimed at generating data records from templates which have been defined previously. This set of tutorials mainly consists of presenting displays of previously completed data records based on a particular template type. The relevant template type is accessed by successive menu selections as shown in Figures 4 through 6, assuming the template building activity is triggered by the analyst's examination of his mail file of messages. For example, Figure 4 shows a simulated message, for which the analyst will develop a template in the activity class labeled "air", which he selects with a lightgun (the selection could of course be performed by using a cursor, joy stick or track ball, depending on the type of terminal utilized). He then narrows the selection further by lightgunning the *penetrate* event type, bringing up the display shown in Figure 6. In order to remind himself of the type of data element required for each field of the template, he requests an example of a filled-in *penetrate*, illustrated in Figure 7. Using the interactive tutorial system, the analyst can request additional filled-in templates as examples, or, by pointing the cursor at a descriptor element of the template, can request the description of the data element required to fill the particular descriptor slot.

```
MSGNO.  BAX396

ZCZCECHO97
KFPSYGX RUEWAY
DE RUIJKL #724 3360615
ZNR UUUU ZOX
RUELRW T BRAVO SIERRA ONE
P R 020830Z DEC 70 ZEDD
FM OPEN DOOR
TO AIG930
BRAVO SIERRA ONE
BT
UNCLAS/EXERCISE CUBAN HEEL/SPOPREP SPIFFY/SPITREP 086/

SUBJECT:  PENETRATION REPORT

AIR RECON ACTIVITY BY SOMALI FIGHTER AIRCRAFT OCCURRED NORTH OF ASMARA AT
011600Z.  THE AIRCRAFT PENETRATED ETHIOPIAN AIRSPACE FOR ABOUT FIVE MILES.
WHEN FRIENDLY AIRCRAFT WERE VECTORED NEAR THE AREA OF PENETRATION THE
SOMALI AIRCRAFT REENTERED THE SOMALI REPUBLIC.

BT
#724
NNNN
-------------------------------------------------------------------
TO BEGIN INTERACTIVE EVENT RECORD GENERATION

      LIGHTGUN ACTIVITY CLASS

            * AIR
            * SUBMARINE
            * SHIP
            * TROOP MOVEMENTS
```

Figure 4.   Simulated Display of Initial
            Selection in Template Filing
                   Procedure

MSGNO. BAX396

ZCZCECH097
KFPSYGX RUEWAY
DE RUIJKL #724 3360615
ZNR UUUU ZOX
RUELRW T BRAVO SIERRA ONE
P R 020830Z DEC 70 ZEDD
FM OPEN DOOR
TO AIG930
BRAVO SIERRA ONE
BT
UNCLAS/EXERCISE CUBAN HEEL/SPOPREP SPIFFY/SPITREP 086/

SUBJECT: PENETRATION REPORT

AIR RECON ACTIVITY BY SOMALI FIGHTER AIRCRAFT OCCURRED NORTH OF ASMARA AT
011600Z. THE AIRCRAFT PENETRATED ETHIOPIAN AIRSPACE FOR ABOUT FIVE MILES.
WHEN FRIENDLY AIRCRAFT WERE VECTORED NEAR THE AREA OF PENETRATION THE
SOMALI AIRCRAFT REENTERED THE SOMALI REPUBLIC.

BT
#724
NNNN

ACTIVITY CLASS = AIR
LIGHTGUN EVENT CHAIN/TYPE

•OBSERVE	•REPORT EVENT STATUS	•FLY	•OVERFLY
DETECT	INITIATE	PROCEED	•PENETRATE
TRACK	CONTINUE	TRAVEL	
NOTE	INCREASE		ENTER
SIGHT	SUSPEND	•LAND	VIOLATE
	STOP	CRASH	

Figure 5. Simulated Display of Second
Selection in Template Fill-in
Procedure

A-21

AIR RECON ACTIVITY BY SOMALI FIGHTER AIRCRAFT OCCURRED NORTH OF ASMARA AT
011600Z. THE AIRCRAFT PENETRATED ETHIOPIAN AIRSPACE FOR ABOUT FIVE MILES.
WHEN FRIENDLY AIRCRAFT WERE VECTORED NEAR THE AREA OF PENETRATION THE
SOMALI AIRCRAFT REENTERED THE SOMALI REPUBLIC.

- -

MESSAGE TIME: 02/0830Z DEC 1970
MESSAGE ORIGINATOR: OPEN DOOR
INFORMATION SOURCE:

ACTIVITY CLASS = AIR
EVENT CHAIN = PENETRATE

   AIRCRAFT

      TYPE:
      NUMBER:
      COUNTRY OF ORIGIN:
      BASE OF ORIGIN:
      ALTITUDE:
      SPEED:

GENERAL AREA PENETRATED:
PARTICULAR AREA OF PENETRATION:
TIME OF PENETRATION:
EXTENT OF PENETRATION:
EVALUATION OF PENETRATION:

TO RECORD FURTHER INFORMATION, LIGHTGUN SELECTION

   • ROUTE OF PENETRATION
   • REACTION TO PENETRATION
   • RELATED EVENT

Figure 6. Simulated Display of Event
Template for a Penetration
Event

```
           EVENT RECORD NUMBER 23

MESSAGE TIME = 02/0830Z DEC 1970
MESSAGE ORIGINATOR = OPEN DOOR
INFORMATION SOURCE =

ACTIVITY CLASS = AIR

EVENT CHAIN = PENETRATE

      AIRCRAFT

          TYPE:  FIGHTER
          NUMBER:
          COUNTRY OF ORIGIN:  SOMALIA
          BASE OF ORIGIN:
          ALTITUDE:
          SPEED:

      GENERAL AREA PENETRATED:  ETHIOPIA
      PARTICULAR AREA OF PENETRATION:  NORTH OF ASMARA
      TIME OF PENETRATION:  011600Z DEC 70
      EXTENT OF PENETRATION:  5 MILES
      EVALUATION OF PENETRATION:  RECON
----------------------------------------------------------
  INDICATE PROCEDURE BY LIGHTGUNNING OR BY PRESSING VARIABLE FUNCTION KEY

          •  REDISPLAY EVENT RECORD MENU

          •  PRINT HARD COPY OF EVENT RECORD

          •  DISPLAY RELATED EVENT RECORD FORWARD

          •  DISPLAY RELATED EVENT RECORD BACKWARD
```

Figure 7.   Sample Filled-In Template for Penetration Event

4.2.2.2 Design of Tutorials for Creating New Templates.

A second set of interactive tutorials and associated support functions allows the creation of new template types on line by the analysts. In the verbose mode, these tutorials will not assume that the operator analyst has previous in depth experience with the template generation procedure. The terse mode -- on the other hand -- will present a template skeleton and a skeleton for indicating obligatory and optional relations to other templates (which the analyst may also be required to create), but will assume that the set of prompting and explanatory displays presented on the following pages represents knowledge that the given operator analyst already has.

For the analyst unfamiliar with procedures for creating new templates, the tutorial he receives when he encounters a message requiring a template type not contained in the existing inventory of templates while reading his daily mail is illustrated in Figures 8 through 11. In Figure 8, the analyst is asked to suggest answers for the basic intelligence questions presented at the bottom of the split screen display, relative to the event described in the simulated message shown at the top of the display. (As discussed in Section 2, these questions and answers are effectively the information parameters of a prototype event template.) The next display, Figure 9, asks the analyst to compare his suggested answers (mental or written suggestions) with the set of example answers. In Figure 10, the analyst is presented with a list of expressions (or descriptors) describing the anticipated answers to the set of basic questions. The analyst is then requested to define a similar set of expressions describing the expected answers for the specific event type represented in the message and illustrated in the right most column. Figure

11 presents an example list of descriptors specific to the given event for comparison. Similarly, structures tutorial displays will lead the analyst through the associated task of defining the obligatory and optional relations to other templates.

4.2.2.3 Automated Implementation of Template Technology

The automated understanding of text is a complex undertaking, since a text-based -- rather than a sentence-based grammar is required. Referential and anaphoric elements (e.g., articles, pronouns, appositives, reference by synonym) operate within a larger discourse context and consequently, are more difficult to unravel. Moreover, a formalism for knowledge representation such as Minsky's "frames", Schank's "scripts", Rieger's "Commonsense Algorithms" is necessary to provide a basis for the computer to infer information implicit in the text in order to reduce its content into a set of discrete information records.

The critical factor in determining the feasibility of fully automated data base generation from unformatted message text is the tractability of the message in terms of its format and content structure.

The language of military messages in effect constitutes a restricted subset of the English language, making possible the development of a capability for automated analysis of messages and synthesis of data base elements. Such a development has been carried out by OSI under several RADC contract efforts in support of an Indications and Warning (I&W) data base in the NMIC Advanced Indications System (AIS). The major thrust of these efforts has been directed toward automated analysis of unformatted message text in the subject domains of air and missile/satellite activities.

A-24

```
ZCZCABC019
OO XADOUL
DE XADBEA #0131 1150930
ZNR UUUUU
KNZO OO ABC JKL
O R 252105Z APR 80 ZZO
FM BROKEN DRUM
TO AIG 602
INFO AMEMBASSY NAIROBI
ZEM
U N C L A S
ZZYYKALQ63
2/SX/0990-80 SPOT REPORT FOLLOW UP NR TWO
AND FINAL TO 2/SX/0983-80, 251030Z APR 80
GUG, GKE
VIOLATION OF KENYAN AIRSPACE
YYGG

A UGANDAN MIG-21 PENTRATED KENYAN AIRSPACE AT 0235Z 25 APRIL 1980.  THE
AIRCRAFT ENTERED KENYAN AT A POINT 6 MILES FROM THE UGANDA BORDER NEAR
SUAM ON A PROBABLE RECONNAISSANCE MISSION.  WHEN KENYAN FIGHTERS WERE
SCRAMBLED FROM NAKURU, THE UGANDAN AIRCRAFT REENTERED UGANDA.

061
#0131
NNNN
```

TO CREATE AN EVENT TEMPLATE FOR ANY EVENT TYPE REPRESENTED IN THE MESSAGE,
THINK OF HOW YOU MIGHT ANSWER THE FOLLOWING SET OF BASIC QUESTIONS
RELATIVE TO THAT EVENT:

- WHAT?
- WHO?
- WHEN?
- WHERE?
- TO WHOM?
- WHY?
- WHAT ELSE HAPPENED?

Figure 8.   Simulated Display of First
Tutorial Frame for Template
Creation

A UGANDAN MIG-21 PENETRATED KENYAN AIRSPACE AT 0235Z 25 APRIL 1980. THE
AIRCRAFT ENTERED KENYA AT A POINT 6 MILES FROM THE UGANDA BORDER NEAR
SUAM ON A PROBABLE RECONNAISSANCE MISSION. WHEN KENYAN FIGHTERS WERE
SCRAMBLED FROM NAKURU, THE UGANDAN AIRCRAFT REENTERED UGANDA.

HERE ARE SOME EXAMPLE ANSWERS TO COMPARE WITH YOUR ANSWERS:

	YOUR ANSWERS	EXAMPLES
• WHAT?		airspace violation
• WHO?		a Ugandan MIG-21
• WHEN?		at 0235Z 25 April 1978
• WHERE?		6 miles from the Uganda border near Suam
• TO WHOM?		Kenya
• WHY?		probable reconnaissance mission
• WHAT ELSE HAPPENED?		e_1: scrambling event

Figure 9. Simulated Display of Second
Tutorial Frame in Template
Creation

SUAM ON A PROBABLE RECONNAISSANCE MISSION. WHEN KENYAN FIGHTERS WERE
SCRAMBLED FROM NAKURU, THE UGANDAN AIRCRAFT REENTERED UGANDA.

EACH BASIC INTELLIGENCE QUESTION APPEARING ON THE LEFT ANTICIPATES A
PARTICULAR TYPE OF ANSWER. THESE BASIC ANSWER TYPES CAN BE DESCRIBED
BY THE SET OF DESCRIPTORS SHOWN BELOW IN THE SECOND COLUMN. USING
THESE BASIC OR GENERAL DESCRIPTORS, HOW MIGHT THE THIRD COLUMN BE FILLED
IN WITH A SET OF SPECIFIC DESCRIPTORS APPLYING ONLY TO THE EVENT TYPE
REPRESENTED IN THE MESSAGE (i.e., AN AIRSPACE VIOLATION)?

BASIC QUESTIONS	BASIC DESCRIPTORS	DESCRIPTORS FOR EVENT TYPE e	COMPLETED EXAMPLE FOR EVENT TYPE e
WHAT	event type		airspace violation
WHO	agent (or object plus owner)		a Ugandan MIG-21
WHEN	time of the event		at 02˜5Z 25 April 1978
WHERE	location at which event took place		6 miles from the Uganda border near Suam
TO WHOM	patient, or entity affected by the event		Kenya
WHY	interpretation of the event		probable reconnaissance mission
WHAT ELSE HAPPENED	indication of relation to other event; nature of relation e.g., causal)		e_1: scrambling event

Figure 10. Simulated Display of Third
Tutorial Frame in Template
Creation

A UGANDAN MIG-21 PENETRATED KENYAN AIRSPACE AT 0235Z 25 APRIL 1980. THE
AIRCRAFT ENTERED KENYA AY A POINT 6 MILES FROM THE UGANDA BORDER NEAR
SUAM ON A PROBABLE RECONNAISSANCE MISSION. WHEN KENYAN FIGHTERS WERE
SCRAMBLED FROM NAKURU, THE UGANDAN AIRCRAFT REENTERED UGANDA.

ARE YOUR SUGGESTED DESCRIPTORS SIMILAR TO THE SET OF SPECIFIC DESCRIPTORS
SHOWN IN COLUMN 3? (USE 'PREVIOUS DISPLAY' FUNCTION TO COMPARE).

BASIC QUESTIONS	BASIC DESCRIPTORS	DESCRIPTORS FOR EVENT TYPE e	COMPLETED EXAMPLE FOR EVENT TYPE e
WHAT	event type	airspace violation	airspace violation
WHO	agent (or object plus owner)	aircraft owned country C	a Ugandan MIG-21
WHEN	time of the event	specific time at which the violation occurred	at 0235Z 25 April 1978
WHERE	location at which event took place	specific location at which the violation occurred	6 miles from the Uganda border near Suam
TO WHOM	patient, or entity affected by the event	owner of the violated airspace	Kenya
WHY	interpretation of the event	probable reason for the violation event	probable reconnaissance mission
WHAT ELSE HAPPENED	indication of relation to other event; nature of relation e.g., causal)	related event: e_1 relation: e cause e_1	e_1: scrambling event

Figure 11. Simulated Display of Fourth
Tutorial Frame in Template
Creation

These efforts are summarized in the following paragraphs.

4.2.2.4 I & W Feasibility Study.

In an initial I&W-oriented exploratory effort sponsored by RADC, I&W 1,) (Contract F30602-76-C-0194), OSI developed a system concept for automated data base generation by drawing upon its own experience in the field (e.g., the ERGO technology developed under a previous RADC contract) and taking into account recent research results. This development effort laid the theoretical and methodological foundation for achieving the ultimate goal of automated data base generation and update based upon a computer "understanding" of natural language text.

In particular, the I&W I project effort was devoted to studying the feasibility of automatically generating I&W data files from electrical messages of all types-- formatted, unformatted, and semi-formatted. This study yielded a comprehensive statement of the problem involved in the analysis of message texts and the synthesis of data base elements. It defined the linguistic methodology required for achieving automated data base generation from un-formatted text, and developed a design concept for automated analysis of formatted fields from a class of messages utilized to update the statistical data elements of the NMIC Advanced Indications System (AIS) data base (RADC-TR-77-194, Vol. I & II.

4.2.2.5 I & W Interactive Testbed Implementation.

I & W resulted in MATRES, an interactive capability for data base generation incorporating a subset of the features requisite for a fully automated system to serve as a Testbed for the progressive elaboration of system components. This version of the testbed utilizes a simple Augmented Transition Network (ATN) sentence acceptor, which

performs a shallow linguistic analysis and produces event representations in the form of formatted records. A rudimentary capability for storage and retrieval was also developed.

Sample output from this version of MATRES is shown in Figure 12.

4.2.3 I & W III.

A Knowledge-Based Automated Message Understanding Methodology for an Advanced Indications System.

Work on a system with full capabilities was initiated under RADC Contract No. F30602-77-C-0144 (I&W III). This was an exploratory developmental effort which addressed the construction of algorithms to expand and augment the capabilities of the interactive system. Essentially, it "reads" and "understands" a message text to the degree necessary to transform each constituent sentence describing an event into a formatted data base record. I & W III involves the definition of an Event Representation Language (ERL), a formal language for the representation of knowledge about events, and the development of a mechanism for the automated transformation of natural language text of AIS-related intelligence messages into a formal content representation. This content representation is the basis for synthesizing indicator and descriptor structures appropriate for AIS.

4.2.3.1 I & W IV. Satellite and Missile Data Base Generation for AIS.

Under RADC Contract No. F30602-78-C-0274, a new subject domain has been added to the MATRES repertoire, and new capabilities which allow analysis of the message into detailed components have been introduced.

Figure 13 shows an output from the MATRES II system, and Figure 14 presents a graphical comparison of the analysis and resulting outputs from MATRES I and II.

INPUT

>> EARLY ON 21 FEBRUARY TWO UNIDENTIFIED AIRCRAFT CONDUCTED A
SURFACE-TO-AIR RECONNAISSANCE MISSION OVER THE GULF OF AQABA.

OUTPUT

EVENT RECORD CONTENTS FOR SENTENCE:
EVENT: ACTIVITY
MISSION: A SURFACE-TO-AIR RECONNAISSANCE MISSION
VERB: CONDUCTED
ACTOR: TWO UNIDENTIFIED AIRCRAFT
REGION: OVER THE GULF OF AQABA
TIME: EARLY ON 21 FEBRUARY

Figure 12.   Input and Output of MATRES I

```
+--------------------------------------------------------------+
|                                                              |
|                           Input                              |
|                                                              |
|                                                              |
| TWO UGANDAN AIRCRAFT FROM REGIMENT A1313 AT ENTEBBE          |
| DEPLOYED TO GULU AT 0200Z ON 21 FEBRUARY.                    |
|                                                              |
+--------------------------------------------------------------+


+--------------------------------------------------------------+
|                                                              |
|                          Output                              |
|                                                              |
|     Event: DEPLOY                                            |
|     Object:                                                  |
|     ...Equipment= UGANDAN ACFT                               |
|     ...Nationality= UGANDAN                                  |
|     ...Subordination= FROM REGIMENT A313                     |
|     ...Stagingbase= AT ENTEBBE                               |
|     ...Number= TWO                                           |
|     Destination= TO GULU                                     |
|     Time= AT 0200Z                                           |
|     Date= ON 21 FEBRUARY                                     |
|     EVENT RECORD COMPLETE.                                   |
|                                                              |
+--------------------------------------------------------------+
```

Figure 13.   Example Input and Output by MATRES II

Figure 14.   MATRES I vs. MATRES II   Data Base Generation

5. EXTENSIONS TO THE TEMPLATE STRUCTURE

The template can be viewed as a fundamental information structure that has a life cycle consisting of the following phases:

[1] Expectations

[2] Information Collection

[3] First Phase Exploitation

[4] Second Phase Exploitation

[5] Third Phase Exploitation

By looking at the life cycle of information, it is possible to explain many of the different perspectives that individual users have of the information and why so many different data formats persist for structuring this information.

In addition to the information parameters of the prototype event template shown in Table 1, there are five descriptor types which are necessary for an information processing implementation in a tactical C^3 environment. These five additional descriptor categories for the template structure are:

[1] Significance/Threat relation

[2] Detection Source

[3] Information Source

[4] Information disposition

[5] Value added

The following paragraphs describe these additional descriptor types and indicate their function in the TAC C^3 environment.

5.1 Significance/Threat-related Descriptors

The importance of intelligence information is directly attributable to the threat relevance of the particular information item. An individual message carrying threat-related data may have more or less information value depending on whether the threat is new, has experienced a significant change from a previous evaluation, or has significantly changed in the immediacy of the threat to the information recipient.

Depending on whether this type of information is being handled in a planning phase, operations management, or indications and warning context, the significance to the user will vary.

Operations data also is characterized by its significance to the recipient. Weather data regularly reported takes on more significance when a major change is identified. Personnel status is significant when it affects readiness of units.

In order for significance to be used in the handling of information within the data processing system, it must be specified in the template structure.

Current message formats have implicit significance by their type code (SPOTREP, HOTPHOTOREP, TACREP, etc.) or by precedence codes (FLASH, CRITIC, OPS IMMEDIATE, PRIORITY, ROUTINE). However, these codes are meaningful only on initial transmission; they become ambiguous to a computer in follow up handling and use in long term planning processes. This is primarily a problem attributed to having significance assigned from the perspective of the information sender only.

5.2 Detection/Information Source Descriptors

Characteristics of the information source have specific impacts on how information is handled and used in the information system. In evaluating sources -- which tends to be a vaguely defined term -- the distinction between the perceiver or detector of the primitive event (a sensor of some time, or a human sensing agent) and the reporter of the information must be distinguished, as this has a bearing on the credibility of the information. Thus, in

A-33

the case of a human observer, the detector and reporter of the event may be the same individual, but this is rarely the case. Often, an observed event is embedded in a report of a report of an event, and the possibilities for introduction of error or additional information at each point in the transformation must be borne in mind.

In a tactical environment, identification of the perceiver or detector of the event along with the time of transmission is normally sufficient to uniquely identify the information item. Based on the collection plan, identification of the observer or detector type (which may be either a mechanical sensor such as a radar or camera, or a human observer) is sufficient to qualify the information in terms of field of view and data quality.

Detection sources are obviously different between SIGINT, IMINT, and HUMINT, as mentioned above, although information or reporting sources may be the same. SIGINT detection sources identify the frequency, mode, target, and collecting station to provide continuity in reports. IMINT platforms are identified by equipment type, mode, mission, frame reference, and platform position. Field of view, period of observation, and scaling parameters are derivable from these platform descriptors. Data quality may vary on a frame by frame basis because of cloud cover, topographic masking, slant range, lighting, sensor performance, or other factors affecting interpretability of the images. Strategic facilities use a quality rating system called NIIRS (Numerical Imagery Interpretability Rating System) to provide a standardized and composite evaluation of resolution and imagery quality. No rating system currently exists for tactical imagery.

HUMINT sources will include IPW reports, combat information, state messages, and special forces reports. The source is normally identified and/or evaluated as an information item in the

format of messages carrying the data. The quality of the data will be a combination of the primitive source of the data plus qualifications added by the reporting agent. The receiver may also qualify the accuracy of the information in terms of its context in his application or because of past experiences in using data from a particular source.

In some instances, identification of the analyst that produced the intelligence product is presented for providing a directory to users requiring followup information or clarification. If information is controlled by a command directive or is under access control restrictions, then a release authority will also be contained in the source identification.

5.3 Disposition Descriptors

The handling of data in hardcopy form is controlled by a combination of flags/codes and physical location (such as in a file cabinet, mail slot, read board, distribution pile, or briefing folder. For computer control, the disposition of data must be explicit in terms of its physical electrical storage location and its logical disposition.

Disposition descriptors fall into the general categories of:

[1] Current state (hypothetical/definition/norm, edit, review, released for output, mail queue, reviewed & acknowledged)

[2] Distribution list (action, information, routed to)

[3] Access control (transmission control, special access, classification, downgrading data)

[4] Perishability (useful life of data, purge date, archive identification)

[5] Audit trail (unique identification of this record, references to master/previous/followup

messages)

The disposition descriptors are necessary in the computerized control of data routing, data storage, and data display.

5.4 Value Added Descriptors

Data which is correlated, fused, filtered, or otherwise categorized by human or computer processes gains value over its former state. These added descriptors may take the form of pointers, physical entry in clusters or files,or addition of descriptor tags to the data record. Set descriptions (queries, report generators,and aggregation algorithms) which may act on the data and increase a user's knowledge can be considered to perform a value-added function if some permanent data remains to tie the record to the process. A hit file for a query or a transaction list for an aggregation computation could be considered to be value-added descriptors.

General categories of value-added descriptors are:

[1] Fusion (across source, mode, time, location)

[2] Correlation (time, activity, space, synonyms)

[3] Clustering (categorization, time, space, organization)

[4] Aggregation (numerical, organization, activity)

[5] Display selectivity/intrarecord ranking (display format control, highlighting, data value significance, data item context, annotation)

[6] Interrecord ranking (precedence, fifo, lifo, relevance ranking)

5.5 Impact of Self-defining Data Structures

The purpose of enumerating the different aspects of value-added descriptors is to show the extent of explication which must exist in a self-defining information structure, such as a template. If self-defining data structures are the key to decentralization of data processing control, then significant attention must be given to implications in software development, the operating system, and technology support requirements. Without centralized control, two cooperating processes which share common data cannot depend on a synchronizing mechanism to ensure the compatibility of the data structure at any point in time.

Software development will be impacted by the need to determine if any information descriptors have changed that affect the shared use of the information. Currently, most systems assume a highly static data structure environment. New software development along the lines of self-defining data structures must consider that front end processing, conversion, and exception case handling will be a normal part of the operating environment.

Because data structures must be self-defining, there must be careful attention paid to the semantics of descriptors as well as primitive content of the data record. Linguistic analysis tools will be required to aid the system designer in developing vocabulary, syntax rules, and disambiguation devices to process the self-defining information structures.

The design of software to achieve efficiency in the self-defining information structure concept will place a high premium on the effective use of time in the processing sequence. Mechanisms such as compilers, initializing processes, functional subsetting, and context dependent user languages can use processing time more effectively by converting the software into a more efficient, run-time form, at the cost of preprocessing.

MISSION
of
Rome Air Development Center

RADC plans and executes research, development, test and selected acquisition programs in support of Command, Control Communications and Intelligence (C³I) activities. Technical and engineering support within areas of technical competence is provided to ESD Program Offices (POs) and other ESD elements. The principal technical mission areas are communications, electromagnetic guidance and control, surveillance of ground and aerospace objects, intelligence data collection and handling, information system technology, ionospheric propagation, solid state sciences, microwave physics and electronic reliability, maintainability and compatibility.