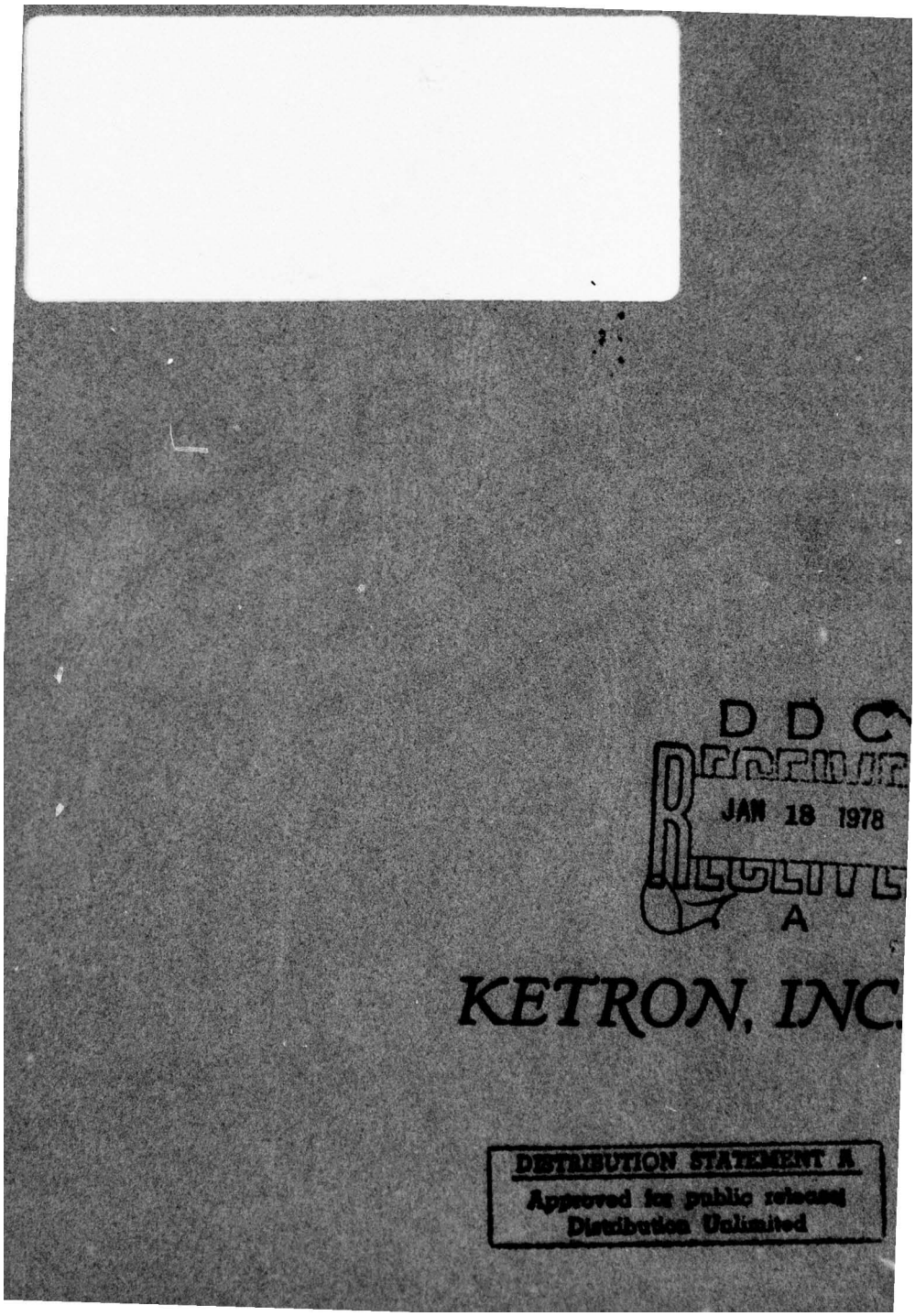
AD-A048 763 KETRON INC ARLINGTON VA F/G 5/1
VP DATA BASE SYSTEM.(U)
SEP 77 S S SHENTON N00014-77-C-0145
UNCLASSIFIED KFR-131-77 NL

1 OF 1
ADA048 763

END
DATE FILMED
2-78
DDC



KETRON, INC.

(1)

VP DATA BASE SYSTEM:
FINAL REPORT

(CONTRACT N00014-77-C-0145)

30 September 1977

KFR 131-77

Submitted to:

Office of Naval Research
Code 230
800 North Quincy Street
Arlington, Virginia 22217

Attn: Mr. Robert J. Miller

Submitted by:

Ketron, Inc.
1400 Wilson Boulevard
Arlington, Virginia 22209

Attn: Mr. Ervin Kapos/
Executive Vice President

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE	READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER [redacted]	2. GOVT ACCESSION NO. / 3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) (6) VP DATA BASE SYSTEM: FINAL REPORT	5. TYPE OF REPORT & PERIOD COVERED (9) Final Report. 1 Nov 1976 - 31 Jul 1977
	6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) (19) Scott S. Shenton	8. CONTRACT OR GRANT NUMBER(s) (15) N00014-77-C-0145
9. PERFORMING ORGANIZATION NAME AND ADDRESS Ketron, Inc. 1400 Wilson Blvd. Arlington, Virginia 22209	10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS NR 128-019
11. CONTROLLING OFFICE NAME AND ADDRESS Scientific Officer, Fleet Analysis and Support Division, Office of Naval Research, 800 North Quincy Street, Arlington, Virginia 22217	(11) 12. REPORT DATE 30 September 1977 / 13. NUMBER OF PAGES 23
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office) (14) KFR-131-77 (12) 29p.	15. SECURITY CLASS. (of this report) UNCLASSIFIED / 15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

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

Approved for public release; distribution unlimited.

18. SUPPLEMENTARY NOTES

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

DATA BASE; VP; PATROL AIRCRAFT; COMPATWINGSLANT; RAINFORM; PURPLE messages

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

This report briefly describes a prototype VP Data Base System (VPDBS) developed for COMPATWINGSLANT, and recommends that the system be installed for regular use. Alternative host computer sites are evaluated; and a dual-site implementation is recommended. Requirements and limitations of a much larger and more complex Advanced VPDBS are examined.

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

S/N 0102-LF-014-6601

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

390 861

LB

TABLE OF CONTENTS

TABLE OF CONTENTS, Continued ....

REFERENCES

A. VP Data Base System: System Documentation, KTR 36-77, 21 September 1977

B. VP Data Base System: Input Data Analysis and Sample Reports, KTR 37-77, 23 September 1977

C. Patrol Plane Operations Data Base Feasibility Study, KFR 77-76, 20 August 1976

I. INTRODUCTION

One of the functions of management is the monitoring and control of subordinate activities to ensure that management goals and objectives are successfully met. Depending on the nature of the organization, these goals may be in terms of dollars, profits, number of widgets manufactured, or number of lives saved. For operational military commands, the goals are frequently "success rates" in the performance of assigned tactical missions against a potential enemy.

To monitor and control performance, the manager or commander needs to have access to information on current performance, and a means to indicate whether observed performance is satisfactory or substandard. To effectively control performance, the system of measures should be able to identify weaknesses or substandard areas of performance and, ideally, should be able to zero in on suspect areas and further pinpoint the source(s) of lower performance, to facilitate taking corrective action.

BACKGROUND

The Commander, Patrol Wings Atlantic Fleet (COMPATWINGSLANT) is responsible for the performance of twelve squadrons of P3 ORION VP ASW aircraft. Some facets and indications of squadron and individual aircrew performance may be determined during training and nonoperational flight activity, but true performance can only be assessed under operational conditions. While COMPATWINGSLANT does not exercise operational control over deployed VP squadrons, overall levels of performance of the squadrons must be measured on what they do, and how well they do it, when deployed.

There are three major sources of information on operational VP flight:

- post-mission message summaries;
- data collected aboard the aircraft during the flight; and
- additional data developed at the Tactical Support Center (TSC) during post-flight assessment.

The latter two data sources may produce voluminous amounts of both raw data and evaluated results. The post-mission summary contains primarily evaluated and confirmed information, in a highly condensed, formatted manner.

Reference (C) reported the results of a feasibility study of the types of Measures of Effectiveness (MOEs) COMPATWINGSLANT could use, the potential data sources to support the MOEs, and the basic requirements of a VP Data Base (VPDB) System to collect and retain the data. The study recommended that a subset of measures and data be selected, and a prototype data base system be developed to test out the concept, and serve as a vehicle for assessing the suitability of a greatly expanded system. This report describes the nature of the prototype VPDBS developed, evaluates the future of the prototype system, and draws some conclusions on the need for an expanded or advanced system.

PROTOTYPE VPDB SYSTEM

The VP Data Base System (VPDBS), as presently constituted, was developed as a limited-purpose prototype system, with the objective of testing and demonstrating the feasibility of computer processing of RAINFORM mission summary (PURPLE) messages. The specific application chosen for the prototype was the automatic calculation of the COMPATWINGSLANT-developed VP Performance Indicators (PIs). The system is completely described in reference (A).

The VPDBS consists of a set of three computer programs which, combined with specified manual procedures and human decision-inputs, reads in actual PURPLE messages from magnetic computer tape, extracts and computes over fifty data items or values which relate to the Performance Indicators, and constructs a single, small data record for each PURPLE message. The new data records are then retained as a permanent data base of information from which the PIs may be calculated. A fourth program then allows the user to access this data base to calculate the PIs based on a variety of user-selected keys. The net result is a system which can replace the present procedure of manual message reading and PI calculation, while offering the following advantages:

- reduced manual (analyst) time;

- permanent retention of data in the data base;
- greatly expanded analysis capabilities; and
- accurate, repeatable results free from human errors.

These advantages are available with the use of surprisingly little computer time -- and therefore costs -- so the overall result of installation of the VPDBS will be greater capability for less cost to COMPATWINGSLANT. The system was developed as a prototype, and a Washington-area computer was used to maximize efficiency in the development. The feasibility and performance of the system have now been demonstrated, and implementation of a production oriented system for regular use by COMPATWINGSLANT personnel is recommended.

II. FUTURE OF THE PROTOTYPE SYSTEM

Given that the development of the prototype system has been successful, and its acceptance and implementation as a regular "production" system has been recommended, it is next necessary to determine where best to implement the system. The present host computer is ideally located for developmental work in Washington, but is unacceptably far removed from the COMPATWINGSLANT Brunswick facilities, by both proximity and accessibility. The best available host site for the system must be selected, based on an evaluation of all pertinent selection criteria.

SYSTEM EVALUATION FACTORS

There are a series of criteria for evaluating alternative sites (i.e., computer systems) to host the VPDBS. Some of the criteria are absolute -- either acceptable or unacceptable -- while others are more subjective and offer a range of good/better/worse type of values. These criteria, and possible host computer systems, were previously evaluated for general applications and reported in reference (C). For the specific needs of the existing prototype VPDBS, the criteria and possible systems were re-evaluated and new information was incorporated. The selection criteria are:

- capability to accommodate data of appropriate security level;
- available time to perform VPDBS tasks;
- adequate size of computer system;
- compatible programming language capability;
- magnetic tape drive availability; and
- accessibility or proximity for use.

A data base management system was an additional criterion in the earlier, and more general, evaluation (reference (C), but does not apply to the existing VPDBS.

There are three computer systems which were examined in detail as candidate

VPDBS host systems, using the above six criteria. The first is the military CP-901 computer in the Brunswick TSC, which is a prime candidate due to its close proximity, security provisions, and apparent time availability. Next is the old IBM-1401 at the Brunswick Naval Air Station, which offers proximity and virtually unlimited availability. Finally, the general purpose CDC-6600 at the Naval Air Development Center which, besides offering the most extensive language capability and other facilities, is already being used by COMPATWINGSLANT for analytical studies.

There is another obvious alternative, which is the acquisition, by purchase or other means, of a new computer system for COMPATWINGSLANT use. Such a system cannot be realistically evaluated here, given the potentially unlimited combinations of possible capabilities and cost; a system big enough to do the job would be ideal, but probably prohibitively expensive unless justified by a large number of other applications. If a medium-to-large scale, general purpose system should become available in the Brunswick area, it would obviously be a prime candidate.

Table 1 shows the results of the evaluation, where "OK" indicates acceptable, "NO" means unacceptable, and "NO?" means probably unacceptable but there may be mitigating circumstances. The single "OK/NO" entry is a special case which is described in full below.

TABLE 1
ALTERNATIVE SYSTEM EVALUATION FACTORS

	Security	Availability	Size	Language	Mag Tapes	Accessibility
Brunswick CP-901	OK	NO?	OK	NO	OK	OK
Brunswick IBM-1401	OK	OK	NO	OK	OK	OK
NADC CDC-6600	OK/NO	OK	OK	OK	OK	NO?

The table shows that there is no standout, obvious "winner"; each system has one or more deficiencies or drawbacks. Before making a selection, each of the six criteria are discussed below, to explain how each influences an eventual selection and how the three systems measure up.

Security

Both the CP-901 and the IBM-1401 can readily process classified data up to the levels necessary for the VPDBS. So can the CDC-6600, but its location in Warminster, Pennsylvania makes it difficult to use from Brunswick. As noted in reference (C), the four ways to get around the Warminster/Brunswick security problem are:

(1) A crypto-covered telephone line -- too expensive for this application; feasible only if other users were to share the need and cost;

(2) CPWL personnel carry data to Warminster, do all work there, and return with results -- does not sound either practical or cost-effective;

(3) Local NADC personnel perform all work -- takes control out of CPWL hands; impractical;

(4) Make maximum use of unclassified data transmission on existing open phone lines, and mail classified portions -- might be unwieldy and would require well-defined procedures.

Overall: a minus for CDC-6600, but not completely unacceptable.

Availability

The IBM-1401 has almost unlimited time available, and the CDC-6600 has more than enough time for the rather small needs of the VPDBS. Availability of the CP-901 is somewhat conjectural. There are large blocks of time when the CP-901 is not in operational use at present, but these frequently occur at odd hours and there is a possibility of being preempted. Use of the 901 for VPDBS would require de-loading of the regular operational system, loading of VPDBS work, and then time-consuming reloading and testing of the operational system to restore to normal use. Additionally, the advent of ASWCCS and TSCIXS may require the TSC/CP-901 to be a round-the-clock, real-time system, precluding extraneous work such as VPDBS.

Overall: a solid minus for CP-901 and, in the future at least, possibly unacceptable.

Size

The CDC-6600 and CP-901 are both completely acceptable in terms of size; the IBM-1401, with only 16K bytes (16,000 characters) of core is not. The IBM-1401 could be used for some of the smaller processing tasks, and other tasks might be "overlayed" into modules or even cut down to run on the 1401. (The 1401 cannot be expanded beyond the 16K size, although some additional capabilities might be added by the use of disk drives -- the 1301 drive (28-56 Mbytes), 1311 (10-15 Mbytes) and 1405 (20 Mbytes) are no longer offered by IBM but might be available through government surplus.)

Overall: a "NO" for the IBM-1401, for the VPDBS as a whole, but some partial use might be made of it.

Computer Language

The CDC-6600 has all the language capabilities necessary or desired, which includes a sophisticated FORTRAN; the IBM-1401 now has a FORTRAN compiler of much lower sophistication, but acceptable. The CP-901, not being a general purpose system but one designed for specialized applications, does not have FORTRAN, COBOL or any common-use high order language (HOL). This would be a severe drawback to both implementation and use of the VPDBS or any other similar software system. The prototype VPDBS is written completely in FORTRAN, and conversion to another language would be onerous. Use of a non-HOL language would impose unacceptable constraints on initial conversion, future modifications and ultimate system use, with the need to have specially-trained programmers for the system or be dependent on contractor support.

Overall: a solid, and probably prohibitive "NO" for the CP-901.

Magnetic Tapes

The most practical means of receiving and processing new PURPLE message data is by magnetic computer tape; additionally, the final data base records must be stored on either tapes or, depending on availability and cost-effectiveness, disk packs. The availability of a tape capability is therefore mandatory. All three candidate systems have acceptable tape drive facilities.

Overall: all three candidate systems acceptable.

Accessibility

Accessibility reflects the ease of access to the VPDBS and the data, both during the input processing and the use of the final data base records for answering ad hoc queries. Both the CP-901 and the IBM-1401 offer excellent accessibility; the CDC-6600 at NADC is accessible via dial-up telephone lines except for the problem of security of classified data. Classified data cannot be transmitted over the dial-up lines, and virtually all data inputs and program outputs would be classified. As mentioned in the discussion of Security, COMPATWING-SLANT use of the NADC computer would therefore require extensive use of mail service to and from NADC, or on-site trips to Warminster. Either procedure would tend to be rather time-consuming and inconvenient for regular use, and inhibit any quick response usage of the VPDBS capabilities.

Overall: a solid minus, if not a final "NO", for the CDC-6600.

EVALUATION RESULTS

A second look at Table 1, in the light of the detailed examination of the evaluation factors, still shows no clear-cut, obvious "winner" between the three candidate systems. There are apparent losers, however. The CP-901 seems to be fairly solidly ruled out, primarily on the basis of the lack of a generalized programming language, while future availability weighs strongly against its use. Similarly, the small size of the IBM-1401 precludes its use as a host for the VPDBS. For the CDC-6600, the major drawback is obviously its location at a substantial, and inconvenient, distance from Brunswick. This deficiency is further aggravated by the inability to use dial-up telephone lines for access because of security considerations. The CDC-6600 at Warminster is thus, if not the best of the candidate systems, the "least worst" of the three. And where the other two are, for all intents and purposes, unacceptable, the 6600 is acceptable but extremely inconvenient.

Unfortunately, the 6600, as a host system for the VPDBSs, is considerably far removed from the ideal system. While rated as acceptable, consideration of

the NADC 6600 as workable is debatable -- this is purely a subjective appraisal which must be made by the ultimate user, answering the question of, "Are the benefits of the VPDBS worth the inconvenience of its use?"

OTHER ALTERNATIVES

Considering the inconveniences of the CDC-6600 as the VPDBS host system, further review of alternatives seems appropriate. The major drawback of the NADC computer is its remoteness from Brunswick. COMPATWINGSLANT presently uses the computer by sending jobs/programs and receiving outputs on a terminal over conventional telephone or AUTOVON lines. Installation of a regular leased line between Brunswick and Warminster, with cryptographic cover to permit transmission of classified data, would almost completely eliminate any objections to the system. The costs of such a line would likely run to several thousands of dollars a month, and thus not be very cost-effective. However, should other uses of such secure lines exist, the combined usage might well justify the cost of a dedicated, secure line.

Up to this point, we have been considering each system as a potential single source host, as if all of the VPDBS functions must reside on a single computer. and that computer must be able to accommodate all parts of the entire VPDBS. While this is the conventional and most straightforward approach, the inconvenience of the "best" single-system alternative and the availability of a local system with at least some useful capabilities indicates some merit in splitting the functions of the VPDBS between two systems, to make the best use of the advantages of each and arrive at the system which is the overall best for COMPATWINGSLANT use.

The VPDBS functions were therefore broken down into their separate components and re-assessed against the different capabilities and advantages of the CDC-6600 and the IBM-1401. The CP-901 was excluded from consideration because of the language barrier as well as the questionable availability. To discuss a component-level evaluation of VPDBS alternatives, some background on the components themselves is necessary.

VPDBS COMPONENTS

The computer programs and manual processing steps which make up the VP Data Base System have been defined in detail in reference (A). Without repeating the detailed documentation, the steps are briefly described below.

1. Receive magnetic computer tapes containing one or more months of PURPLE messages.

2. Using program DECONVRT, convert tapes into a new format more suitable for processing.

3. Using program PURPLE, review each of the PURPLE messages for applicability, then for data errors, omission and conflicts. During this step, each of the messages may be printed for future reference.

4. Manually review the PURPLE-program output, and identify corrections to be made to the individual messages to make them complete and internally consistent. The analyst may need to refer to other PURPLE messages or to non-PURPLE data to resolve ambiguous points.

5. Prepare the changes, on either punched cards or magnetic tape, to be made to the original data.

6. Using program PCHANG, create a new message data tape by adding the changes to the original message tape.

7. Using program PURPLE, process the new corrected messages for undetected/uncorrected errors as in Step 3. Steps 4 through 7 may need to be repeated until all significant errors are corrected.

8. Using program PURPLE, extract the necessary data from each message and create a new data base record for each one.

9. Merge the new data records with those from previous months, to form the complete data base.

10. Using the data base and program DELRAP, access the data base and produce reports desired by the analysts and other users.

These ten steps, or at least those which are performed on a computer (2,3,6,7, 8,9,10), may all be done on the NADC CDC-6600 computer. However, some of them could alternatively be done on the nearby IBM-1401, and some steps which cannot be done on the 1401 could be tailored so that the 6600 outputs are unclassified and can be returned to Brunswick over the dial-up telephone lines. The revised procedural steps are described in the following section.

DUAL-SYSTEM PROCEDURES

The basic procedural steps identified above are expanded on here, with particular reference to the steps used by the two different computer systems and the points at which the two systems exchange data or connect via manual procedures.

Step 1. Data is received on magnetic tape. NADC presently is getting the PURPLE messages on tape for its own uses; making the data available for COMPATWINGSLANT use on the NADC computer should be a simple administrative matter.

Step 2. Conversion of the data into a more suitable format is accomplished by DECONVERT. The program can be transmitted from Brunswick to Warminster by the Brunswick terminal over dial-up lines; unclassified run output can be received at Brunswick over the same line.

Step 3. Program PURPLE is run for the first data examination. The program is entered on the CDC-6600 computer via the dial-up line. A brief, unclassified output, consisting of a summary of number of messages processed and an indication of whether the job ran successfully or not, is returned to the Brunswick terminal. An extensive, classified output is printed at NADC, and mailed to Brunswick; this output contains a complete listing of all PURPLE messages on the data tape, and all errors identified in each of the messages. This output will become the analyst's master reference listing for all subsequent processing.

Step 4. The COMPATWINGSLANT analyst reviews the master listing, identifying those missions/messages to be deleted from the data file and determining what additions or corrections must be made to each message to clear up

any reported (or unreported) errors. The corrections should be annotated on the master listing for future reference.

Step 5. The changes decided upon by the analyst must be coded into a computer-readable form, starting with punched cards. A portion of program PCHANG should be installed on the IBM-1401 to allow a checking step of the change data before it is sent to Warminster -- this would help to eliminate some of the errors which can be introduced by improper key punching, changes not entered or entered improperly, etc. These changes will then be sent to NADC. For ease of mailing, the deck of change cards would be converted to magnetic tape on the IBM-1401, and the classified tape mailed to Warminster. While some of the change cards will be unclassified, and could therefore be sent directly to the CDC-6600 via phone lines, classified cards or card-records on tape must be mailed.

Step 6. Program PCHANG is entered to the CDC-6600 from the Brunswick terminal, to add the change data to the original message data and produce the updated message data tape. An unclassified output, detailing the results and any program failures during the run, would be returned to the Brunswick terminal.

Step 7. Program PURPLE is entered into the CDC-6600 from the Brunswick terminal for a second (or subsequent) check for errors and inconsistencies. This is similar to the run of PURPLE in Step 3, except that there should be significantly less errors detected this time. Also, in Step 3, a classified output containing a full listing of each message was made; in this Step a follow-up listing is unnecessary, and error messages may be keyed to the text so that they are completely unclassified and can be returned directly to Brunswick over the terminal. Results of this run will indicate either a return to Step 4, or continuation on to Step 8.

Step 8. Once the PURPLE runs are error-free, or at least free of significant errors, program PURPLE is entered into the CDC-6600 via the Brunswick terminal. The program takes the most recent version of the message data tape, and produces an unclassified output, directed back to the Brunswick terminal, which is a summary of all missions for which data records have been constructed.

The other output of the program run is a magnetic tape (or a disk file) containing the classified data records just created.

Step 9. At this point the new data records must be combined with data records for previous periods to make a data base containing the records of all missions to date. The data base may be maintained in Warminster or Brunswick. If the final data base is to be kept in Brunswick, then the tape containing the new records from this run must be mailed to Brunswick. There they will be merged with the older records, using the IBM-1401, to form the complete data base.

Step 10. Program DELRAP may be run as often as desired on the IBM-1401 to produce the Performance Indicator results keyed in whichever manner the COMPATWINGSLANT analysts desire. From this point on, all processing of the data can be performed on the Brunswick 1401. The NADC CDC-6600 will not be needed until a new set of monthly (or quarterly) PURPLE messages are to be processed for inclusion in the data base.

EVALUATION OF A DUAL-SYSTEM

The schema described above takes into account the advantages of both the IBM-1401 and the CDC-6600 -- the proximity and ease of access of the former, and the size and processing capabilities of the latter. The Dual-System implementation suggested above does appear to be somewhat unwieldy to use, but, considering the alternatives of no VPDBS or doing all work at Warminster, it is definitely the preferred option. The key issues regarding suitability of the system are:

- the mailing volume and inconvenience;
- the processing delays;
- the location of the final data base; and
- the actual capabilities of the IBM-1401.

The inconvenience of mailing classified material back and forth to Warminster is reduced (compared to doing <u>all</u> processing at Warminster), but not completely

avoided. By the use of the terminal access to the 6600 and the local availability of the 1401, use of the mails is reduced to the following requirements during the construction of the final data base records:

- one-time mailing of large printout from CDC-6600 to Brunswick (Step 3);
- mailing of data corrections on magnetic tape from IBM-1401 to Warminster, possibly several times (Step 5);
- one-time mailing of completed data records from the CDC-6600 to Brunswick (Step 9).

Additional classified outputs of programs run on the 6600 could also be mailed to Brunswick if desired by the analyst, but the mailings noted above are the minimum required. The transmission of programs from Brunswick to Warminster could also be reduced, as the programs themselves could be stored in the computer and merely called up for execution, and provision of necessary control cards, via the Brunswick terminal. Note that all the processing steps apply to each new set of input messages, and are the same whether the inputs are received monthly, quarterly, or with some other frequency. The total volume of mail processing would thus be greatly reduced if data were processed bi-monthly or quarterly rather than monthly.

Once processing of a set of new input data has begun, the ideal situation would be to follow through all the processing steps and complete the process in a few days. This is not practical for two reasons:

- Some significant portion of time (on the order of an estimated two days for each month of data) is required for the analyst to review the errors and inconsistencies identified by the programs, and delays caused by mail delivery.
- The analyst's time is, of course, controllable by the analyst, but the mail delay time will cause interruptions in the processing flow. The degree to which this may be a burden or irritation to the analyst, or present unacceptable or undesirable delays to the command, must be assessed by the user.

Under the proposed dual-system implementation, the Warminster computer could be used solely to extract data from the PURPLEs, and construct the final data base records. The records -- the whole data base itself -- would be on magnetic tape, and retained at COMPATWINGSLANT for use on the IBM-1401. Therefore any queries of the data base could be run overnight, or sooner if required, and CPWL would have complete control over access to the data base.

The use of the 1401 is suggested in several of the steps, assuming that those particular functions can in fact be implemented on the smaller computer. The functions are:

- The use of part of PCHANG to pre-check punched change cards before they are sent to NADC (Step 5);
- Copying of the change card deck to tape, for ease of transmission (also Step 5);
- Merging of the newly created data records with the existing data base (Step 9); and
- Executing data base queries on the full data base (Step 10).

The first three are not difficult to implement, with no concern over the limited size of the IBM-1401. Implementation of DELRAP, the query processor program, is not completely certain, and would require some testing. Sizing estimates are difficult to make, as the 1401 storage process uses different sizes of "words" for characters, integer numbers and real numbers. Preliminary sizing indicates DELRAP should fit, possibly with some scaling down of array sizes; if not, the alternative step of breaking the program into modules is available. Sizing tests of the program on the 1401 itself would be one of the first steps of any Dual-System implementation.

RECOMMENDATIONS

We recommend that implementation of the prototype VPDBS be made on the Dual-System basis described above. The implementation should be performed

in two phases, with commencement of Phase two to be contingent upon successful completion of Phase one.

Phase one tasks will ensure that all foreseeable conflicts of usage and procedure at Warminster are resolved, the Brunswick IBM-1401 can be used for data base querying, and that any fundamental logic or structural changes are identified and resolved.

Phase two tasks will include actual implementation of the individual programs, system documentation, and test runs by COMPATWINGSLANT personnel.

IMPLEMENTATION TASKING

Phase one is to be structured into three tasks as follows:

1-A. Contractor and CPWL representatives discuss proposed implementation with NADC administrative and technical personnel, to ensure implementation is acceptable to NADC and that computer and administrative (primarily mail processing) functions are well-defined and approved in advance.

1-B. Contractor makes preliminary tests of program DELRAP to confirm that DELRAP can be implemented on IBM-1401. Any doubts about implementation success are to be absolutely resolved at this point; this is a go/no-go milestone.

1-C. Contractor and CPWL representatives mutually review capabilities and processing logic of program PURPLE and DELRAP, as well as the format of the data base record, to ensure that the programs and records adequately reflect then-current Performance Indicator definitions, and RAINFORM line formats, as well as any changes or modifications desired by COMPATWINGSLANT.

Assuming successful conclusion of all Phase one tasks, and authorization from COMPATWINGSLANT, Phase two will consist of the following six tasks, with work proceeding on several of the tasks concurrently and independently.

2-A. Make all programming changes necessary to program PURPLE and DELRAP, implementing modifications determined in Task 1-C.

2-B. Implement DELRAP and PCHANG subset on Brunswick IBM-1401.

2-C. Implement PURPLE on Warminster CDC-6600.

2-D. Implement DCONVERT and PCHANG on 6600.

2-E. Produce complete set of draft system documentation, consisting of both a technical portion (program listings and programming guide) and an operation portion (user's manual).

2-F. Deliver system to COMPATWINGSLANT and observe complete run of system -- from new data input to data base query -- by COMPATWINGSLANT personnel. Upon successful run, update documentation to reflect any discrepancies encountered in Task 2-E.

III. THE ADVANCED SYSTEM

The original feasibility study for the VPDBS surveyed the data sources and Measures of Effectiveness (MOEs) necessary for a more complete data base system, one which would permit detailed analysis across the whole spectrum of VP activity. The prototype VPDBS is only a small subset of such a complete system, which will be referred to as the Advanced VPDBS, or VPDBS(A).

VPDBS(A) DATA SOURCES

The data sources for a VPDBS(A) are the post-mission PURPLE messages, the in-flight computer tapes (for the P3C), and the post-flight analysis data developed by the TSC. In all, for a "typical" mission, this includes an enormous amount of data, especially for the P3C with its in-flight computer tape. The PURPLE messages are available by a number of means: by message (on paper copy), on magnetic tape from CINCLANTFLT on a periodic basis, through the ASWCCCS, or from individual TSC files. The P3C aircraft tapes are presently retained on site for at least several months, and periodic copying and retention for a data base could be arranged. The MPE data is available in the same manner, and the TSC data base information is available at the TSC, from the system itself for current data and from dump tapes for older missions.

Except for the PURPLE messages, the other data would be somewhat difficult to collect on a regular basis, as making it available to COMPATWINGSLANT or some other centralized user would require a not inconsiderable amount of care and effort on the part of personnel at each of the TSCs. A number of reels of magnetic tape would have to be copied and mailed, with some sort of master list and check-off procedure to ensure all the appropriate data for each mission during the collection period (probably one month) was forwarded in its entirety. Similar procedures would also be necessary at the receiving site.

PRE-PROCESSING REQUIREMENTS

The majority of data collected and used by the TSCs is for the purposes of individual mission replay and analysis. The volume of data available for each

mission is enormous, especially for P3C missions. A data base to retain all or most of this data, for all TSCs and for several years, is patently out of the question. The raw source data must therefore be carefully screened to select and retain only that smaller portion which could realistically be stored, and which has a high probability of being useful for future analysis. This process of excision and compression is most efficiently done at the receiving site, rather than at each of the TSCs. This ensures that standardized compression techniques and procedures are utilized, rather than a number of not-quite-identical processes. However, the receiving site user is not as intimately familiar with each data set as the originating TSC user would generally be, so that lost, distorted or excess data would be almost impossible to prevent and to detect.

In order to develop a procedure for data selection and extraction from the raw data, data must be selected or excised using two separate criteria: overall suitability and specific suitability. In the first instance, data selection/excision would be done based on the type or content of data -- e.g., accept aircraft position data items, but reject environmental items. Therefore all environmental data would be excluded as a type or category, while the aircraft position type would be retained. Within those retained data types, some will be further excised or selected based on more specific criteria, such as time. Therefore we might exclude all aircraft position data which is recorded prior to on-station time and subsequent to off-station time. Additionally, only one item of aircraft position might be retained every minute except during critical maneuvers, when four or five per minute might be necessary.

The pre-processing requirements for raw data tapes must be primarily automated on computers, due to the sheer volume of data involved. Most likely, a completely automated step would first perform the overall or generic excision, with a second semi-automated step to perform the specific excision/selection process. A comprehensive and fairly sophisticated logic system must also be included for the entire pre-processing stage to test for data completeness, format and context correctness, and consistency among individual data elements.

The programs and procedures to satisfactorily implement the pre-processing functions can be seen to require substantial expertise and effort, in both the design and development stages.

VPDBS(A) CAPABILITIES

Capabilities of any Advanced VPDBS are primarily a function of the spectrum of data available, and the abilities of available programs or software to effect interactions and relations between the individual data elements. User-developed programs, whatever their degree of complexity, are generally fairly specific in that they are designed to perform a single function. The prototype VPDBS may be considered an upper-end user program, as it is primarily oriented to performing the calculation of the set of pre-specified Performance Indicators (PIs). Considerable capabilities are available to select data records based on a variety of factors, but the end product is a set of PIs.

A number of extensions may be made to the prototype system, to extract and retain additional data in the data records, and to add additional pre-programmed PI-type calculations or user-selectable math operations in the query processor. However, data manipulation capabilities beyond that will almost certainly require a generalized data base management system (DBMS).

Truly generalized DBMS software is generally offered either by the manufacturer of commercial mainframes or by independent software firms. Military computer systems generally have one of three options:

- no DBMS;
- heavily-tailored, non-general ones; or
- full-scale DBMSs developed at enormous cost, such as the WWMCCS "WDMS" software.

VPDBS(A) SITE ALTERNATIVES

The site selection criteria used for the prototype VPDBS apply to the VPDBS(A)

evaluation, except for the addition of the data base management system and higher cut-off requirements on size and availability. The IBM-1401 must be immediately ruled out on both size and DBMS requirements. The CP-901 at the TSC should be acceptable on the size criteria, but the requirements of time for both input processing and data base use would conflict severely with the TSC's operational performance. Additionally, the data management capabilities designed into the CP-901 are heavily predicated on manipulating a single mission/event at one time; and while that amount of data may be quite large, the system is not designed for extracting and manipulating data across hundreds of missions rather than within a single mission.

The Warminster CDC-6600 is a large-scale, general purpose computer, which has a commercial data base management system called System-2000. This DBMS has not been specifically evaluated for its fit with likely VPDBS(A) requirements, but the system is large enough and general enough to handle virtually any such demands as an advanced system would make on it.

Based on size, DBMS and availability criteria, the Warminster computer is the only possible current site for a VPDBS(A). The problem of availability is the same as for the prototype VPDBS, only in the case of the VPDBS(A) it would be much more severe in that much of the probable use of an advanced system would be more speculative and interactive than the VPDBS, and all processing/accessing would have to be done at Warminster rather than on the Brunswick 1401. As with computer selection for the prototype VPDBS, installation of a new computer at or near Brunswick has not been specifically considered.

VPDBS(A) RECOMMENDATIONS

No matter what the size or scale of any VPDBS(A), the potential value of the system drops off precipitously if the COMPATWINGSLANT analysts do not have direct control over at least the final data base, either by being located in close proximity or by direct-access terminals. Thus site location and access is a major consideration and, unless other operational considerations override, must be

considered a major problem for a VPDBS(A). The TSC's CP-901 is the only viable local system available to COMPATWINGSLANT, and DBMS and availability factors definitely preclude it from consideration.

Especially given the lack of availability of any accessible systems, there are a number of other factors which mitigate against the development of an advanced system significantly more sophisticated and complex than the prototype VPDBS:

- the previously mentioned burden of data collection on the individual TSCs;
- the extensive amount of data selection/rejection necessary to reduce the raw data to manageable proportions;
- the difficulty of selecting the data to be excised or retained, without ruling out too many areas of potentially fruitful analysis; and
- the almost impossible task of continually modifying the data base structure and the pre-processing programs to keep up with updates of the P3C, MPE and TSC software.

Additionally, while it is difficult to quantify and define, the maintenance of any non-trivial data base imposes a considerable burden of effort on the custodian of the data base. The VPDBS requirements should be minimal in this respect, in that its small size and rather restricted purpose should not make it subject to many maintenance demands. But a large-scale data base would require substantial maintenance efforts which, without extra personnel specifically assigned for the task, would not be cost-effective for COMPATWINGSLANT.

If the prototype VPDBS is implemented for COMPATWINGSLANT use, the additional (incremental) advantages of an Advanced System would be difficult to justify in view of the incremental labor and effort required.

It is therefore recommended that consideration of a sophisticated Advanced VP Data Base System for COMPATWINGSLANT not be pursued for the present.

Instead, consideration of modest enhancements to the existing VPDBS, such as the utility of user-selectable math operations, should be made. In the event that COMPATWINGSLANT subsequently gets secure access to a general purpose computer (such as WWMCCS), with technical support at the computer site itself, the question of an Advanced System may be re-opened.

For interim analytical requirements beyond the capabilities of the existing VPDBS, essentially one-time analysis projects may be undertaken by collecting samples of the P3C/TSC data, without the intention of retaining the data or continuing the effort beyond the initial analysis. Such one-time efforts could be used to resolve questions of unaccountably poor performance in specific areas, or to allow detailed evaluations of performance variations between TSCs, squadrons, months, or other selected elements.

OTHER ALTERNATIVES

The negative recommendations for a COMPATWINGSLANT VPDBS(A) are predicated on such a system being for the use of COMPATWINGSLANT use alone, and having to use only the facilities and resources available to COMPATWINGSLANT. A system for joint or shared use by WINGSLANT and WINGSPAC would spread both development and operating costs between two users, and make any applications much more cost-effective. Additionally, COMPATWINGSPAC does presently have access to several secure systems, one on site and one by secure line to NOSC. Locating a system at CPWP/NOSC would allow direct secure access for data preparation and data base construction, and data base accessing from WINGSPAC, although there would still be no way to provide secure access for WINGSLANT users. This alternative would have to be explored directly between LANT and PAC representatives.

One additional factor should be considered, although it is not an immediate alternative. The S3 community has somewhat similar requirements for analysis of air ASW performance, and a separate review of data processing needs and solutions is presently being conducted under the cognizance of MASWSP/PM-4. While the input data and the eventual analysis needs are certainly not identical, substantial economies and increased capabilities should be available by combining the needs of the LANT and PAC P3 and S3 communities.