MITRE PAPER

TEASG Step 3 Report on

APB Step 3 Test, Evaluation, and Analysis Process

April 2000

Michael Beasley, Digital Systems Resources David Colella, The MITRE Corporation, Chair Ronald Fico, Johns Hopkins University, Applied Physics Laboratory W. Robert Lane, Naval Undersea Warfare Center Gregory C. Rice, Johns Hopkins University, Applied Physics Laboratory Michael K. Seil, Naval Undersea Warfare Center

Sponsor: Dept. No.: ASTO W097 Contract No.: Project No.: DAAB07-99-C-C201 0700V190AA

The views, opinions and/or findings contained in this report are those of the MITRE Corporation and should not be construed as an official Government position, policy, or decision, unless designated by other documentation.

© 2000 The MITRE Corporation



EXECUTIVE SUMMARY

A subcommittee of the Test, Evaluation, and Assessment Support Group was formed to address issues and procedures relating to the Advanced Processor Build Step 3 test and analysis procedures. This examination was prompted by the recent APB-99 Step 4 test wherein a number of system deficiencies were noted with little or no perceived forewarning from the APB-99 Step 3 evaluation. This report discusses the subcommittee's examination of these issues. This critical examination led to an assessment of the overall Step 3 process and its implementation. The report therefore provides conclusions regarding the implementation of the APB Step 3 mechanism and recommendations for the modification of the Step 3 procedures so that more effective functional and integrated system testing can be attained. The failure to adequately address the Step 3 procedural issues raised here can increase performance risk associated with each new generation APB sonar system build.

A majority of the system deficiencies highlighted during the APB-99 Step 4 sea test were either observed during prior lab tests or could not be tested during Step 3 lab tests due to a lack of testing capabilities. Nonetheless, the appearance of these problems was not sufficiently disseminated to the general community in a timely fashion. There were a number of reasons for this. Most significant was the fact that the primary focus of APB-99 testing was to enable a stable and fully functional system in preparation of the Step 4 sea test. This distracted the test team from assessing functional performance. This was partly a result of a large portion of the novel APB-99 functionality being too immature or inadequately tested prior to being incorporated into the APB system.

More importantly, examination of the APB-99 deficiencies highlighted shortcomings of the overall Step 3 procedures. The APB-99 system suite continually underwent change throughout the Step 3 test period, including modifications, functional changes, and algorithm tuning. Such activity during testing seriously hinders the ability to adequately examine functional and overall system performance. Recommendations are presented to help correct some of the shortcomings of the testing implementation. These include minimizing system modifications and retuning during the Step 3 test, increasing communication between the testing groups and other support groups and developers, providing a more focused approach to Step 3 testing, and, finally, extending the Step 3 test period to provide a find, fix, repair, and retest phase that examines the impact of post-test modifications.

1. INTRODUCTION

This report contains the conclusions and recommendations of the Step 3 Subcommittee of the Test, Evaluation, and Assessment Support Group (TEASG) that was formed to address issues relating to the Advanced Processor Build (APB) Step 3 test and analysis process. Our examination was prompted by two factors: (1) concerns raised following the APB-99 Step 4 sea tests regarding the effectiveness of the Step 3 process to identify system performance limitations, and (2) the need to assess the impact of the Step 3 test constraints created by the current APB test cycle, limited resource availability, and test focus.

The overall goal of our inquiry was to develop a more effective test procedure to facilitate and improve the transition of developmental signal and data processing algorithms into integrated APB system functionality. Our examination of the APB-99 process did not address the question of why any of the noted deficiencies appeared in the APB system nor did we consider ongoing attempts to correct the problems. Instead, the initial focus was on whether or not the Step 3 process effectively identified and reported these issues. The follow-up assessment of the overall APB Step 3 process was then aimed at improving the Step 3 process to more effectively evaluate system performance, identify system characteristics that are likely to degrade performance, and report problem issues to the community in a timely fashion for effective remediation.

In the next section, we present a list of APB-99 functional deficiencies that were specifically cited during the APB-99 Step 4 at-sea test as areas of major concern. The particular deficiencies are summarized in Table 1. In each case, we summarize the activities from Step 3 associated with the problem to assess the effectiveness for finding and repairing such problems prior to Step 4 testing. Conclusions drawn from an examination of these deficiencies are then given. In Section 3, we provide the subcommittee's conclusions from an assessment of the APB Step 3 in general. These conclusions are broader in scope and involve issues associated with Step 1 and Step 2 activities as well as the evolution of functional components through the APB process. Section 4 provides recommendations that the subcommittee feels, if implemented, would improve Step 3 testing. The recommendations are limited to those fitting within the current construct of the APB process without requiring a wholesale modification of the APB process. In Section 5, we offer concluding remarks.

2. APB-99 TEST AND EVALUATION

Step 4 Issues. Following the APB-99 Step 4 at-sea test, a number of significant deficiencies were identified in the APB-99 system. These deficiencies are listed in Table 1 as inkblot, washout, range discrimination, PNB tracker performance, full spectrum normalization, sasquash, array shape estimation, and barber pole effects. The question then posed to the TEASG was whether these APB-99 deficiencies had been identified during the Step 3 phase prior to Step 4 testing. In this section that question is addressed by providing an item by item summary of when the deficiencies were recorded and the extent to which they were reported to the community. We also begin to examine the effectiveness of the Step 3 process in recording and reporting system problems.

Table 1 itemizes the eight deficiencies that were identified during the APB-99 Step 4 tests. For each case, the table includes a brief summary of the results of the Step 3 Subcommittee assessment as related to each deficiency. Evidence suggests that many of the performance issues

Item	Step 4 Test	Remarks
Inkblot	Severe inkblot effects are present and occur often.	<i>Observed during Step 3 testing</i> . Effects noted and reported to Joint Test Group; issue discussed but not disseminated. ^{2,5}
Washout	Noted during specific range and bearing cases.	Not observed for APB-99 Step 3 testing. This problem was known prior to APB-99 testing as it was observed in the APB-98 TB-23 data but was not acknowledged as a major issue. ^{1,2,5}
Range	Inability to identify proper range	Noted during Step 3 testing. Related
Discrimination	cell in PBB display.	problems with OR-ing effects, which drive the PBB displays, were prominent during testing and noted in reports. ^{2,5}
PNB Tracker Performance	PNB trackers provided unreliable initialization and often were unable to hold track.	<i>Some problems noted during Step 3</i> <i>testing.</i> Problems with initialization, tracking high dynamic targets, and excessive random errors in range estimates noted and reported. ^{2,4,5,6,7,8}
FS Normalization	Invocation of full spectrum normalization processing was inoperable.	<i>Problems noted during Step 3 testing.</i> General problems with the FSN were noted during in-lab testing, reported, and modifications made; normalization problems reappeared during Step 4 when a software switch prevented FSN from turning on. ^{2,3,5}
Sasquash Effect	Identified at sea by sonar operators.	<i>Not observed during Step 3 testing.</i> This problem was first noted during the Step 4 at-sea tests.
ASE Auto Sensor Selection	Improper ASE sensor data noted at sea from comparison of port/starboard display tracks and variation of tracker range estimates.	Not addressable during Step 3 testing. ASE sensor reliability could not be tested during lab tests since there was no capability to replay recorded ownship navigation data back through the system in the lab.
Barber Pole Effect	Noted during at-sea testing on towed array data.	Not identified during Step 3 testing. The barber pole effect had been observed in data from other systems and arrays and is not unique to APB processing; this effect is not yet well understood and at present is not considered a high priority problem.

¹APB-98 Step 3 Final Report. ²APB-99 Step 3 JTG status briefings and meeting minutes. ³PTR report number APB99_156. ⁴APB-99 Operator Survey results briefing. ⁵APB-99 NUWC status reports to ASTO and PMS-425. ⁶PTR report numbers APB99_237, APB99_297, APB99_189. ⁷APB-99 Tracker Performance Study briefing presented to SDWG. ⁸APB-99 Tracker Performance Study Summary briefing presented to JTG.

Table 1. APB-99 Step 4 Issues.

were noted prior to Step 4 testing. This includes issues that remained unresolved from APB-98 testing. It should be noted that occasionally (e.g., inkblot, washout), the currently used label was not associated with the particular problem until the APB-99 Step 4 tests. There were several cases where the problem was not reported to the community in a timely and effective manner. There were also cases where insufficient in-lab testing capabilities prevented the particular problem from being examined during Step 3 tests.

Step 3 Findings. We now discuss the issues listed in Table 1. This discussion presents observations from the subcommittee examination of Step 3 testing relative to the Step 4 deficiencies and provides comments germane to Step 3 efficacies. We reiterate that the focus of this discussion is on the ability of the Step 3 process to identify and report the particular problems. It addresses neither the technical reasons for the appearance of the reported problems nor the effort made to correct them.

- *Inkblot*. The inkblot effect occurs when significant energy is spread across many of the PBB beams in the SPED processor, creating an "inkblot" appearance in the PBB display. The appearance of these inkblot effects was identified during APB-99 Step 3 tests, although the nomenclature "inkblot" was not adopted until the recent sea tests. Significant degradation with Energy Detect SPED (ED) processing was observed during Step 3 testing when compared to APB SPED Energy Detect Clutter Suppress (EDCS) and A-RCI processing. General problems with SPED were noted and testing was halted during the week of 30 June 1999 to implement modifications. When testing resumed the following week, improved SPED performance was observed; however, it was then noted that a re-tuning of the SPED processor was required.
- *Washout*. The PBB washout effect occurs when traces in the PBB display become washed out and appear as white stripes (indicating reduced energy) as opposed to darkened tracks (indicating excessive energy). This effect did not appear in the Step 3 testing for APB-99 because the real-world data sets available for TB-29 testing at that time did not generate the effects of the problem for the processor. However, it should also be noted that although the washout effect can be reproduced in the lab for the APB-99 system using synthetic data, scenarios that would generate this effect were not included in the Step 3 test regimen. This was due at least in part to the fact that not enough significance was attached to the washout problem prior to APB-99 testing. In retrospect, it was later found that washout effects can be seen in display grams from APB-98 testing for the TB-23. The washout problem has since been identified in the Step 3 and Step 4 final reports for APB-98, although the Step 4 report is only now being disseminated.
- *Range Discrimination.* The range discrimination problem occurs when a target of interest appears in multiple range fields in the PBB displays with nearly equal energy levels, making it difficult for an operator to accurately estimate target ranges from the visual display. The inability to identify target range from the PBB displays was noted during Step 3 tests. This and other problems believed associated with OR-ing losses incurred by the PBB display processing were flagged as important issue areas.
- *PNB Tracker Performance*. Tracker performance issues primarily concern the initiation of the PNB tracker function and the ability of a tracker, once assigned, to hold track. These

problems are particularly manifested for medium and high target dynamics scenarios. Problems with the PNB tracker function were observed throughout APB-99 Step 3 testing and noted in reports and briefings. One particular recurrent issue was the excessive random error noted in tracker range estimates. Several Problem Trouble Reports (PTRs) that addressed tracker function issues were opened. Tracker initialization was included on the list of necessary system fixes that led to a halt of testing during the week of 30 June 1999. Tracker problems were also identified in operator evaluations on the APB-99 Operator Questionnaire following Step 3 testing. Finally, warnings regarding degraded tracker function performance were presented prior to Step 4 testing from the Modeling and Prediction Support Group.

- *FS Normalization*. Issues regarding full spectrum normalization (FSN) concern the proper implementation and invocation of the FSN process as visually noted on the PNB displays. Problems with FSN processing were noted and reported during Step 3 at the JTG meetings and in PTR APB99_156. These problems were immediately addressed and presumed fixed when they did not reappear in subsequent testing. Later, during Step 4 tests, it was noted that operators were unable to invoke the FSN. Post sea-test examination showed that a software error prevented the FSN from turning on during the sea tests. The precise relationship between Step 3 in-lab failures of the FSN and subsequent failures during the Step 4 sea tests is unknown and is likely to remain unresolved.
- *Sasquash Problem.* The sasquash effect is the observance of traces that alternately appear and disappear on both left and right PBB display during assumed straight tow operations. This problem was first noted at sea during Step 4 testing and not identified earlier. It is now believed to arise from slight deformations in the towed array under conditions for which there is no apparent change in ownship heading.
- *Array Shape Estimation*. Array shape estimation issues deal with retrieving accurate estimates of the shape of the towed array. Array shape estimation problems due to the self-monitoring of the heading sensor elements were observed during Step 4. These effects were not noted during Step 3 testing since the array shape estimation module could not be effectively analyzed as the available playback system did not support ownship navigational data playback. Future tests, including APB-00 Step 3 testing, will support this feature.
- *Barber Pole Effect.* The barber pole effect concerns the appearance of parallel diagonal traces, much like "barber pole" stripes, on the PBB display. This effect has been identified in the output of other sensors and systems. It was not noted during Step 3 tests and since its cause is not well understood, tests were not designed to address this issue. Future analysis should identify both its source and impact on overall system performance.

Step 3 Lessons Learned. The lessons reported by the Step 3 Subcommittee regarding the effectiveness of the APB-99 Step 3 process to report the above deficiencies to the community prior to Step 4 testing are as follows.

- (2.1) The majority of the deficiencies appearing in the APB-99 Step 4 at-sea test were documented and reported from either the APB-99 Step 3 test phase or earlier APB-98 testing.
- (2.2) The APB-99 Step 3 process did not provide adequate communication between the test groups (TEASG and JTG) and the wider community and support groups (SPWG, AWG, COSG, and MPSG) before, during, and after the Step 3 phase.
- (2.3) APB-99 Step 3 testing and assessment was significantly impacted by activities endemic to the APB Step 3 process as currently implemented.

Details and specific items from which the above conclusions were derived are provided below.

- Most of the deficiencies that emerged following Step 4 at-sea testing of APB-99 (inkblot, washout, range discrimination, PNB track initialization and track loss, and FS normalization) were noted and reported during either the APB-99 Step 3 tests or previous APB testing. Two of the deficiencies (ASE, sasquash) could not be addressed during Step 3 given available in-lab playback capabilities. The barber pole effect was not observed during Step 3 testing but had been observed previously from other sensors; it is not currently considered a serious problem for APB.
- Many of the deficiencies and related problems were reported to the JTG and to the ASTO and PMS-425 program offices. Reports to the JTG were made during weekly public forums that served as a review of ongoing test activities and status. However, other support groups, including the Signal Processing Working Group (SPWG), the Automation Working Group (AWG), the Concept of Operations Support Group (COSG), the Modeling and Prediction Support Group (MPSG), and development laboratories and universities were not directly notified regarding the occurrence of the noted effects.
- In the case of washout, appropriate data sets that would have highlighted the effect for TB-29 processing were not included for testing because there was limited concern prior to Step 3 testing of the need to address this issue. This case in particular highlights the need for all working and support groups to help identify system problems prior to the Step 3 test phase.
- The original goal of the Step 3 test was to identify APB-99 performance strengths and weaknesses. However, this goal was difficult to achieve due to ongoing modifications and re-tunings implemented during the Step 3 tests. Because of this, the major focus of the Step 3 test evolved into being able to provide a functional and stable baseline system for the APB-99 Step 4 sea test. The modifications that were implemented appear to be the result of immature and inadequately-tested system components that were undergoing their first major integrated system testing. Furthermore, much of the system functionality was not available at the onset of testing. Instead, increased functionality was gradually phased in during the first few weeks of the Step 3 phase. This imposed a significant burden for an effective evaluation of system performance and presented a challenge to the operators and analysts responsible for testing a system that was constantly changing.

• The APB-99 tests suffered from a lack of baseline metrics for almost all of the TB-29 processing functionality that served as a major focus for the APB-99 build. Expected performance was extracted from an interpolation of TB-23 performance which did not provide sufficient forewarning of problems particular to TB-29 processing and the novel functionality.

3. STEP 3 PROCESS ISSUES

Throughout our examination of the APB-99 Step 3 testing, the Step 3 test team was continually faced with issues that seemed to epitomize the current implementation of the APB Step 3 process rather than issues specific to APB-99 testing. For this reason, we also undertook an extensive examination of the general Step 3 procedure. Having completed the Step 3 process twice over the past two years, we felt it was an opportune time for this process assessment. Whereas the above conclusions pertain particularly to APB-99 Step 3 testing, the conclusions provided in this section address broader-scope issues regarding Step 3 testing as they pertain to the recurring APB process and demanding schedule. We provide these conclusions in hopes of improving our current testing methodology for the Step 3, and subsequently, Step 4 processes. This assessment has helped us to further refine the goals, objectives, and expectations of the Step 3 test phase.

The following conclusions are related to issues the Step 3 Subcommittee felt represented areas for improvement and that, if ignored, could impose unnecessarily high risk for APB operability during Step 4 testing.

(3.1) Testing is typically performed on a system that is constantly being modified or retuned during testing, which has more than occasionally led to confusion regarding the overall impact of any single modification.

For example, a completely functional APB-99 system was not ready for testing on 1 June 1999, the initial test date. The system cycled through more than seven builds during the Step 3 phase; although these builds were regarded as upgrades to incorporate added functionality, often modified or re-tuned functions were incorporated as well. The modifications presented a severe strain on operator and analyst attempts to provide an accurate assessment of system performance. Furthermore, much of the new functionality had not undergone adequate testing prior to Step 3 system integration tests. Inadequate testing of system functionality can be traced to a number of causes. Some of these were the urgency to include a technology in the APB system that was too immature, delays in the presentation of final functional specifications prior to system integration, and schedule demands that did not permit adequate time for proper tuning of the function being integrated.

(3.2) There is insufficient interaction between the other support groups and the TEASG/JTG prior to and during testing to help define effective test scenarios and presage possible system strengths and weaknesses.

Greater interaction with the other support groups is required if effective testing is to be implemented. The developers and members from the other support groups that recommend a particular function for the APB system are best qualified to identify and quantify specific strengths and weakness for the given functionality. In particular, more detailed information regarding functional performance during the Step 1 and Step 2 test evaluation phases are needed. Information on how to stress a given function and what impact its processing might have on other system components are best provided by those with in-depth knowledge of the function.

(3.3) Reporting from the test group to the general community is insufficient to ensure that the community is aware of possibly significant problems in the system prior to Step 4 testing. Furthermore, a more effective method for the timely reporting of results and, especially, problem issues needs to be implemented.

Although many of the problems arising during the Step 3 testing were reported to the JTG and the program offices at ASTO and PMS-425, the problems that were encountered were not successfully raised to the community at large so that concerns over many of these issues could be addressed. Also, although at least some of the signal processing and data processing problems discussed earlier regarding APB-99 Step 4 deficiencies were considered known to the community from prior testing, adequate and timely feedback on functional problems arising during testing was not provided. The mechanism to do this, regular JTG meetings, often did not accomplish this task. One consideration here is that there is insufficient time during testing to report back to the support groups at large and that meetings of the JTG are rarely attended by others not directly involved in testing.

(3.4) The current APB Step cycle (Steps 1-4), and in particular the Step 3 and Step 4 test schedules, does not allow sufficient time to examine functional problems in the APB system and subsequent testing of modified functionality after being implemented. The APB-99 test cycle allotted no time following the APB-99 Step 3 in-lab tests for modifications to be sufficiently tested prior to their incorporation in the APB system tested at sea.

For APB-99, limited Step 2 results were available to support system evaluation and the definition of meaningful test scenarios. Lack of a sufficient baseline made it difficult to determine whether or not the integrated system performed as expected. All of the following significantly impact the effective in-lab evaluation of system performance: limited time allotted for in-lab testing; limited resource availability, including personnel, data, simulation signal generators, and hardware; and lack of significant pre- and post-Step 3 testing and re-evaluation of integrated and modified functionality.

(3.5) Operators involved in the Step 3 process are often learning a new APB system while they are testing. Overall system performance can be better evaluated if operators, watchstanders, observers, and testers are all adequately trained on the new APB functionality prior to Step 3 testing.

The ability to effectively operate and evaluate a system is directly proportional to knowledge about how that system should be used and what its strengths and limitations are. This knowledge can only be gained through adequate training prior to Step 3

testing and is needed for all active participants of the test process. In general, pre-test training on the system and OMI familiarity is inadequate.

4. **RECOMMENDATIONS**

The following recommendations are provided to improve future APB Step 3 and Step 4 testing.

(4.1) APB Step 3 testing should focus on stressing the current build to find strengths and weaknesses. The build should remain unchanged during Step 3 testing or a minimum set of modifications and tuning should occur once testing has begun.

Changes to the APB Step 3 baseline should be limited to fixing problems that are required in order to continue testing and modifications necessitating evaluation prior to the Step 4 sea test. This recommendation requires more extensive testing of integrated system functionality prior to the Step 3 process and demands that Performance Verification Tests (PVTs) be completed prior to Step 3. The software build agent with government oversight and TEASG observation should execute the PVT tests. TEASG observation is desired because it provides the opportunity to become familiar with integrated system functionality prior to the Step 3 tests. The recommendation is based on focusing attention on the ability to examine system functionality as opposed to system stability, a primary concern during the APB-99 test cycle. This requires developers to provide the system integrator with sufficient lead-time to implement and test the given functionality as part of the integrated system for stability and tuning prior to Step 3.

(4.2) Direct interaction between the TEASG/JTG and the other support groups should increase, especially prior to and during the Step 3 test phase.

Step 2 test results and expected performance summaries for each recommended function should be provided to the TEASG for consideration. The TEASG should directly solicit participation from the other support groups in order to define effective test scenarios and data sets and assist in providing performance evaluations during Step 3 testing. Specific areas of participation include: the AWG should provide a summary of strengths and weaknesses for each function recommended for inclusion in APB from the Step 2 process; the SPWG should provide a summary of strengths and weaknesses for each function or modification that it recommends and expected overall system impact for recommendations (e.g., such as band selection); the COSG should provide a prioritized list of OMI objectives for novel display features as well as a draft concept of operation that should be utilized by the test operators; the MPSG should provide appropriate functional performance bounds and parameter constraints for test scenarios. The interaction should increase communication between the working groups and the test group without sacrificing the objectivity provided by the test group for independent assessment of system performance. Furthermore, more timely and detailed reports of test results during Step 3 testing should be provided by the JTG to appropriate representatives of the working groups for consideration. It is recommended that the JTG provide daily reports on the status of testing and, when available,

initial results to the general community and in particular to the working group chairs and technical leads.

(4.3) The Step 3 testing procedure should utilize a more focused approach to testing that would allow more direct comparison of a given function across various data sets and more rapid availability of functionality tests.

This can be effected by the use of teams of operators and analysts that focus on a given functionality, improved operator/analyst reporting through immediate survey of impressions of system performance, and immediate and dedicated next-day analysis of test results. The effective use of sonar technicians and operators of various levels of expertise should continue. Also, the TEASG should a utilize a test approach that focuses on specific functionalities rather than overall system performance whenever constraints in time, personnel, and/or hardware significantly restrict full system testing.

(4.4) *Efforts should be made to improve testing capabilities in order that more effective tests can be implemented.*

There are several areas where improvements can be achieved that would directly impact testing effectiveness. Some of these improvements are: a more robust simulator for signal generation and test data (PNB dynamic targets, swath-like signal generation, multi-target and multipath signal generation) and array shape modeling. Furthermore, the development of an interface for ownship course, speed, and depth data would facilitate testing of functionality utilizing these data. Finally, suitable data sets should be identified and secured to provide more appropriate real-world stresses of system function from lab testing.

(4.5) Adequate time should be provided for identifying and correcting observed problems in the APB system. This should entail an extended Step 3 process whereby a find, fix, repair, and retest phase is added following the current Step 3 tests to ensure that sufficient testing of modifications (as part of the integrated system) can be completed.

The failure to allow for adequate re-testing of modified functionality can seriously jeopardize system operability and performance. Attempts to provide quick fixes to system failures without sufficient follow-up testing can often be shortsighted and introduce new unanticipated problems. Such activity should therefore be avoided. Furthermore, regression testing at dockside prior to the sea test should be performed to ensure the system is performing as indicated during lab tests.

5. CONCLUDING REMARKS

The conclusions and recommendations contained within this report are aimed at providing an improved APB test process. Of particular importance is the goal of reducing the accepted risk for at-sea system failure or degraded performance. In order to do this, more effective testing of the APB system during the Step 3 test phase must be implemented. This requires more effective test procedures and utilization of test personnel. It will also entail the incorporation of more mature functional components at the Step 3 test phase. Furthermore, an extended Step 3 process is desirable so that post-test modifications can be adequately assessed and impact on overall system performance evaluated. Given current constraints in the testing schedule and the limited resources currently available, serious consideration must be given to reducing the magnitude of the functional changes incorporated from one build to the next. Finally, we feel that failure to address the issues raised in this report would increase performance risk for fielded APB sonar systems.